ESTYLF 2012

XVI CONGRESO ESPAÑOL SOBRE TECNOLOGÍAS Y LÓGICA FUZZY

Actas

Valladolid, 1- 3 de febrero de 2012

Palacio de Congresos *Conde Ansúrez*

Portada: *Claustro del Palacio de Santa Cruz*,
Fotografía de Jesús Ángel Hernández de Rojas
Licencia Creative Commons

Actas del
XVI Congreso Español sobre Tecnologías y Lógica Fuzzy

Valladolid, 1-3 de febrero de 2012

ORGANIZADORES:

Universidad de Valladolid
European Society for Fuzzy Logic and Techonology

EDITORES:

Gregorio I. Sainz Palmero
Jesús Alcalá Fernández
Rubén García Pajares
Bonifacio Llamazares
María Jesús de la Fuente Aparicio

ENTIDADES COLABORADORAS:

Ministerio de Ciencia e Innovación. Gobierno de España.
Universidad de Valladolid.
VIAS y CONSTRUCCIONES, S. A.
Fundación CARTIF.

ISBN: 978-84-615-6653-2

Presentación

Nos es grato presentar este volumen electrónico donde se recogen las comunicaciones de los 122 trabajos aceptados de la XVI Edición del Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF 2012) celebrado en Valladolid entre los días 1 y 3 de Febrero de 2012. Se trata del mayor número de contribuciones en la historia de los congresos ESTYLF, lo cual no es una sorpresa sino la confirmación de la vitalidad y trabajo creciente de la comunidad *fuzzy* española.

Esperamos que durante las sesiones se produzca el habitual y deseable intercambio de ideas, comentarios y sugerencias que den lugar a la mejora y profundización de los trabajos presentados para que en un futuro próximo éstos puedan ser sometidos a revistas científicas de prestigio reconocido. También esperamos que de la actividad desarrollada se estrechen y renueven los vínculos existentes entre los participantes.

Con objeto de incentivar la investigación de calidad de nuestros jóvenes investigadores, se concederán dos premios (en las modalidades de aplicaciones y fundamentos/modelos) a los mejores trabajos cuyo autor principal se encuentre en fase de realización de la tesis o la haya defendido a partir del año 2011. Así mismo se concederán placas conmemorativas a los investigadores que hayan acreditado 25 años de dedicación a las tecnologías y lógica fuzzy.

Hemos de agradecer a todos aquéllos que han hecho posible que este congreso llegue a buen fin: autores, organizadores de sesiones especiales, evaluadores, miembros de los comités organizador y de programa, patrocinadores (Ministerio de Ciencia e Innovación, Universidad de Valladolid, Fundación CARTIF y VIAS y CONSTRUCCIONES S.A.), participantes en el *Foro Colaboración Técnico Científica Empresa - Universidad* y todas aquellas personas que de forma desinteresada han ayudado a la organización del congreso.

Esperamos que todos los participantes se encuentren a gusto en Valladolid, disfruten de la hospitalidad castellana, de su arte, vinos y gastronomía.

José Luis García Lapresta

Presidente del Comité Organizador
de ESTYLF 2012

Gregorio I. Sainz Palmero
Francisco Herrera Triguero

Presidentes del Comité de Programa
de ESTYLF 2012

Comité Organizador

Presidente de Honor: Enric Trillas

Presidente: Jose Luis García Lapresta

Jesús Alcalá (UGr) Sesiones invitadas
José Candau Pérez (UVa)
Mª Jesús de la Fuente Aparicio (UVa)
Marta Galende Hernández (CARTIF)
Rubén García Pajares (CARTIF)
Ángel M. Gento Municio (UVa)
Bonifacio Llamazares Rodríguez (UVa)
Miguel Martínez Panero (UVa)
Luis Carlos Meneses Poncio (UVa)
Mª Teresa Peña García (UVa)

Comité de Programa

Presidentes: Gregorio I. Sainz Palmero
Francisco Herrera Triguero

Rafael Alcalá (UGr)
Jesús Alcalá (UGr)
José M. Alonso (ECSC)
Senen Barro (USC)
Edurne Barrenechea (UPNa)
Angel Barriga Barros (IMSE-CSIC)
Iluminada Baturone (IMSE-CSIC)
José Manuel Benítez (UGr)
Alberto Bugarín (USC)
Pedro Burillo (UPNa)
Humberto Bustince (UPNa)
José Manuel Cadenas (UM)
Tomasa Calvo (UAH)
Jose Manuel Cano (UPCt)
Pablo Carmona (UEx)
Jorge Casillas (UGr)
Juan Luis Castro (UGr)
José Jesús Castro-Schez (UCLM)
Elena Castiñeira (UPM)
Oscar Cordón (ECSC)
Inés Couso (UO)
Cris Cornellis(UGent)
Susana Cubillo (UPM)
Mª José Del Jesús (UJa)
Miguel Delgado (UGr)
Francesc Esteva (IIIA-CSIC)
María Jesús de la Fuente (UVa)
Pere García (IIIA-CSIC)
Luis Garmendia(UCM)
Mª Ángeles Gil (UO)
Lluís Godo (IIIA-CSIC)
Daniel Gómez (UCM)
Antonio González (UGr)
Sergio Guadarrama (ECSC)
Julio Gutiérrez (UPM)
Enrique Herrera-Vied. (UGr)
Joan Jacas (UPC)
Luis Jiménez (UCLM)
Mª Teresa Lamata (UGr)
Vicente Liern (UV)
Enrique López (ULE)
Bonifacio Llamazares (UVa)
Luis Magdalena (ECSC)
Roque Marín (UM)
María José Martín (UGr)
Luis Martínez (UJa)
Francisco Mata (UJa)
Gaspar Mayor (UIB)
José Mª Merigó (UB)
José M. Molina (UC3m)
Javier Montero (UCM)
Susana Montes (UO)
Eduard Montseny (UPC)
Claudio Moraga (ECSC)
Francisco José Moreno (UHu)
Manuel Mucientes (USC)
Ana Palacios (UO)
Antonio Peregrín(UHu)
Manuel Ojeda (UMa)
José A. Olivas (UCLM)
José Ignacio Peláez (UMa)
Raúl Pérez (UGr)
Héctor Pomares (UGr)
Ana Pradera (URJC)
Jordi Recasens (UPC)
Adolfo R. de Soto (ULE)
Ignacio Rojas (UGr)
Santiago Sánchez (IMSE-CSIC)
Luciano Sánchez (UO)
Daniel Sánchez (ECSC)
Miguel Angel Sicilia (UAH)
Pilar Sobrevilla (UPC)
Alejandro Sobrino (USC)
Vicenç Torra (CSIC)
Joan Torrens (UIB)
Gracian Triviño (ECSC)
Enric Trillas (ECSC)
Aida Valls (URV)
José Luis Verdegay (UGr)
Amparo Vila (UGr)
Pedro Villar (UGr)

ÍNDICE

Sesión Regular

Sesión Regular: Teoría

On the study of some measures of comparison of IF-sets …….1
(Ignacio Montes, Vladimír Janis, Susana Montes)

Estudio de sucesiones de contextos L-fuzzy…….……...8
(Cristina Alcalde, Ana Burusco, Ramón Fuentes-González)

Un algoritmo rápido para obtener retículos de conceptos utilizando el lenguaje de programación Python…….14
(Juan Carlos Díaz, Jesús Medina, Rafael Rodríguez)

Cardinality of IF-Sets…….20
(Patricia Díaz, Tania Iglesias, Vladimír Janis, Susana Montes)

T-convexity for lattice-valued fuzzy sets……..……..……26
(Tania Iglesias, Ignacio Montes, Vladimir Janis, Susana Montes)

Short note on a deductive scheme in fuzzy logic …….31
(Laura Menéndez-Arduengo, Itziar García-Honrado, Enric Trillas)

Algunos modelos computacionales de razonamiento conjetural……...……37
(Itziar García-Honrado, David P. Pancho, Laura Menéndez-Arduengo)

Un método para calcular openings transitivos…….43
(Dionís Boixader, Jordi Recasens)

Autocontradicción en el conjunto de las funciones de $[\mathbf{0},\mathbf{1}]^{[0,1]}$ normales y convexas.48
(Pablo Hernández, Susana Cubillo, Carmen Torres-Blanc)

Fuzzy morphological operators in a general context……...54
(J. Elorza, R. Fuentes-González, J. Bragard, P. Burillo)

Sesión Regular: Similaridades e Indistinguibilidades

Decomposition of IV—similarities……60
(Ramón González-del-Campo, Luis Garmendia)

Isomorfismos entre Indistinguibilidades, Conjuntos de Extensionales y Aproximaciones Superiores e Inferiores……...65
(Gabriel Mattioli, Jordi Recasens)

A recursive algorithm to compute a basis of a similarity……71
(Luis Garmendia, Jordi Recasens, Adela Salvador)

Sesión Regular: Sobre Soft Computing e Incertidumbres

"Rough Sets" y "Soft Computing"……77
(Rafael Bello, José Luis Verdegay)

Una tipología conjunta de la bipolaridad y la incertidumbre borrosa……..84
(J. Tinguaro Rodríguez, Camilo A. Franco, Javier Montero)

Abstracción en lógica fuzzy……...…….90
(Adolfo R. de Soto)

Sesión Regular: Predicción

Optimización de modelos estadísticos y difusos para el análisis de series temporales mediante evolución diferencial……...96
(Christoph Bergmeir, Isaac Triguero, Francisco Velasco, José Manuel Benítez)

Series temporales y clustering difuso. Caso de estudio: centro de atención de llamadas de emergencia……102
(Rubén García, Francisco Barrientos, José M. Benítez, Gregorio Sainz-Palmero)

Neurofuzzy model of an industrial process, reducing complexity by using Principal Component Analysis…..……...108
(Juan Manuel Escaño, Carlos Bordons)

Sesión Regular: Optimización e imprecisión en ambiente de incertidumbre

Modelo difuso de preferencia-aversión……...114
(Camilo Franco, Javier Montero, J. Tinguaro Rodríguez)

Fuzzy quadratic programming solving a portfolio selection problem…….120
(Carlos Cruz, Ricardo C. Silva, José Luis Verdeday)

Decisión multicriterio borrosa de la distribución potencial de abies pinsapo boiss con tecnologías GIS……126
(Verónica Bonis Martín , Luis Garmendia Salvador, Alfonso Garmendia Salvador)

Standar goal programming with fuzzy hierarchies: a secuential approach……132
(Mariano Jiménez, Mar Arenas-Parra, Amelia Bilbao-Terol)

Sesión Regular: Implicaciones y T-normas

Implicaciones residuadas en el conjunto de números borrosos discretos……138
(J. Vicente Riera, Joan Torrens)

De Morgan triples revisited……..……144
(Francesc Esteva, Lluís Godo)

Caracterización de las implicaciones *f* y *g*-generadas de Yager……...150
(Sebastià Massanet, Joan Torrens)

Sesión Regular: Modelos de Información

Solutions of systems of fuzzy relation equations as concepts of a formal context…...156
(Juan Carlos Díaz, Jesús Medina, Rafael Rodríguez)

Possibilistic evaluation of fuzzy temporal intervals……...163
(José Enrique Pons, Antoon Bronselaer, Olga Pons, Guy De Tré)

Incorporacion de una capa ELT para el uso de etiquetas lingüísticas en un data warehouse en Pentaho……...169
(Angélica Urrutia, Javier Jirón)

Implementación de una máquina HTM……..175
(Adolfo R. de Soto, Conrado Andreu Capdevila)

Sesiones Especiales

Sesión Especial: Toma de decisiones con información difusa

Organizadores:

- Luis Martínez
- Rosa Rodríguez
- Francisco Herrera

Imprecisión de los votantes en los sistemas de votación. Una propuesta basada en múltiples términos lingüísticos…….181
(Edurne Falcó Díaz de Cerio, José Luis García Lapresta, Llorenç Roselló Saurí)

An evaluation model for company environmental practices based on interacting criteria.....187
(Rocío de Andrés Calle, Teresa González-Arteaga, Luis Martínez)

Modelling influence among individuals in group decision making problems ……193
(I.J. Pérez, F.J. Cabrerizo, S. Alonso, E. Herrera-Viedma)

Metodología owa-topsis para la evaluación de tecnologías de fabricación de células fotovoltaicas……...199
(E. Cables Pérez, M. S. Garcia-Cascales, M. T. Lamata Jiménez, J. M. Sánchez-Lozano)

Evaluación de la sostenibilidad de proyectos con lógica borrosa…….205
(Raúl Martín, Raquel Caro, Luis Garmendia, Alfonso Garmendia)

Uso de operadores OWA para modelar la actitud hacia el consenso en problemas de toma de decisión en grupo ……...211
(Iván Palomares, Jun Liu, Yang Xu, Luis Martínez)

Modelo lingüístico de toma de decisiones multicriterio con expresiones lingüísticas comparativas……..217
(Rosa M. Rodríguez, Luis Martínez, Francisco Herrera)

Application of the AHP in corporate reputation management…….……...223
(Ana María Casado Molina, José Ignacio Peláez Sánchez, Jesús Doña Fernández**)**

Toma de decisión multicriterio usando integrales de Choquet intervalo-valoradas……230
(Humberto Bustince, Javier Fernández, Aranzazu Jurio, Radko Mesiar, Anna Kolesárová, Benjamin Bedregal)

Type-1 OWA based multi-granular consensus model……...235
(Francisco Mata, Francisco Chiclana, Shang-Ming Zhou)

Método difuso para identificación de variables relevantes en prospectiva tecnológica …….....241
(Pablo Villacorta, Dagoberto Castellanos, Antonio Masegosa)

Lógica difusa en sistemas de fusión de información visual: aplicaciones, extensiones y propuestas……..247
(Juan Gómez-Romero, Jesús García, Miguel A. Patricio, Miguel A. Serrano, José M. Molina)

Impacto de las etiquetas en la interpretación de la escala de Likert……253
(Nuria Martínez, Daniel Gómez, Javier Montero)

Evaluación de un método de ponderación de atributos multivaluados en sistemas de recomendación basados en contenido …….259
(Manuel J. Barranco, Jorge Castro, Luis Martínez)

Sesión Especial: Sistemas Difusos Evolutivos

Organizadores:

- María José Gacto
- Rafael Alcalá

Aprendizaje evolutivo de un sistema basado en reglas difusas para un sistema de control de entorno por medio de puntero láser……..265
(Francisco Chávez, Francisco Fernández, Rafael Alcalá, Jesús Alcalá, Francisco Herrera)

Ajuste genético lateral de las etiquetas lingüísticas en descubrimiento de subgrupos ……..271
(C.J. Carmona, P. González, M.J. Gacto, M.J. del Jesús)

Un estudio experimental del uso de dominios con intensificaciones……..……...277
(David García, Antonio González, Enrique Leyva, Raúl Pérez)

Reconocimiento del habla utilizando conjuntos difusos intervalo-valorados ajustados con técnicas evolutivas …….283
(Juan Cerrón, Raúl Orduna, Mikel Galar, Edurne Barrenechea)

Una propuesta cooperativa entre conjuntos difusos intervalo-valorados y ajuste evolutivo para mejorar el rendimiento de árboles de decisión difusos……289
(José Antonio Sanz, Alberto Fernandez, Humberto Bustince, Francisco Herrera)

Aprendizaje evolutivo de sistemas aproximativos de tipo TSK para problemas de alta dimensionalidad…….295
(M.J. Gacto1, R. Alcalá, F. Herrera)

Algoritmo genético para decisiones inteligentes en cruces de vehículos con función de coste dinámica……301
(E. Onieva, J. Villagrá, V. Milanés, J. Godoy)

Un operador de conjunción adaptativo para modelado difuso lingüístico de problemas de alta dimensionalidad……...….307
(César Serrano, Antonio A. Márquez, Francisco A. Márquez, Antonio Peregrín)

Un sistema de clasificación basado en reglas difusas jerárquico con programación genética para problemas de clasificación altamente no balanceados……...313
(Victoria López, Alberto Fernández, María José del Jesús, Francisco Herrera)

Sesión Especial: Datos de baja calidad

Organizadores:

- Luciano Sánchez
- Ana María Palacios
- Inés Couso

Linear and nonlinear regression models based in a concept of distance for imprecise response……..319
(Concepción Roldán, Antonio Roldán, Juan Martínez-Moreno)

Extracción de reglas de asociación difusas a partir de datos de baja calidad……325
(A.M. Palacios, J. Alcalá-Fdez)

Disimilitud esperada de una variable aleatoria difusa……..331
(Laura Garrido)

Análisis espectral singular para datos de baja calidad……..336
(Luciano Sánchez, Inés Couso, Luis Junco, Ana Palacios)

Introducing a genetic fuzzy linguistic combination method for bagging fuzzy rule-based multiclassification systems…….342
(Krzysztof Trawinski, Oscar Cordón, Arnaud Quirin, Luciano Sánchez)

Comparación de métodos de evaluación en algoritmos GAP con representación de imprecisión…….348
(José R. Villar, Enrique de la Cal, Marco García-Tamargo, Javier Sedano)

Sesión Especial: Aplicaciones reales

Organizadores:

- María Jesús de la Fuente
- Marta Galende

Estimador recursivo subóptimo para el ajuste de modelos borrosos TS……..….354
(Antonio Javier Barragán Piña, Mariano José Aznar Torres, José Manuel Andújar Márquez)

Modelos neuro-difusos basados en FasArt. Aplicaciones……360
(Jose-Manuel Cano-Izquierdo, Julio Ibarrola, Miguel Almonacid)

Análisis de eficiencia con datos inciertos. Una aplicación a la industria textil……366
(Leonor Plá, Trinidad Casasús, Juan Carlos Pérez, Vicente Liern)

Controladores borrosos para la dirección de vehículos autónomos en maniobras dentro de entornos urbanos……..372
(Joshué Pérez Rastelli, Teresa de Pedro Lucio, Matilde Santos Peñas)

Detección de incoherencias documentales mediante el uso de sistemas neuro-difusos y conocimiento experto……378
(Susana Martín, Susana San José, Gregorio I. Sáinz)

Selección de características para la conservación de vía ferroviaria mediante PCA, FA y operadores OWA, MA-OWA……384
(Rogelio Contreras, Marta Galende, María Jesús Fuente, Gregorio I. Sáinz)

Árboles de decisión no supervisados difusos: aplicación a un centro de atención de llamadas de emergencia ……...390
(Francisco Barrientos, Gregorio I. Sáinz, Rubén García)

Aplicación de técnicas borrosas en sistemas altamente no lineales……..396
(Juan Carlos Sanz Villanueva, Alberto Herreros López)

Takagi-Sugeno fuzzy modelling for control of greenhouse climate……..402
(Meriem Nachidi, Fernando Tadeo, Francisco Rodriguez, José Luis Guzmán)

Control borroso para el drenaje de aguas pluviales……..… 408
(Mercedes Ramírez Mendoza, Manuel González Valdez)

Sesión Especial: Software para Soft Computing

Organizador:

- Francisco Moreno

Diseñando datasets con datos de baja calidad para algoritmos de aprendizaje computacional……414
(José M. Cadenas, M. Carmen Garrido, Raquel Martínez)

Un API para modelado difuso aproximativo…….....420
(Mercedes Valdés-Vela, Fernando Terroso-Sáénz, Antonio Gómez-Skarmeta)

Síntesis automática de sistemas difusos mediante Xfuzzy……..……..…..426
(Santiago Sánchez Solano, María Brox Jiménez)

XFSML: un lenguaje de modelado de sistemas difusos basado en xml ……..432
(Ramón Santano Ruiz, Francisco José Moreno Velo)

Sesión Especial: Computing y Recuperación en Información en Internet

Organizadores:

- José Ángel Olivas
- María José Martín-Bautista
- Alejandro Sobrino
- Enrique Herrera

Impacto de las relaciones jerárquicas inter/intra-recursos en la similitud semántica entre conceptos del dominio biomédico……...438
(Israel Alonso Martínez, David Contreras Bárcena, Francisco P. Romero Chicharro)

Revisión sistemática de estrategias para el modelado automático de ontologías borrosas……..444
(Emilio Fdez-Viñas, Jesús Serrano-Guerrero, Francisco P. Romero, José A. Olivas, Victor H. Menendez-Dominguez)

Un sistema automático de Control de Autoridades para Bibliotecas Digitales ……..……...450
(Irene Diaz-Valenzuela, Maria J. Martin-Bautista, M. Amparo Vila)

Recuperación de Preguntas Frecuentes a través de SMS……..456
(David Pinto, Darnes Vilariño, Saúl León, Esteban Castillo, Mireya Tovar)

Mecanismo para la extracción y almacenamiento de grados borrosos a partir de las relaciones semánticas de WordNet……..462
(Francis C. Fernández-Reyes, José A. Olivas, Exiquio C. Leyva , Rogelio Lau)

Técnicas borrosas de clasificación y clustering de documentos aplicadas a la anotación semántica de contenidos web……468
(Francisco P. Romero, Jose A. Olivas, Jesús Serrano-Guerrero, Mateus Ferreira-Satler, Andres Soto)

Clasificación de documentos a partir de la definición conceptual de las categorías ……474
(Andrés Soto Villaverde, Francisco P. Romero, Jesús Serrano-Guerrero, José A. Olivas)

Modelo heurístico para la gestión de necesidades especiales de información….……721
(José M. Morales-del-Castillo , Eduardo Peis , José Andrés Delgado López)

Sesión Especial: Soft Computing en visión artificial

Organizadores:

- Jesús Chamorro
- Edurne Barrenechea
- Miguel Pagola

An image retrieval approach based on fuzzy dominant color descriptors……..480
(J. Chamorro-Martínez, J.M. Soto-Hidalgo, D. Sánchez, P. Martínez-Jiménez)

Umbralización de imágenes utilizando funciones de agrupamiento……486
(Aranzazu Jurio, Miguel Pagola, Luis Gonzalez-Jaime, Humberto Bustince, Ana Pradera)

Clasificación no supervisada y jerárquica basada en coloración en grafos: una aplicación a imágenes astronómicas……492
(Daniel Gómez, Edwin Zarrazola, Javier Montero, Javier Yáñez)

Evaluación de algoritmos de clustering difuso espaciales para el uso de vecindades 3D en el análisis de estudios tacs de alta resolución……498
(Alberto Rey, Alfonso Castro, Bernardino Arca)

Detección de bordes en imágenes intervalares……...504
(C. Lopez-Molina, B. De Baets, E. Barrenechea, H. Bustince)

Determinación de la mejor pareja t-norma-implicación en el gradiente morfológico ……510
(Manuel González-Hidalgo, Sebastià Massanet, Arnau Mir)

Sesión Especial: Nuevas técnicas para la enseñanza universitaria en el contexto de las tecnologías y la lógica difusa

Organizadores:

- Luciano Sanchez
- Ines Couso
- Itziar Garcia

Aprendizaje intuitivo de lógica fuzzy combinando teoría y práctica……...516
(David P. Pancho, Jose M. Alonso, Luis Magdalena)

Números borrosos y principio de extensión de Zadeh……...522
(Ana Belén Ramos-Guajardo, Itzíar García-Honrado)

Usos docentes del análisis exploratorio gráfico de datos borrosos ……..528
(Luciano Sánchez, Inés Couso, José Otero)

KEEL: una herramienta docente para sistemas difusos……...534
(Joaquín Derrac, Julián Luengo, Alberto Fernández, Salvador García, Jesús Alcalá-Fdez)

On fuzzy sets and their linguistic labeling – a reflexion for those teaching fuzzy set theory – ……..540
(Enric Trillas)

Sesión Especial: Funciones de agregación y sus aplicaciones

Organizadores:

- Tomasa Calvo
- Humberto Bustince
- Francisco Javier Fernánadez
- Javier Montero

Funciones de disimilaridad restringidas y funciones penalty……...545
(Humberto Bustince, Javier Fernandez, Mikel Galar, Radko Mesiar, Ana Pradera, Gleb Beliakov)

Funciones T–S lineales convexas: una generalización de la ecuación de Frank…….552
(Tomasa Calvo, Javier Martín, Gaspar Mayor)

New results on metrics aggregation…….558
(Sebastià Massanet, Oscar Valero)

Algunas propiedades de consistencia de las familias de operadores de agregación ..564
(Karina Rojas, Daniel Gómez, J.Tinguaro Rodríguez, Javier Montero)

Sobre t-conormas y uninormas migrativas ……569
(M. Mas, M. Monserrat, D. Ruiz-Aguilera, J. Torrens)

A new proof of a theorem of Ling……575
(Jesús Sols)

Una tipología conjunta de la bipolaridad y la incertidumbre borrosa……579
(J. Tinguaro Rodríguez, Camilo A. Franco, Javier Montero)

Sesión Especial: El problema de la interpretabilidad en los sistemas basados en reglas difusas

Organizadores:

- José Alonso
- Luis Magdalena

FINGRAMS: una nueva herramienta para análisis visual de sistemas fuzzy ……585
(José M. Alonso, David P. Pancho, O. Cordón, A. Quirin, L. Magdalena)

Contribución a la interpretación de sistemas Mamdani……...591
(Claudio Moraga)

Integración de la precisión en la mejora de la interpretabilidad de modelos difusos mediante colonias de hormigas ……...596
(Pablo Carmona, Juan Luis Castro, José Luis Herrero)

Sesión Especial: Soft Computing en minería, descripción y resumen de datos

Organizadores:

- María José Martín Bautista
- Nicolás Marín
- Daniel Sánchez

Sistemas de clasificación basados en reglas difusas y sistemas nítidos robustos entrenados en presencia de ruido de clase: un caso de estudio…….602
(José A. Sáez, Julián Luengo, Francisco Herrera)

Quality assessment of linguistic description of data …….608
(M. Pereira-Fariña, Luka Eciolaza, Gracian Triviño)

Linguistic description of the traffic evolution in roads …….614
(Daniel Sanchez-Valdes, Alberto Alvarez-Alvarez, Gracian Triviño)

Building linguistic summaries with f-cube factory……620
(Rita Castillo-Ortega, Nicolás Marín, Carlos Molina, Daniel Sánchez)

Sistemas de clasificación basados e reglas borrosas bipolares…….626
(J. Tinguaro Rodríguez, Begoña Vitoriano, Javier Montero)

Una propuesta para la mejora de los resultados de predicción en modelos basados en Prototipos Deformables Borrosos ……...632
(Ma. del Rosario Vázquez, Francisco P. Romero, José A. Olivas)

Selección de prototipos basada en conjuntos rugosos difusos ……..638
(Nele Verbiest, Chris Cornelis, Francisco Herrera)

Construcción de resúmenes lingüísticos informativos sobre series de datos meteorológicos: informes climáticos de temperatura ……...644
(A. Ramos-Soto, F. Díaz-Hermida, A. Bugarín)

Sesión Especial: Soft Computing en bioinformática y biología computacional

Organizadores:

- Rocio Romero
- Coral del Val
- Igor Zwir

Finding motifs in DNA sequences……650
(David Gutiérrez-Avilés, Francisco Martínez-Álvarez, Cristina Rubio-Escudero, José C. Riquelme)

Predicting gene network kinetics from *cis*-actingregulatory features ……656
(Oscar Harari, Coral del Val, Igor Zwir)

Perfiles comparativos de transcripción de especies silvestres de *solanum* en condiciones de sequía……662
(M.F. Lucca, E. Hopp, R. Romero-Zaliz)

Identificación de perfiles de promotores utilizando un enfoque multi-objetivo basado en restricciones …….668
(R. Romero-Zaliz, J. Arnedo-Fdez, I. Zwir, C. del Val)

A time series gene expression data analysis pipeline with a fuzzy clustering method to assess the human adenovirus infection process……..674
(P. Yankilevich, P. R. Barrero, I. Zwir)

A graphical time series analysis tool to assess synchrony: application on neuronal activity of Parkinson's disease ……678
(Patricio Yankilevich, Luis Riquelme, Mario Gustavo Murer)

Sesión Especial: Fuzziness and Philosophy

Organizadores:

- Alejandro Sobrino
- Martín Pereira

Fuzziness in bioethics ……...684
(Txetxu Ausín)

Algunos problemas filosóficos sobre imprecisión que permanecen en el ámbito computacional ……...689
(Santiago Fernández Lanza)

Método para el descubrimiento de conocimiento abstracto basado en proximidad y su incorporación al sistema Bousi_Prolog……...695
(Pascual Julián-Iranzo, Clemente Rubio-Manzano)

Razonamiento silogístico aproximado con cuantificadores generalizados ……701
(M. Pereira-Fariña, Juan C. Vidal-Aguiar, P. Montoto, F. Díaz-Hermida, A. Bugarín)

On the fuzziness of scientific disciplines – using the example of computer science…707
(Rudolf Seising)

¿Por qué usamos tanto lenguaje impreciso? a favor de una semántica comparativa para los predicados vagos…….715
(Alejandro Sobrino)

On the study of some measures of comparison of IF-sets

Ignacio Montes[1], **Vladimir Janiš**[2], **Susana Montes**[1]

[1]*University of Oviedo*, {imontes,montes}@uniovi.es

[2]*Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, Slovak Republic*, Vladimir.Janis@umb.sk

Abstract

This work is devoted to the study of some different measures of comparison between intuitionistic fuzzy sets. In particular, we focus on distances, IF-dissimilarities, IF-dissimilitudes and IF-divergences, a concept we have recently introduced. Here we study the possible relationships between these measures, we provide a common theoretical approach to them, and we also continue developing the notion of IF-divergence, trying to relate it with the notion of divergence for fuzzy sets.

Keywords: IF-set, IF-dissimilarity, IF-dissimilitude, IF-divergence.

1 INTRODUCTION

When we try to model real situations with the classical set theory, sometimes we are losing information, since these sets do not represent all the available information. In order to avoid such problem, Zadeh ([15]) introduced the theory of fuzzy sets. However, there are situations in which even this theory is not sufficient. In 1983 Atannasov ([1]) proposed a generalization of fuzzy sets, the theory of intuitionistic fuzzy sets (IF-sets, for short). In the following years he continued developing his idea ([2, 3]), and now it has become a commonly accepted generalization of fuzzy sets.

IF-sets allow us to assign a degree of membership and a degree of non-membership for every element on the set. Then, every IF-set has a degree of indeterminacy or uncertainty, that is one minus the sum of the degrees of membership and non-membership. In this sense we can see that every fuzzy set is in particular an IF-set. Moreover, the indeterminacy degree of a fuzzy set equals zero, and consequently fuzzy sets do not allow indeterminacy.

Several measures of comparison between fuzzy sets can be found in the literature. The more common are dissimilarities ([10]), dissimilitudes ([6]) and divergences ([11, 12]), in addition to classical distances. Other authors, like Bouchon-Meunier ([4]) tried to define a general measure of comparison between fuzzy sets, that includes the cited measures as particular cases. However, in the framework of IF-sets the literature is restricted to distances and some examples of IF-dissimilarities (see for example [14]).

In [13] we presented a first approach to a theoretical analysis of the measures of comparison of IF-sets. In particular, we introduce the notion of divergence for IF-sets (IF-divergence, for short), and we study the relationship among distances, IF-dissimilarities and IF-divergences.

In this work we continue developing this theory. We introduce the generalization of dissimilitudes for IF-sets, and we complete the relationships between the four measures. We also try to generalize the work of Bouchon-Meunier et al. ([4]), in order to define a general measure of comparison for IF-sets. We will obtain either distances, or IF-dissimilarities or IF-divergences depending on the conditions imposed to the function.

Moreover, we also study how to relate IF-divergences and divergences for fuzzy sets. In fact, we show a method to build IF-divergences from divergences and, conversely, how to build divergences from IF-divergences.

This paper is organized as follows: in Section 2 we make an overview on the basic concepts of the theory of IF-sets. Then, in Section 3 we introduce different measures of comparison between IF-sets and we study the relationships among them. In Section 4 we define a general measure of comparison of IF-sets and we study how, depending on the conditions imposed to such function, we will obtain either distances, or IF-dissimilarities or IF-divergences. We study the relationship between IF-divergences and divergences in Section 5. Finally, in Section 6 we present some concluding remarks and we provide ideas for future work.

2 INTUITIONISTIC FUZZY SETS

This section is devoted to introduce the basic notions of the theory of intuitionistic fuzzy sets. Intuitionistic fuzzy sets (see [1, 2]) appeared as a generalization of fuzzy sets (see [15]) that allows us one more freedom degree, and consequently they are very useful to model real situations.

Given a finite frame $X = \{x_1, \ldots, x_n\}$, an IF-set A is defined by:

$$A = \{(x_i, \mu_A(x_i), \nu_A(x_i)) \mid i = 1, \ldots, n\},$$

where μ_A and ν_A are functions $\mu_A, \nu_A : X \to [0,1]$ satisfying the restriction $0 \leq \mu_A(x_i) + \nu_A(x_i) \leq 1$ for $i = 1, \ldots, n$.

Then, $\mu_A(x_i)$ denotes the degree in which the element x_i belongs to A or satisfies the property defined by A; $\nu_A(x_i)$ denotes the degree in which the element x_i does not belongs to A. In this way, if $\nu_A = 1 - \mu_A$, A can be identified with a fuzzy set. Otherwise, we can define the function $\pi_A : X \to [0,1]$, given by $\pi_A(x_i) = 1 - \mu_A(x_i) - \nu_A(x_i)$ for $i = 1, \ldots, n$.

This function is called the intuitionistic fuzzy index or the hesitation index, and it expresses lack of knowledge of whether x belongs to A or not. Note that if A is a fuzzy set, then $\pi_A = 0$.

One of the possible applications of these sets is the human voting, where three possibilities are able: *yes* (that corresponds to the membership function), *no* (that corresponds to the non-membership function) and *does not apply* (that corresponds to the IF-index).

Let us denote by $IFSs(X)$ the set of all the IF-sets defined on X, and by $FS(X)$ the set of fuzzy sets on X. Consequently, $FS(X) \subseteq IFSs(X)$. In the theory of IF-sets, operations like union, intersection, complement and the inclusion relation can be defined as follows.

Definition 1. *Let us consider two IF-sets $A, B \in IFSs(X)$.*

- *The intersection of A and B, $A \cap B$, is defined by:*

$$A \cap B = \{(x_i, \mu_{A\cap B}(x_i), \nu_{A\cap B}(x_i)) \mid i = 1, \ldots, n\},$$

where

$$\mu_{A\cap B}(x_i) = \min\{\mu_A(x_i), \mu_B(x_i)\} \text{ and}$$
$$\nu_{A\cap B}(x_i) = \max\{\nu_A(x_i), \nu_B(x_i)\}.$$

- *The union of A and B, $A \cup B$, is defined by:*

$$A \cup B = \{(x_i, \mu_{A\cup B}(x_i), \nu_{A\cup B}(x_i)) \mid i = 1, \ldots, n\},$$

where

$$\mu_{A\cup B}(x_i) = \max\{\mu_A(x_i), \mu_B(x_i)\} \text{ and}$$
$$\nu_{A\cup B}(x_i) = \min\{\nu_A(x_i), \nu_B(x_i)\}.$$

- *The complement of A, A^c, is defined by:*

$$A^c = \{(x_i, \nu_A(x_i), \mu_A(x_i)) \mid i = 1, \ldots, n\}.$$

- *A is said to be included in B, $A \subseteq B$, if:*

$$\mu_A(x_i) \leq \mu_B(x_i) \text{ and } \nu_A(x_i) \geq \nu_B(x_i) \text{ for } i = 1, \ldots, n.$$

Let us recall that if A is a fuzzy set, then these operations coincide with the usual operations on fuzzy sets. Moreover, we have to add that the intersection and the union can be defined on a more general way, substituting the minimum by a t-norm and the maximum by a t-conorm (see [7] for information about T-intersection and S-union of IF-sets, and [9] for a complete survey on t-norms). However, we restrict to these basic definitions because they are the most usual in the framework of these work.

We are also going to introduce a difference operator between IF-sets (IF-difference, for short).

Definition 2. *A map $- : IFSs(X) \times IFSs(X) \to IFSs(X)$ is an IF-difference if it satisfies the following properties:*

D1 *If $A \subseteq B$, then $A - B = \emptyset$.*
D2 *If $B \subseteq B'$, then $B - A \subseteq B' - A$.*

Every function satisfying D1 and D2 is a difference operator. Nevertheless, there are other interesting properties that IF-differences may satisfy:

D3 $(A \cap C) - (B \cap C) \subseteq A - B$.
D4 $(A \cup C) - (B \cup C) \subseteq A - B$.
D5 $A - B = \emptyset \Rightarrow A \subseteq B$.

An example of IF-difference that also fulfills D3, D4 and D5 is the following:

$$A - B = \{(x_i, \mu_{A-B}(x_i), \nu_{A-B}(x_i)) \mid i = 1, \ldots, n\},$$

where:

$$\mu_{A-B}(x_i) = \max\{0, \mu_A(x_i) - \mu_B(x_i)\}.$$
$$\nu_{A-B}(x_i) = \begin{cases} 1 - \mu_{A-B} & \text{if } \nu_A(x_i) > \nu_B(x_i). \\ \min(1 + \nu_A(x_i) - \nu_B(x_i), & \\ \qquad 1 - \mu_{A-B}(x_i)) & \text{if } \nu_A(x_i) \leq \nu_B(x_i). \end{cases}$$

3 MEASURES OF COMPARISON OF IF-SETS

Several measures have been proposed in the literature for the pairwise comparison of IF-sets. In this section we are going to make an overview on the most usual ones, and we are going to study the possible relationships among them.

Given two IF-sets A and B defined on X, we can measure how much different they are by means of a distance. Let us notice that a distance is a map $d : IFSs(X) \times IFSs(X) \to \mathbb{R}$ that is symmetric, non-negative and it satisfies the identity of indiscernibles and the triangle inequality: given $A, B, C \in IFSs(X)$, it holds that

$$d(A,B) = 0 \Leftrightarrow A = B \text{ and}$$
$$d(A,C) \geq d(A,B) + d(B,C).$$

Another popular measures are IF-dissimilarities, that are functions $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$ satisfying the following properties:

IF-Diss.1 $D(A,A) = 0$ for every $A \in IFSs(X)$.
IF-Diss.2 $D(A,B) = D(B,A)$ for every $A,B \in IFSs(X)$.
IF-Diss.3 For every $A,B,C \in IFSs(X)$ such that $A \subseteq B \subseteq C$ it holds that $D(A,C) \geq \max\{D(A,B),D(B,C)\}$.

On the theory of the comparison of fuzzy sets, one of the most important measures are divergences, that were introduced by Montes et al. [11, 12]. A divergence for fuzzy sets is a map $D : FS(X) \times FS(X) \to \mathbb{R}$ that satisfies the following axioms:

Div.1 $D(A,A) = 0$ for every $A \in FS(X)$.
Div.2 $D(A,B) = D(B,A)$ for every $A,B \in FS(X)$.
Div.3 $D(A \cap C, B \cap C) \leq D(A,B)$ for every $A,B,C \in FS(X)$.
Div.4 $D(A \cup C, B \cup C) \leq D(A,B)$ for every $A,B,C \in FS(X)$.

In the case of the comparison of IF-sets, we have already defined (see [13]) the IF-divergences as functions $D_{\text{IF}} : IFSs(X) \times IFSs(X) \to \mathbb{R}$ satisfying:

IF-Diss.1 $D(A,A) = 0$ for every $A \in IFSs(X)$.
IF-Diss.2 $D(A,B) = D(B,A)$ for every $A,B \in IFSs(X)$.
IF-Div.3 $D(A \cap C, B \cap C) \leq D(A,B)$ for every $A,B,C \in IFSs(X)$.
IF-Div.4 $D(A \cup C, B \cup C) \leq D(A,B)$ for every $A,B,C \in IFSs(X)$.

In [13], the following result is given.

Proposition 3. *Every IF-divergence is also an IF-dissimilarity.*

In the same reference it has been proved that there is not a relationship of inclusion between distances and IF-dissimilarities and IF-divergences.

Couso et al. (see [6]) proposed a measure of comparison between fuzzy sets called IF-dissimilitude, defined as a function $D : FS(X) \times FS(X) \to \mathbb{R}$ satisfying the following properties:

Div.1 $D(A,A) = 0$ for every $A \in FS(X)$.
Div.2 $D(A,B) = D(B,A)$ for every $A,B \in FS(X)$.
Diss.3 If $A,B,C \in FS(X)$ such that $A \subseteq B \subseteq C$, then $D(A,C) \geq \max\{D(A,B),D(B,C)\}$.
Div.4 $D(A \cup C, B \cup C) \leq D(A,B)$ for every $A,B,C \in FS(X)$.

This measure can be easily translated to the comparison of IF-sets.

Definition 4. *An IF-dissimilitude is a function $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$ satisfying the following axioms:*

IF-Diss.1 $D(A,A) = 0$ *for every* $A \in FS(X)$.
IF-Diss.2 $D(A,B) = D(B,A)$ *for every* $A,B \in FS(X)$.
IF-Diss.3 *If* $A,B,C \in FS(X)$ *such that* $A \subseteq B \subseteq C$, *then* $D(A,C) \geq \max\{D(A,B),D(B,C)\}$.
IF-Div.4 $D(A \cup C, B \cup C) \leq D(A,B)$ *for every* $A,B,C \in IFSs(X)$.

We can see that IF-dissimilitudes are stronger than IF-dissimilarities but less restrictive than IF-divergences. In fact, we can find examples of IF-divergences that are not IF-dissimilitudes and examples of IF-dissimilitudes that are not IF-dissimilarities.

Then, we have introduced four different measures of comparison between IF-sets. In the following theorem we complete the study of the possible relationships between the different measure of comparison.

Theorem 5. *If we consider distances, IF-dissimilarities, IF-divergences and IF-dissimilitudes, it holds that:*

- *IF-divergence $\Rightarrow$ IF-dissimilitude $\Rightarrow$ IF-dissimilarity, and the converse implications do not hold in general.*
- *If the measure is also a distance, the converse implications of the previous item do not hold in general.*
- *There is not a general relationship of inclusion between distances and IF-dissimilarities, distances and IF-dissimilitudes or distances and IF-divergences.*

A summary of the possible relationships between these measures of comparison between IF-sets can be seen in Figure 1.

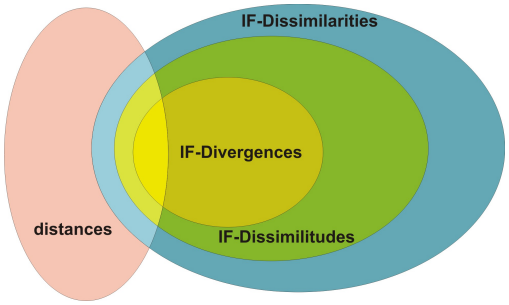


Figure 1: Relationships among distances, IF-dissimilarities, IF-divergences and IF-dissimilitudes.

As we have seen, IF-divergences are more restrictive than both IF-dissimilarities and IF-dissimilitudes. Then, the conditions imposed for IF-divergences are stronger than the ones of IF-dissimilarities and IF-dissimilitudes. Sometimes it can be an advantage, because the more restrictive the conditions, the more "robust" the measure is, (we call a measure "robust" when it is difficult to find counterintuitive

examples). In the following example we present a counterintuitive example of IF-dissimilarity that is neither an IF-dissimilitude, and consequently, nor an IF-divergence.

Example 6. *Let us consider a finite frame $X = \{x_1, \ldots, x_n\}$ and the IF-dissimilarity defined by Chen ([5]):*

$$D_C(A,B) = \frac{1}{2n}\sum_{i=1}^{n} |S_A(x_i) - S_B(x_i)|,$$

for every $A,B \in IFSs(X)$, where $S_A(x_i) = |\mu_A(x_i) - \nu_A(x_i)|$ and $S_B(x_i) = |\mu_B(x_i) - \nu_B(x_i)|$. This measure is an IF-dissimilarity, but it is neither an IF-dissimilitude nor an IF-divergence. We have to note that, whenever $S_A(x_i) = S_B(x_i)$ for every $i = 1, \ldots, n$, $D_C(A,B) = 0$. In fact, we can consider the IF-sets defined by:

$$A = \{(x_i, 0.5, 0.5) \mid i = 1, \ldots, n\}.$$
$$B = \{(x_i, 0, 0) \mid i = 1, \ldots, n\}.$$

For these sets there is $S_A(x_i) = S_B(x_i) = 0$ for $i = 1, \ldots, n$, and consequently $D_C(A,B) = 0$. However, these two IF-sets seem to be quite different.

Thus, this example shows that sometimes we must impose strong conditions to the measure of comparison in order to avoid counterintuitive situations.

4 THEORETICAL APPROACH TO THE COMPARISON OF IF-SETS

Bouchon-Meunier et al. proposed in [4] a general measure of comparison for fuzzy sets that generates some particular measures depending on the conditions imposed to such general measure.

Following the same idea, in this section we will try to define a general measure of comparison between IF-sets that, depending on the imposed properties, it generates either a distance, or an IF-dissimilarity or an IF-divergence. It will provide us sufficient conditions in order to obtain these types of functions.

For this aim, let us consider a function $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$, and assume that there is a function G_D:

$$G_D : IFSs(X) \times IFSs(X) \times IFSs(X) \to \mathbb{R}^+ \qquad (1)$$

such that D can be expressed by:

$$D(A,B) = G_D(A \cap B, B - A, A - B), \qquad (2)$$

where $-$ is a difference operator between IF-sets (see Definition 2) that fulfills D3, D4 and D5. We shall see that depending on the conditions imposed on G_D, we can obtain that D is either a distance, an IF-dissimilarity or an IF-divergence.

We begin by determining which conditions must be imposed on G_D in order to obtain a distance for IF-sets.

Proposition 7. *Let $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$ that can be expressed as in (2), where G_D is a map as in (1). If the function G_D satisfies the properties:*

- *S-Dist.1* $G_D(A,B,C) = 0$ *if and only if* $B = C = \emptyset$;
- *S-Dist.2* $G_D(A,B,C) = G_D(A,C,B)$ *for every* $A,B,C \in IFSs(X)$;
- *S-Dist.3* *For every* $A,B,C \in IFSs(X)$, $G_D(A \cap C, C - A, A - C) \leq G_D(A \cap B, B - A, A - B) + G_D(B \cap C, C - B, B - C)$;

then D is a distance for IF-sets.

Let us now consider IF-dissimilarities. We have proven the following result:

Proposition 8. *Let $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$ that can be expressed as in (2), where G_D is a function as in (1). Then D is an IF-dissimilarity if G_D satisfies the following properties:*

- *S-Diss.1* $G_D(A, \emptyset, \emptyset) = 0$ *for every* $A \in IFSs(X)$.
- *S-Dist.2* $G_D(A,B,C) = G_D(A,C,B)$ *for every* $A,B,C \in IFSs(X)$.
- *S-Diss.3* $G_D(A,B,\emptyset)$ *is non-decreasing in B.*
- *S-Diss.4* $G_D(A,B,\emptyset)$ *is non-increasing in A.*

Concerning IF-divergences, we have proven the following proposition.

Proposition 9. *Let $D : IFSs(X) \times IFSs(X) \to \mathbb{R}$ be a function that can be expressed as in (2), where G_D is a map as in (1). Then, D is an IF-divergence if G_D satisfies the following properties:*

- *S-Diss.1* $G_D(A, \emptyset, \emptyset) = 0$ *for every* $A \in IFSs(X)$.
- *S-Dist.2* $G_D(A,B,C) = G_D(A,C,B)$ *for every* $A,B,C \in IFSs(X)$.
- *S-Div.3* $G_D(A,B,C)$ *is non-decreasing in B and C.*
- *S-Div.4* $G_D(A,B,C)$ *is independent of A.*

From Theorem 5 we see that every IF-divergence is an IF-dissimilitude, and also every IF-dissimilitude is an IF-dissimilarity. Consequently, conditions of Proposition 9, that are sufficient conditions for IF-divergences, are also sufficient conditions for both IF-dissimilarities and IF-dissimilitudes. In fact, conditions of Proposition 9 imply conditions of Proposition 4, and this means that sufficient conditions for IF-divergences imply sufficient conditions for IF-dissimilarities.

We can note that we have not presented sufficient conditions for IF-dissimilitudes (different than the ones for IF-divergences). The reason is that, in this framework, it is not possible to distinguish IF-divergences and IF-dissimilitudes.

5 PROPERTIES OF IF-DIVERGENCES

In [13] we studied some properties of the IF-divergences. One of them was based on an interesting property that some

IF-divergences may satisfy, under which both axioms IF-Div.3 and IF-Div.4 are equivalent. We denote this property by IF.Div.5.

Definition 10. *If D_{IF} is an IF-divergence, it satisfies IF-Div.5 if:*

$$D_{\text{IF}}(A,B) = D_{\text{IF}}(A^c,B^c) \text{ for every } A,B \in IFSs(X).$$

This property is very useful since axioms IF-Div.3 and IF-Div.4 are equivalent when it is fulfilled.

Proposition 11. *If $D_{\text{IF}} : IFSs(X) \times IFSs(X) \to \mathbb{R}$ is a function satisfying IF-Diss.1, IF-Diss.2 and IF-Div.5, then it satisfies IF-Div.3 if and only if it satisfies IF-Div.4.*

Consequently, in order to prove that a function D_{IF} is or not an IF-divergence, if it satisfies axioms IF-Diss.1, IF-Diss.2 and IF-Div.5, it is enough to prove that is satisfies either IF-Div.3 or IF-Div.4, because both axioms are equivalent.

This property can also be defined for divergence of fuzzy sets.

Definition 12. *Let $D : FS(X) \times FS(X) \to \mathbb{R}$ be a divergence of fuzzy sets. Then D satisfies the property Div.5 if:*

$$D(A,B) = D(A^c,B^c) \text{ for every } A,B \in FS(X).$$

In [13] we presented a method to build IF-divergences from divergences. Now, we are going to extend this result, making it more general; we are going to provide a method to build divergences conversely from IF-divergences; and finally we are also going to study if these methods preserve axioms Div.5 and IF-Div.5.

First of all, we present the method to build IF-divergences from divergences, and we study if it preserves such properties.

Proposition 13. *Let D be a divergence for fuzzy sets. If $f : [0,\infty) \times [0,\infty) \to [0,\infty)$ is a function satisfying*

1. *$f(0,0) = 0$, and*
2. *$f(\cdot,t)$ and $f(t,\cdot)$ are non-decreasing,*

we know (see [13]) that the function D_{IF}, defined by:

$$D_{\text{IF}}(A,B) = f(D(\mu_A,\mu_B),D(\nu_A,\nu_B))$$

for every $A,B \in IFSs(X)$, is an IF-divergence. Then, we obtain the following results:

- *If f is symmetric, D_{IF} satisfies property IF-Div.5 (regardless if D satisfies or not axiom Div.5).*
- *If f is not symmetric, then although D satisfies axiom Div.5, D_{IF} may not satisfies axiom IF-Div.5.*

However, this result can be easily generalized in the following way:

Proposition 14. *Let D_1 and D_2 be two divergences for fuzzy sets. If $f : [0,\infty) \times [0,\infty) \to [0,\infty)$ is a function satisfying*

1. *$f(0,0) = 0$, and*
2. *$f(\cdot,t)$ and $f(t,\cdot)$ are non-decreasing,*

then the function D_{IF}, defined by:

$$D_{\text{IF}}(A,B) = f(D_1(\mu_A,\mu_B),D_2(\nu_A,\nu_B))$$

for every $A,B \in IFSs(X)$, is an IF-divergence. Moreover, if f is symmetric and $D_1 = D_2$ then D_{IF} satisfies IF-Div.5 (regardless if D satisfies Div.5 or not). Otherwise, if either f is not symmetric or $D_1 \neq D_2$, then D_{IF} may not satisfies axiom IF-Div.5.

Conversely, let us now show a method that allows us to build divergences from IF-divergences.

Proposition 15. *If D_{IF} is an IF-divergences, then the mapping D, defined by:*

$$D(A,B) = D_{\text{IF}}(A,B), \text{ for every } A,B \in FS(X),$$

is a divergence for fuzzy sets. Moreover, if D_{IF} satisfies the axiom IF-Div.5, then D satisfies the axiom Div.5.

Thus, we have seen that our method to build divergences from IF-divergences preserves properties Div.5 and IF-Div.5. However, the methods to build IF-divergences from divergences do not preserve them.

We conclude this section showing an example of the use of these results.

Example 16. *Let us consider a frame $X = \{x_1,\ldots,x_n\}$, and the normalized Hamming distance for fuzzy sets (see [14]) defined by:*

$$l(A,B) = \frac{1}{n}\sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|$$

for every $A,B \in FS(X)$. If we consider the function $f(x,y) = \frac{x+y}{2}$, it satisfies conditions of Proposition 13, and consequently we can build the following IF-divergence:

$$\begin{aligned} D_H(A,B) &= f(D(\mu_A,\mu_B),D(\nu_A,\nu_B)) \\ &= \tfrac{1}{2n}\sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| \end{aligned}$$

for every $A,B \in IFSs(X)$. In fact, this is an IF-dissimilarity (and also an IF-divergence) that were defined by Hong and King (see [8]). Moreover, the Hamming distance satisfies property Div.5 and, as f is a symmetric function, then D_H satisfies IF-Div.5.

If we now apply Proposition 15, we obtain the following divergence for fuzzy sets:

$$D(A,B) = D_{\mathrm{IF}}(A,B) = \frac{1}{2n}\sum_{i=1}^{n} 2|A(x_i) - B(x_i)| = l(A,B).$$

Consequently, we have obtained the original divergence for fuzzy sets that, as we already known, satisfies property Div.5.

6 CONCLUSIONS

In this work we have presented some contributions to the theory of the measures of comparison between IF-sets. Following the same line as in previous works, we have considered two classic measures, distances and IF-dissimilarities, the measures we have recently introduced, IF-divergences, and another possible measure, IF-dissimilitudes. Using these four measures of comparison, we have studied the possible relationships among them. In particular, we have concluded that IF-divergences are stronger than IF-dissimilitudes, and IF-dissimilitudes are stronger than IF-dissimilarities, while there is not a general relationship between them and distances.

We have also introduced a general measure of comparison of IF-sets that provides sufficient conditions to obtain either distances, or IF-dissimilarities or IF-divergences. However, these general measure does not distinguish IF-dissimilitudes from the other measures.

Finally, we continued exploring the notion of IF-divergences. In fact, we have presented a method to build IF-divergences from divergences for fuzzy sets and, conversely, a method to build divergences from IF-divergences.

With respect to this topic of research, we have found a general measure of comparison between IF-sets. It provides us sufficient conditions for obtain different measures. In this sense, we would like to find a characterization, that is, both sufficient and necessarily conditions. Moreover, in the future we hope to study the possible definition of local IF-divergences. In addition, it is also possible to look for applications of these measures of comparison in different areas like multicriteria decision making or pattern recognition.

Acknowledgements

The research in this paper is partly supported by the Science and Education Ministry FPU grant AP2009-1034, by the Agency of the Slovak Ministry of Education for the Structural Funds of the EU, under project ITMS:26220120007, by grant 1/0297/11 of the Slovak grant agency VEGA and the Spanish Ministry of Science and Innovation grant MTM2010-17844.

References

[1] K. Atanassov: Intuitionistic fuzzy sets. In: *Proceedings of VII ITKR*, Sofia (Bulgaria), 1983.

[2] K. Atanassov: Intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 20, pp. 87–96, 1986.

[3] K. Atanassov: Intuitionistic fuzzy sets past, present and future. In: *Proceedings of the 2003 EUSFLAT conference*, Zittau, 2003.

[4] B. Bouchon-Munier, M. Rifqi, S. Bothorel: Towards general measures of comparison of objects. *Fuzzy Sets and Systems* 84, pp. 143–153, 1996.

[5] S. M. Chen: Measures of similarity between vague sets. *Fuzzy Sets and Systems* 74(2), pp. 217–223, 1995.

[6] I. Couso, L. Garrido, S. Montes, L. Sánchez: Expected pairwise comparison of the outcomes of a fuzzy random variable. In: C. Borgelt, G. González-Rodríguez, W. Tretsching, M. A. Lubiano, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), *Advances in Intelligent and Soft Computing*, pp. 105–113, 2010.

[7] G. Deschrijver, E. Kerre: A generalization of operators on intuitionistic fuzzy sets using triangular norms and conorms. *Notes on Intuitionistic Fuzzy Sets* 1, pp. 19–27, 2002.

[8] D. H. Hong, C. Kim: A note on similarity measures between vague sets and between elements. *Information Sciences* 115, pp. 83–96, 1999.

[9] E. P. Klement, R. Mesiar, E. Pap: *Triangular Norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[10] X. Lui: Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems* 52, pp. 305–318, 1992.

[11] S. Montes, I. Couso, P. Gil, C. Bertoluzza: Divergence measure between fuzzy sets. *International Journal of Approximate Reasoning* 30, pp. 91–105, 2002.

[12] S. Montes: *Partitions and divergence measures in fuzzy models*. Doctoral Thesis, University of Oviedo, 1998.

[13] I. Montes, V. Janiš, S. Montes: An axiomatic definition of divergence for intuitionistic fuzzy sets. In: *Proceedings of the 2011 EUSFLAT conference*, Aix-Les-Bains (French), 2011.

[14] E. Szmidt, J. Kacprzyk: Analysis of similarity measures for Atanassov's intuitionistic fuzzy sets. In: J. P. Carvalho, D. Dubois, U. Kaymak, J. M. C. Sousa (Eds.), *Proceeding of IFSA-EUSFLAT 2009*, pp. 1416–1421, 2009.

[15] L. A. Zadeh: Fuzzy Sets. *Information Control* 8, pp. 338–353, 1965.

ESTUDIO DE SUCESIONES DE CONTEXTOS L-FUZZY

Cristina Alcalde[1], **Ana Burusco**[2], **Ramón Fuentes-González**[2]

[1]*Departamento de Matemática Aplicada. UPV/EHU.*
Plaza de Europa 1, 20018 San Sebastián, Spain., c.alcalde@ehu.es
[2]*Departamento de Automática y Computación. Universidad Pública de Navarra*
Campus de Arrosadía, 31006 Pamplona, Spain., {burusco, rfuentes}@unavarra.es

Resumen

En algunos casos, disponemos de una sucesión de contextos L-Fuzzy que establecen la relación existente entre un conjunto de objetos X y un conjunto de atributos Y. Un ejemplo de esta situación se tiene al estudiar la evolución en el tiempo de un contexto L-Fuzzy. En este trabajo analizaremos este tipo de situaciones y propondremos modelos que permitan realizar predicciones así como estudiar comportamientos pasados.

Palabras Clave: Contextos L-Fuzzy, Conceptos L-Fuzzy, operadores OWA.

1 INTRODUCCIÓN

La Teoría de Conceptos L-Fuzzy analiza la información existente en un contexto L-Fuzzy mediante el uso de los conceptos L-Fuzzy. Estos contextos L-Fuzzy son tuplas (L,X,Y,R), con L retículo completo, X e Y conjuntos de objetos y atributos respectivamente y $R \in L^{X\times Y}$ una relación L-Fuzzy entre los objetos y atributos.

En algunas situaciones disponemos de información sobre cómo es la evolución de la relación existente entre este conjunto de objetos X y el de atributos Y, almacenándola en una sucesión de contextos L-Fuzzy. Podemos ser más ambiciosos en estos casos e intentar predecir tendencias futuras además de estudiar comportamientos pasados. El estudio de esta sucesión de contextos L-Fuzzy será el objetivo principal de este trabajo.

Partiremos de una sucesión de contextos L-Fuzzy $(L,X,Y,R_i)_{i\in I}$, con $I \subseteq \mathbb{N}$ finito, donde X e Y son dos conjuntos de objetos y atributos respectivamente y R_i representa la relación i-ésima existente entre los objetos de X y los atributos de Y.

El objetivo final será estudiar esta sucesión de contextos L-Fuzzy y los conceptos L-Fuzzy que de ella se derivan. Para hacerlo, analizaremos dos tipos de situaciones distintas: en la primera lo relevante será estudiar los valores que destacan de los contextos L-Fuzzy, independientemente del contexto en el que estén, y en la segunda será importante mantener el orden de los contextos ya que para nosotros representará una evolución en el tiempo.

Muchas son las posibles aplicaciones de esta técnica. Por ejemplo, podemos pensar en una sucesión de contextos L-Fuzzy que recoja las ventas mensuales de artículos deportivos en unos determinados establecimientos a lo largo de un periodo de tiempo. Este estudio nos permitirá establecer tendencias en las que basarnos para tomar decisiones.

Veamos en primer lugar algunos resultados importantes de la Teoría de Conceptos L-Fuzzy:

2 CONTEXTOS L-FUZZY

El Análisis de Conceptos Formales de R. Wille ([12]) extrae información de una tabla binaria que representa un contexto formal (X,Y,R) con X e Y conjuntos finitos de objetos y atributos repectivamente y $R \subseteq X \times Y$. Esta información se obtiene por medio de los conceptos formales que son pares (A,B) con $A \subseteq X$, $B \subseteq Y$ cumpliendo $A^* = B$ y $B^* = A$, siendo $*$ un operador derivación que asocia a cada subconjunto de objetos A los atributos relacionados con los elementos de A, y a cada subconjunto de atributos B los objetos relacionados con los elementos de B. Estos conceptos se pueden interpretar como un grupo de objetos A que comparten unos atributos B.

En trabajos anteriores ([4],[5]) hemos definido los contextos L-Fuzzy (L,X,Y,R), con L retículo completo, X e Y conjuntos de objetos y atributos respectivamente y $R \in L^{X\times Y}$ una relación L-Fuzzy entre los objetos y atributos. Ésta es una extensión al caso Fuzzy de los contextos formales de R. Wille cuando se quieren estudiar relaciones entre los objetos y los atributos con valores en un retículo completo L, en lugar de valores binarios.

En nuestro caso, para trabajar con estos contextos L-Fuzzy,

utilizamos los operadores derivación 1 y 2 definidos mediante las siguientes expresiones:

$\forall A \in L^X, B \in L^Y$

$$A_1(y) = \inf_{x \in X} \{\mathscr{I}(A(x), R(x,y)\}$$

$$B_2(x) = \inf_{y \in Y} \{\mathscr{I}(B(y), R(x,y)\}$$

con $\mathscr{I}$ una implicación fuzzy definida en $(L, \leq)$ y donde A_1 representa de forma difusa a los atributos relacionados con todos los objetos de A y B_2, a los objetos relacionados con todos los atributos de B.

La información almacenada en el contexto se visualiza mediante los conceptos L-Fuzzy que son pares $(M, M_1) \in (L^X, L^Y)$ con $M \in fix(\varphi)$, conjunto de puntos fijos del operador φ, siendo éste definido a partir de los operadores derivación 1 y 2 antes mencionados como $\varphi(M) = (M_1)_2 = M_{12}$. Estos pares, cuyas primera y segunda componentes son la extensión e intensión respectivamente, representan de forma vaga a un grupo de objetos que comparten unos atributos.

El conjunto $\mathscr{L} = \{(M, M_1)/M \in fix(\varphi)\}$ con la relación de orden $\leq$ definida como:

$\forall (M, M_1), (N, N_1) \in \mathscr{L}, (M, M_1) \leq (N, N_1)$ si $M \leq N$

es un retículo completo al que llamamos ([4],[5]) retículo de conceptos L-Fuzzy.

Por otra parte, dado $A \in L^X$, (o $B \in L^Y$) podemos calcular su concepto asociado sin más que aplicar sucesivamente los operadores derivación. En el caso de utilizar la implicación de Lukasiewicz (como ocurrirá en este trabajo), el concepto asociado tiene la forma (A_{12}, A_1) (o (B_2, B_{21})).

Otros resultados importantes sobre esta teoría se recogen en [1], [6], [9] y [10].

3 SUCESIONES DE CONTEXTOS L-FUZZY

3.1 ESTUDIO GENERAL DE LA SUCESIÓN DE CONTEXTOS L-FUZZY

Si disponemos de una sucesión de contextos L-Fuzzy $(L, X, Y, R_i)_{i \in I}$ con X e Y conjuntos de objetos y atributos respectivamente y $R_i \in L^{X \times Y}, \forall i \in I$, y queremos definir un nuevo contexto L-Fuzzy que resuma la información que aportan los distintos contextos de la sucesión, tendremos que agregar los datos de las relaciones R_i. Así, podemos tomar la media (con o sin ponderación), construir los intervalos cuyo extremo inferior sea el mínimo de las correspondientes observaciones y cuyo extremo superior sea el máximo de las mismas, obteniendo de esa forma un contexto L-Fuzzy intervalo-valorado, o trabajar con contextos con varios valores. Hemos publicado algunos trabajos relacionados con este tema [7, 8].

La utilización de medias ponderadas para construir una relación que describa la información existente en las distintas relaciones nos permite asignar pesos distintos a los distintos contextos L-Fuzzy pudiendo dar de esa manera más importancia a unos que a otros. Así, la nueva relación R estará definida como:

$$R(x,y) = \sum_{i \in I} w_i . R_i(x,y), \forall x \in X, y \in Y$$

verificándose que $\sum_{i \in I} w_i = 1$, $\forall (w_i)_{i \in I}$.

Sin embargo, en muchas ocasiones ocurre que cuando tenemos una familia de contextos L-Fuzzy puede que haya ciertas observaciones de un determinado contexto que nos resulten interesantes mientras que otras no lo sean tanto.

Para estudiar situaciones similares mediante contextos con varios valores, se utilizaron en [5] multiconjuntos y expertones. En aquel caso, se analizaban todas los observaciones en su conjunto sin poder establecer distintos estudios en función de distintos niveles de exigencia. Ésta será la principal novedad de este trabajo.

Veamos el siguiente ejemplo:

Ejemplo 1. Sea $(L, X, Y, R_i)_{i \in I}$ una sucesión de contextos L-Fuzzy que representa las ventas de artículos deportivos (X) en unos establecimientos (Y) a lo largo de un periodo de tiempo (I), y lo que queremos estudiar son los lugares donde se producen las mayores ventas teniendo en cuenta que hay artículos que son de temporada (por ej. esquíes, bañadores) y de una determinada zona (se venderán más esquíes en Huesca que en Sevilla).

En este caso, el modelo de media ponderada no será válido ya que no sabríamos cómo asignar un mayor peso a un contexto L-Fuzzy determinado (en unos meses se venden más bañadores y en otros más esquíes).

Para analizar esta situación, podría ser interesante utilizar operadores OWA[11] con la mayoría de los pesos cerca de los valores más grandes. De esta forma, estaríamos dando más valor a las observaciones mayores, independientemente de en qué momento se hayan producido y, por otra parte, evitaríamos algunos valores pequeños en la relación final resultante (que como ya hemos estudiado en [2], nos pueden traer problemas a la hora de calcular los conceptos L-Fuzzy.

Veamos cómo definió Yager estos operadores OWA[11]:

Definición 1. *Una función F de $L^n \longrightarrow L$, donde $L = [0,1]$ se llama* operador OWA de dimensión *n si asociada a F está una n-tupla de pesos $W = (w_1, w_2 \ldots w_n)$ tal que $w_i \in [0,1]$ y $\sum_{1 \leq i \leq n} w_i = 1$, donde $F(a_1, a_2, \ldots a_n) = w_1.b_1 + w_2.b_2 + \cdots + w_n.b_n$, con b_i el i-ésimo mayor elemento en la colección $a_1, a_2, \ldots a_n$.*

Hay dos casos particulares de especial interés:

W_* definido como la n-tupla de pesos que tiene $w_n = 1$ y $w_j = 0, \forall j \neq n$. Y W^* definido como la n-tupla de pesos que tiene $w_1 = 1$ y $w_j = 0, \forall j \neq 1$.

Está probado que $F_*(a_1, a_2, \ldots a_n) = \text{mín}_j(a_j)$ y $F^*(a_1, a_2, \ldots a_n) = \text{máx}_j(a_j)$. Estos operadores se llaman *and* y *or*, respectivamente.

Para un estudio general de la sucesión de contextos L-Fuzzy, estaremos interesados en utilizar operadores próximos al *or*. Para medir cuánto de próximos son, podemos utilizar la definición de grado de *optimismo (orness)* dada en [11]:

Definición 2. *Sea F un operador de agregación OWA con una n-tupla de pesos $W = (w_1, w_2, \ldots w_n)$. El grado de optimismo asociado a este operador se define como:*

$$orness(W) = (1/n-1)\sum_{i=1}^{n}((n-i).w_i)$$

Ejemplo 2. Si $W = (1,0,0,\ldots 0)$, $orness(W) = 1$, y si $W = (0,0,0,\ldots 1)$, $orness(W) = 0$.

Volviendo a la situación inicial, podemos dar la siguiente definición:

Definición 3. *Sea la sucesión de contextos L-Fuzzy $(L,X,Y,R_i)_{i\in I}$, podemos obtener una relación R que resuma la información de los distintos contextos en el caso que queramos estudiar los valores grandes mediante la siguiente expresión:*

$$\begin{aligned} R(x,y) =& F(R_1(x,y), R_2(x,y) \ldots R_{Card(I)}(x,y)) = \\ =& w_1.b_1 + w_2.b_2 + \cdots + w_{Card(I)}.b_{Card(I)}, \\ & \forall x \in X, y \in Y \end{aligned}$$

donde W sea tal que $orness(W)$ sea mayor que un umbral que establezcamos.

En definitiva, estaremos interesados en establecer tuplas de pesos W cuyos primeros valores sean una cadena decreciente de valores no nulos y que, a partir de una posición, tengan todos sus valores nulos.
Un caso particular consiste en tomar la media de los k mayores valores (con $k \in \mathbb{N}, k \leq Card(I)$) utilizando una tupla de pesos W tal que $w_i = 1/k$, si $i \leq k$ y $w_i = 0$, si $i > k$.

Mediante estos operadores OWA, estaríamos representando cuantificadores (la mayoría, al menos la mitad, etc.) Variando los pesos asociados a los operadores OWA podemos pasar del mínimo (for all) al máximo (exists).

Por otra parte, para un estudio más en profundidad de la sucesión de contextos L-Fuzzy, daremos la siguiente definición que nos permite establecer relaciones utilizando una tabla de pesos con un único valor no nulo: W con $w_k = 1$, para un determinado $k \leq Card(I)$.

Definición 4. *Dada una familia de contextos L-Fuzzy $(L,X,Y,R_i)_{i\in I}$ con X e Y conjuntos de objetos y atributos respectivamente y $R_i \in L^{X\times Y}, \forall i \in I$, y dado un $k \in \mathbb{N}, k \leq Card(I)$, definimos la relación $R^{(k)}$:*

$$R^{(k)}(x,y) = \min_{z\in J_{xy}^k}\{z\}$$

donde J_{xy}^k es el conjunto formado por los k mayores valores asociados el par (x,y) en las distintas relaciones R_i.

Utilizando un operador OWA que tiene asociada la tabla de pesos W con un único valor no nulo $w_k = 1$:

$$R^{(k)}(x,y) = [F(R_1(x,y), R_2(x,y) \ldots R_{Card(I)}(x,y))],$$
$$\forall x \in X, y \in Y.$$

De esta forma, estamos diciendo que hay por lo menos k observaciones por encima de los valores de la relación $R^{(k)}$ y, por tanto, para nosotros esta relación mide el grado en que x está al menos k veces relacionado con y.

Ejemplo 3. Volvamos al ejemplo de la sucesión de contextos L-Fuzzy $(L,X,Y,R_i)_{i\in I}$ que representa las ventas de artículos deportivos ($X = \{x_1, x_2, x_3\}$) en unos establecimientos ($Y = \{y_1, y_2, y_3\}$) a lo largo de un periodo de tiempo. En las siguientes relaciones $R_i, i \in I$ que toman valores en $L = \{0, 0.1, 0.2 \ldots 0.9, 1\}$ se recogen los datos sobre el porcentaje de ventas de cada producto en cada establecimiento en función del stock y a lo largo de 5 meses.

$$R_1 = \begin{pmatrix} 0.7 & 1 & 0.8 \\ 0 & 0.1 & 0.1 \\ 0 & 0.1 & 0 \end{pmatrix} \quad R_2 = \begin{pmatrix} 1 & 0.8 & 1 \\ 0.2 & 0.4 & 0.1 \\ 0 & 0 & 0.2 \end{pmatrix}$$

$$R_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0.6 & 0.5 & 0.7 \\ 0 & 0.1 & 0.2 \end{pmatrix} \quad R_4 = \begin{pmatrix} 0.5 & 0.4 & 0.6 \\ 0.1 & 0.5 & 0.3 \\ 0.6 & 0.8 & 0.8 \end{pmatrix}$$

$$R_5 = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0.8 & 1 & 0.9 \end{pmatrix}$$

Mediante la Teoría de Conceptos L-Fuzzy queremos estudiar en qué establecimientos hay una mayor venta de cada producto sin importarnos cuándo se ha realizado la venta.

Como ya hemos comentado antes, hay productos estacionales que se venden en unos determinados periodos de tiempo y no en otros (esquíes, bañadores ...). Por tanto, intentar resumir la información de la familia de conceptos mediante la media, por ejemplo, no daría buenos resultados (si un

producto sólo se vende durante un par de meses al año, al calcular la media con los demás meses nos daría un valor proximo a 0 y no obtendríamos buenos resultados aplicando la Teoría de Conceptos L-Fuzzy.)

Por otra parte, si queremos estudiar en qué lugares tienen más exito ciertos artículos deportivos, por ejemplo x_1 y x_3, podríamos calcular para cada una de las relaciones los conceptos L-Fuzzy que se derivan. En este caso, obtendríamos resultados ligados al mes de donde provienen los datos.

Sin embargo, si construimos la nueva cadena de relaciones:

$$R^{(1)} = \begin{pmatrix} 1 & 1 & 1 \\ 0.6 & 0.5 & 0.7 \\ 0.8 & 1 & 0.9 \end{pmatrix} R^{(2)} = \begin{pmatrix} 1 & 1 & 1 \\ 0.2 & 0.5 & 0.3 \\ 0.6 & 0.8 & 0.8 \end{pmatrix}$$

$$R^{(3)} = \begin{pmatrix} 0.7 & 0.8 & 0.8 \\ 0.1 & 0.4 & 0.1 \\ 0 & 0.1 & 0.2 \end{pmatrix} R^{(4)} = \begin{pmatrix} 0.5 & 0.4 & 0.6 \\ 0 & 0.1 & 0.1 \\ 0 & 0.1 & 0.2 \end{pmatrix}$$

$$R^{(5)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

podemos fijar nuestro nivel de exigencia en un k. Por ejemplo, si $k = 2$ y tomamos como contexto L-Fuzzy $(L,X,Y,R^{(2)})$, podemos calcular los conceptos que se obtienen asociados a los puntos básicos x_1 y x_3 utilizando la implicación de Lukasiewicz ($\mathscr{I}(a,b) = \text{mín}(1,1-a+b)$):

$$x_1 \longrightarrow (\{x_1/1, x_2/0.2, x_3/0.6\}, \{y_1/1, y_2/1, y_3/1\})$$
$$x_3 \longrightarrow (\{x_1/1, x_2/0.5, x_3/1\}, \{y_1/0.6, y_2/0.8, y_3/0.8\})$$

Entonces, podríamos decir que el artículo x_1 ha tenido mucho éxito en los tres establecimientos, por lo menos durante dos meses, y que hay, por lo menos en dos meses, ventas altas de los artículos x_1 y x_3, en mayor medida en los establecimientos y_2 e y_3.

Análogamente, podemos tomar otros niveles de k distintos.

En resumen, el cálculo de los conceptos asociados a la relación $R^{(1)}$ nos permitirá analizar en qué tiendas se han producido las mayores ventas de cada artículo durante un mes (independientemente de qué mes sea). Si pasamos a una relación $R^{(k)}$ estaremos relajando la exigencia y mirando las k ventas mayores para nuestro estudio.

Estos estudios nos permitirían no tener en cuenta ciertos valores pequeños que pueden ir apareciendo (si un producto no es de temporada, su venta es próxima a 0) y que si tomásemos la media de las relaciones que forman la familia, nos distorsionarían los resultados.

La observación de estos conceptos L-Fuzzy nos da la idea para las siguientes proposiciones:

Proposición 1. *Sea $k \in \mathbb{N}$, con $k \leq Card(I)$. Si (A,B) es un concepto L-Fuzzy del contexto $(L,X,Y,R^{(k)})$, $\forall h \in \mathbb{N}, h \leq k$, existirá un concepto L-Fuzzy (C,D) del contexto $(L,X,Y,R^{(h)})$ tal que $A \leq C$ y $B \leq D$.*

Demostración. Si $k = h$ es obvio. En otro caso, teniendo en cuenta que $R^{(k)} \leq R^{(h)}$ con $h < k$, entonces B, que es el subconjunto derivado de A en $(L,X,Y,R^{(k)})$, estará contenido en el subconjunto derivado de A en $(L,X,Y,R^{(h)})$, al que llamaremos D. Por tanto, $B \leq D$.

Si ahora derivamos de nuevo D en $(L,X,Y,R^{(h)})$, obtenemos un D_2 al que llamaremos C y, por otro lado, teniendo en cuenta que un conjunto está siempre contenido en el que se obtiene derivando dos veces si la implicación es residuada [3]: $A \leq A_{12} = D_2 = C$. Por tanto, la otra desigualdad es también cierta. Además, es obvio que si utilizamos una implicación residuada (C,D) es un concepto L-Fuzzy teniendo en cuenta cómo se ha definido.

□

El siguiente resultado establece relaciones entre los conceptos L-Fuzzy asociados a un mismo conjunto de partida (ver apartado 2) en los distintos contextos L-Fuzzy.

Proposición 2. *Sean $k,h \in \mathbb{N}$, con $k,h \leq Card(I)$ y sea $A \in L^X$. Llamaremos $(\hat{A}^k, \hat{B}^k)$ y $(\hat{A}^h, \hat{B}^h)$ a los conceptos asociados a A en los contextos L-Fuzzy $(L,X,Y,R^{(k)})$ y $(L,X,Y,R^{(h)})$ respectivamente. Si $k \leq h$ entonces $\hat{B}^k \geq \hat{B}^h$.*

Si además $\mathscr{I}$ es una implicación residuada y A es un punto básico, $\hat{A}^k(x_i) = \hat{A}^h(x_i) = 1$ siendo x_i el elemento de X donde el punto básico A toma el valor 1.

Un resultado similar se obtiene si partimos de un conjunto L-Fuzzy de atributos $B \in L^Y$.

Demostración. Sea $A \in L^X$. Si desarrollamos las extensiones de los dos conceptos L-Fuzzy:

$$\begin{aligned} \hat{B}^k(y) &= \inf_{x \in X} \{\mathscr{I}(A(x), R^{(k)}(x,y))\} \\ &\geq \inf_{x \in X} \{\mathscr{I}(A(x), R^{(h)}(x,y))\} = \hat{B}^h(y) \end{aligned}$$

Esto se cumpliría para cualquier A y cualquier implicación.

Por otra parte, siempre que se trabaja con puntos básicos:

$$A(x) = \begin{cases} 1 & \text{si } x = x_i \\ 0 & \text{en otro caso} \end{cases}$$

y una implicación residuada, es cierto que el valor de la extensión de los conceptos L-Fuzzy en x_i es 1:

$$\hat{B}^k(y) = \inf_{x \in X} \{\mathscr{I}(A(x), R^{(k)}(x,y))\} = R^{(k)}(x_i, y)$$

$$\hat{A}^k(x) = \inf_{y\in Y}\{\mathscr{I}(B^{(k)}(y), R^{(k)}(x,y))\}$$
$$= \inf_{y\in Y}\{\mathscr{I}(R^{(k)}(x_i,y), R^{(k)}(x,y))\}.$$

Por tanto, $\hat{A}^k(x_i) = 1$.

Análogamente se puede demostrar el resultado para el otro conjunto L-Fuzzy.

□

Sin embargo, la desigualdad $\hat{A}^k \leq \hat{A}^h$ no siempre se cumple, como se puede ver si volvemos al ejemplo anterior y comparamos la extensión del concepto derivado de $A(x) = \{x_1/1, x_2/0, x_3/0\}$ en el contexto L-Fuzzy $(L,X,Y,R^{(2)})$ y en $(L,X,Y,R^{(4)})$:

En $(L,X,Y,R^{(2)})$, $\hat{A}^2 = \{x_1/1, x_2/0.2, x_3/0.6\}$ mientras que en $(L,X,Y,R^{(4)})$ el resultado es $\hat{A}^4 = \{x_1/1, x_2/0.5, x_3/0.5\}$.

3.2 EVOLUCIÓN EN EL TIEMPO DE LA SUCESIÓN DE CONTEXTOS L-FUZZY

Fijado un valor de k, y un par objeto-atributo (x,y), la Definición 4 trabaja con el mínimo de las k mayores observaciones $R_i(x,y), i \in I$, de la sucesión de contextos L-Fuzzy, pero no nos permite hacer un análisis de su evolución en el tiempo. Para abordar este asunto, la siguiente definición nos dirá cuánto vale como mínimo la relación existente entre cada objeto y cada atributo a partir de un instante h.

Definición 5. *Dada una familia de contextos L-Fuzzy $(L,X,Y,R_i)_{i\in I}$ con X e Y conjuntos de objetos y atributos respectivamente y $R_i \in L^{X\times Y}$, construiremos una relación $\bar{R}^{(h)}$ que indicará cuál es el valor mínimo que toman las distintas relaciones a partir del instante h, con $h \leq Card(I)$.*

$$\bar{R}^{(h)}(x,y) = \min_{i\geq h}\{R_i(x,y)\}, \forall x \in X, y \in Y$$

O lo que es lo mismo, utilizando un operador OWA que tiene asociada la tabla de pesos W de dimensión $k = Card(I) - h + 1$ con un único valor no nulo $w_k = 1$:

$$\bar{R}^{(h)}(x,y) = [F(R_h(x,y), R_{h+1}(x,y) \ldots R_{Card(I)}(x,y))],$$
$$\forall x \in X, y \in Y.$$

Igual que en el apartado anterior, en lugar del mínimo (que es muy exigente) podríamos tomar el máximo, la media u otra agregación sin más que cambiar la tabla de pesos W del operador OWA. En el caso de utilizar el máximo, los grados de pertenencia que se obtendrían en los conceptos L-Fuzzy correspondientes a las relaciones $\bar{R}$ serían los más grandes.

Ejemplo 4. Volviendo al ejemplo anterior, la sucesión de contextos L-Fuzzy que se obtiene con la Definición 5 está asociada a las siguientes relaciones:

$$\bar{R}^{(1)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bar{R}^{(2)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$$

$$\bar{R}^{(3)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0.1 & 0.2 \end{pmatrix} \bar{R}^{(4)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0.6 & 0.8 & 0.8 \end{pmatrix}$$

$$\bar{R}^{(5)} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0.8 & 1 & 0.9 \end{pmatrix}$$

Si queremos estudiar las tendencias futuras de la sucesión, tomaremos un valor de h próximo al $Card(I)$ y analizaremos los conceptos L-Fuzzy que se obtienen para dicho contexto L-Fuzzy:

Por ejemplo, si $h = 4$, obtenemos los siguientes resultados tomando como contexto L-Fuzzy $(L,X,Y,\bar{R}^{(4)})$ y utilizando la implicación de Lukasiewicz en el cálculo de los conceptos L-Fuzzy asociados a los puntos básicos:

$$x_1 \longrightarrow (\{x_1/1, x_2/0.9, x_3/1\}, \{y_1/0.1, y_2/0, y_3/0\})$$
$$x_2 \longrightarrow (\{x_1/0.9, x_2/1, x_3/1\}, \{y_1/0, y_2/0.1, y_3/0\})$$
$$x_3 \longrightarrow (\{x_1/0.2, x_2/0.2, x_3/1\}, \{y_1/0.6, y_2/0.8, y_3/0.8\})$$

Podríamos decir que la tendencia futura es que únicamente el artículo x_3 se venderá bien y lo hará en todos los establecimientos mientras que x_1 y x_2 se venderán poco y siempre en compañía de x_3, el primero en el establecimiento y_1 fundamentalmente y el segundo en y_2.

Además, podemos establecer comparaciones entre los distintos conceptos L-Fuzzy que se obtienen a partir de las distintas relaciones $\bar{R}^{(i)}, \forall i \in I$.

Proposición 3. *Sea $A \in L^X$. Sean $(\hat{\hat{A}}^k, \hat{\hat{B}}^k)$ y $(\hat{\hat{A}}^h, \hat{\hat{B}}^h)$ los conceptos asociados a A en los contextos L-Fuzzy $(L,X,Y,\bar{R}^{(k)})$ y $(L,X,Y,\bar{R}^{(h)})$ respectivamente, con $k,h \leq Card(I)$. Si $k \leq h$ entonces $\hat{\hat{B}}^k \leq \hat{\hat{B}}^h$.*

Si además $\mathscr{I}$ una implicación residuada y A es un punto básico, $\hat{\hat{A}}^k(x_i) = \hat{\hat{A}}^h(x_i) = 1$ siendo x_i el elemento de X donde el punto básico A toma el valor 1.

Un resultado similar se obtiene si partimos de un conjunto L-Fuzzy de atributos $B \in L^Y$.

Demostración. Similar a la de la Proposición 2 teniendo en cuenta que en este caso, si $k \leq h$ entonces $\bar{R}^{(k)} \leq \bar{R}^{(h)}$.

□

Ejemplo 5. En nuestro ejemplo, los conceptos que se obtienen partiendo del punto básico asociado al objeto x_3 en los contextos L-Fuzzy $(L,X,Y,\bar{R}^{(4)})$ y $(L,X,Y,\bar{R}^{(5)})$, utilizando la implicación de Lukasiewicz son:

$$\bar{R}^{(4)} : (\{x_1/0.2, x_2/0.2, x_3/1\}, \{y_1/0.6, y_2/0.8, y_3/0.9\})$$

$$\bar{R}^{(5)} : (\{x_1/0, x_2/0.1, x_3/1\}, \{y_1/0.8, y_2/1, y_3/0.9\})$$

que cumplen la proposición anterior.

Por otra parte, como si un objeto y un atributos están relacionados a partir de un instante h, estarán por lo menos relacionados $Card(I)-h+1$ veces, entonces se puede ver que existe un resultado análogo entre los conceptos L-Fuzzy que se obtienen al utilizar las relaciones definidas en 4 y 5.

Proposición 4. *Si partimos de cualquier $A \in L^X$, entonces, para un valor de $h \in I$, la intensión del concepto L-Fuzzy $(\hat{\hat{A}}^h, \hat{\hat{B}}^h)$ que se obtiene en $(L,X,Y,\bar{R}^{(h)})$ estará contenida en la intensión del concepto L-Fuzzy $(\hat{A}^k, \hat{B}^k)$ que se obtiene en $(L,X,Y,R^{(k)})$ con $k = Card(I)-h+1$.*

De igual forma, obtenemos un resultado análogo partiendo de $B \in L^Y$.

Demostración. Inmediata teniendo en cuenta la demostración de la proposición anterior y que $\bar{R}^{(h)} \leq R^{(k)}$ siendo $k = Card(I)-h+1$.

□

Ejemplo 6. Si tomamos $h = 4$ y $k = 2$, y los contextos L-Fuzzy representados por las relaciones $(L,X,Y,\bar{R}^{(4)})$ y $(L,X,Y,R^{(2)})$, partiendo del objeto x_1 y utilizando la implicación de Lukasiewicz, se obtienen los conceptos L-Fuzzy:

$$\bar{R}^{(4)} : (\{x_1/1, x_2/0.9, x_3/1\}, \{y_1/0.1, y_2/0, y_3/0\})$$

$$R^{(2)} : (\{x_1/1, x_2/0.2, x_3/0.6\}, \{y_1/1, y_2/1, y_3/1\})$$

Y se cumple la proposición anterior.

4 CONCLUSIONES Y TRABAJO FUTURO

En este trabajo vemos cómo los operadores OWA pueden ser de gran utilidad cuando queremos estudiar una sucesión de contextos L-Fuzzy y los conceptos que de ella se derivan.

Por otra parte, los contextos que evolucionan con el tiempo se pueden generalizar si estudiamos contextos L-Fuzzy en los que las observaciones sean a su vez otros contextos L-Fuzzy. Esta tarea la abordaremos en próximos trabajos.

Agradecimientos

Artículo parcialmente subvencionado por el Grupo de Investigación de la Universidad Pública de Navarra *Adquisición de conocimiento y minería de datos, funciones especiales y métodos numéricos avanzados* y por el Grupo de Investigación del Gobierno Vasco *Sistemas Inteligentes y Energía (SI+E)* (Proyecto IT519-10).

Referencias

[1] C. Alcalde, A. Burusco, R. Fuentes-González, I. Zubia: Treatment of L-fuzzy contexts with absent values, *Information Sciences* 179 (1-2), pp.1-15, 2009.

[2] C. Alcalde, A. Burusco, R. Fuentes-González, I. Zubia: The use of linguistic variables and fuzzy propositions in the L-Fuzzy Concept Theory, *Computers and Mathematics with Applications* 62, pp.3111-3122, 2011.

[3] R. Bělohlávek: Fuzzy Galois Connections, *Math. Logic Quarterly* 45 (4), pp. 497–504, 1999.

[4] A. Burusco: R. Fuentes-González: The Study of the L-fuzzy Concept Lattice, *Mathware and Soft Computing* 1 (3), pp. 209-218, 1994.

[5] A. Burusco: R. Fuentes-González: Construction of the L-fuzzy Concept Lattice, *Fuzzy Sets and Systems* 97 (1), pp. 109-114, 1998.

[6] A. Burusco: R. Fuentes-González: Concept lattices defined from implication operators, *Fuzzy Sets and Systems* 114 (1), pp. 431-436, 2000.

[7] A. Burusco: R. Fuentes-González: Contexts with multiple weighted values, *The International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* 9 (3), pp. 355-368, 2001.

[8] A. Burusco: R. Fuentes-González: The study of the interval-valued contexts, *Fuzzy Sets and Systems* 121, pp. 69-82, 2001.

[9] J. Medina, M. Ojeda-Aciego: Multi-adjoint t-concept lattices, *Information Sciences* 180 (5), pp. 712-725, 2010.

[10] Y. Djouadi, H. Prade: Interval-Valued Fuzzy Galois Connections: Algebraic Requirements and Concept Lattice Construction, *Fundamenta Informaticae* 99 (2), pp. 169-186, 2010.

[11] R.R. Yager: On ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Transactions on Systems, Man and Cibernetics* 18, pp. 183-190, 1988.

[12] R. Wille: Restructuring lattice theory: an approach based on hierarchies of concepts, in: *Rival I. (Ed.), Ordered Sets, Reidel, Dordrecht-Boston*, pp. 445-470, 1982.

Un algoritmo rápido para obtener retículos de conceptos utilizando el lenguaje de programación *Python*

Juan Carlos Díaz, Jesús Medina, Rafael Rodríguez

Department of Mathematics. University of Cádiz. Spain, {juancarlos.diaz,jesus.medina,rafael.rodriguez}@uca.es

Resumen

Python es un lenguaje de programación versátil y rápido, que integra sistemas muy eficientemente.

Este artículo presenta una implementación, de un algoritmo recientemente introducido para encontrar los puntos fijos de operadores de clausura, mediante el lenguaje de programación *Python*, con el fin de obtener los retículos de conceptos de varios entornos, tales como los retículos de conceptos multi-adjuntos, los retículos de conceptos orientados a propiedades, etc.

Keywords: Análisis formal de conceptos; lenguaje de programación *Python*; conjuntos difusos.

1. INTRODUCCIÓN

El Análisis Formal de Conceptos es una teoría del análisis de datos que identifica estructuras conceptuales entre conjuntos de datos. Fue introducida por Rudolf Wille en 1982 [20] y se ha desarrollado rápidamente por su gran utilidad, tanto en un ambiente clásico [7] como en ambientes con información imprecisa, falta de información, etc. [2, 3, 6, 8, 9, 10, 21]. Se ha aplicado en multitud de campos tales como: la medicina, la psicología, la musicología, en las bases de datos lingüísticas, las ciencias de la información, en ecología, etc.

Recientemente, se ha introducido un enfoque más flexible y versátil que el clásico, una extensión difusa del Análisis Formal de Conceptos, los "retículos de conceptos multi-adjuntos", cuya esencia principal es la filosofía multi-adjunta, la cual permite que se pueda utilizar el nuevo enfoque en un mayor número de situaciones, obteniendo mayor información y mejores resultados a partir de la base de datos inicial [18, 17, 16, 15, 12]. Por ejemplo, esta técnica puede usarse en la extracción de información como herramienta para el desarrollo del software que permita, por ejemplo, la gestión eficiente de la energía en una infraestructura.

Además, esta teoría se ha relacionado con la teoría de los rough sets, la cual también es otra herramienta de extracción y manipulación de información en bases de datos relacionales con información imprecisa, falta o pérdida de información, etc. [14, 13].

En este trabajo se presenta un programa para el cálculo del retículo de conceptos asociado a un contexto difuso que ha sido desarrollado con un doble objetivo: la flexibilidad, desde el punto de vista de la rápida implementación de distintos algoritmos, y la posibilidad de extensión hacia la paralelización y la computación de alto rendimiento.

La versión actual, que utiliza el algoritmo recientemente propuesto en [5], el cual calcula los puntos fijos de operadores de clausura y se basa en un cálculo de los vecinos superiores de cada punto fijo del operador de clausura, ha sido implementada en *Python*, un lenguaje de programación interpretado de propósito general que ha visto incrementada drásticamente su popularidad en la última década.

Se ha intentado explotar algunas de sus ventajas, entre ellas su versatilidad y rapidez de desarrollo, así como su orientación a objetos. Aunque *Python*, como lenguaje interpretado, no destaca por sus grandes prestaciones de cálculo, se valoró especialmente la variedad de alternativas existentes para adaptar los algoritmos hacia la computación de alto rendimiento, una vez estos han sido debidamente probados y validados. Entre estas posibilidades se encuentra todo un rango que abarca desde el uso de bibliotecas orientadas al cálculo matricial, como *Scipy/Numpy*, hasta la programación de módulos específicos en *C* o *C++* (las integración de *Python* y *C/C++* es natural y existen numerosas

posibilidades para ello), que podrían llegar a comprender la totalidad del programa, haciendo así posible el alcanzar tiempos de cómputo óptimos.

Esta preocupación por la velocidad de proceso es especialmente importante en el ámbito del análisis formal de conceptos, donde el número de operaciones crece drásticamente cuando aumenta la talla del conjunto de objetos y atributos. En la versión actual del programa se ha hecho un importante esfuerzo para, en estas operaciones, maximizar el uso de la aritmética de enteros y eliminando el uso de operaciones en punto flotante, disminuyendo así el tiempo de cálculo. Además, se han experimentado algunas variantes orientadas a la paralelización de los algoritmos, lo cual es un requisito ineludible si realmente deseamos profundizar en la computación de alto rendimiento en el marco del análisis formal de conceptos difusos.

2. RETÍCULOS DE CONCEPTOS MULTI-ADJUNTO

Los bloques básicos de los retículos de conceptos multiadjuntos son los *triples adjuntos*, que son generalizaciones de los pares adjuntos en el caso en el que el conjuntor sea no conmutativo.

Puesto que el conjuntor $\&$ es no conmutativo, existen dos formas diferentes de generalizar la propiedad de adjunción usual, dependiendo de qué argumento del conjuntor fijemos. Esto nos conduce a dos implicaciones diferentes $\swarrow$ y $\nwarrow$.

Definición 1 *Sean* $(P_1, \leq_1)$, $(P_2, \leq_2)$, $(P_3, \leq_3)$ *posets y* $\&\colon P_1 \times P_2 \to P_3$, $\swarrow\colon P_3 \times P_2 \to P_1$, $\nwarrow\colon P_3 \times P_1 \to P_2$ *aplicaciones, entonces la tupla* $(\&, \swarrow, \nwarrow)$ *es un* triple adjunto *con respecto a* P_1, P_2, P_3 *si:*

- $\&$ *es creciente en ambos argumentos.*
- $\swarrow$ *y* $\nwarrow$ *es creciente en el primer argumento (consecuente) y decreciente en el segundo (antecedente).*
- $x \leq_1 z \swarrow y$ *si y sólo si* $x \,\&\, y \leq_3 z$ *si y sólo si* $y \leq_2 z \nwarrow x$, *donde* $x \in P_1$, $y \in P_2$ *y* $z \in P_3$.

La tercera propiedad es la *propiedad de adjunción* y proporciona una generalización de la regla del modus ponens en un marco multi-valuado. Obsérvese que no se impone ninguna codición de frontera a diferencia con las definiciones de retículo multiadjunto y de triple de implicación introducida en [19, 1].

2.1. MARCO Y CONTEXTO MULTI-ADJUNTO

La definición de contexto multi-adjunto de presenta como una generalización natural de la de contexto clásico, aunque para esto hay que fijar los soportes en los que se realizarán los cálculos y los triples adjuntos que se podrán usar, que vendrán dados por la noción de marco multi-adjunto. Estas definiciones, la de marco y contexto multiadjuntos, resultan ser básicas para la definición de retículo de conceptos multiadjunto.

Definición 2 *Dados los retículos completos* $(L_1, \preceq_1)$ *y* $(L_2, \preceq_2)$, *el conjunto parcialmente ordenado (poset)* $(P, \leq)$ *y la tupla de triples adjuntos* $(\&_i, \swarrow^i, \nwarrow_i)$ *con respecto a* L_1, L_2, P *para todo* $i = 1, \ldots n$. *Un* marco multi-adjunto, $\mathcal{L}$, *es una tupla*

$$(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$$

Por comodidad en la notación, a un marco multi-adjunto se le denotara simplemente como $(L_1, L_2, P, \&_1, , \ldots, \&_n)$.

Una vez introducida la noción de marco multi-adjunto podemos pasar a definir la de *contexto multi-adjunto*, la cual está compuesta, como en su forma clásica, por objetos, atributos y una relación difusa entre ellos, y, además, por una función que asigna a cada objeto (o atributo) un triple adjunto, la cual nos permite, en la práctica, dar grados de preferencia a subconjuntos de objetos con respecto al resto. Formalmente:

Definición 3 *Dado un marco multi-adjunto* $(L_1, L_2, P, \&_1, , \ldots, \&_n)$, *un* contexto multiadjunto *es una tupla* (A, B, R, σ) *tal que* A *y* B *son conjuntos no vacíos, usualmente interpretados como objetos y atributos respectivamente,* R *es una relación* P*-difusa* $R\colon A \times B \to P$ *y* $\sigma\colon B \to \{1, \ldots n\}$ *es un aplicación que asocia a cada elemento de* B *un triple adjunto del marco.*

2.2. RETÍCULO DE CONCEPTOS

A partir de un marco y contexto multi-adjunto podemos construir el retículo de conceptos asociado a éstos usando las propiedades de dos operadores que forman una conexión de Galois.

Definición 4 *Dados dos posets* $(P_1, \leq_1)$ *y* $(P_2, \leq_2)$, *y las aplicaciones* $^{\downarrow}\colon P_1 \to P_2$, $^{\uparrow}\colon P_2 \to P_1$, *se dice que el par* $(^{\uparrow}, ^{\downarrow})$ *forma una* conexión de Galois *entre* P_1 *y* P_2 *si y sólo si:*

1. $^{\uparrow}$ *y* $^{\downarrow}$ *conservan el orden.*
2. $x \leq_1 x^{\downarrow\uparrow}$, *para todo* $x \in P_1$.

3. $y \leq_2 y^{\uparrow\downarrow}$*, para todo* $y \in P_2$*.*

La aplicaciones que generalizan a los operadores clásicos son de la forma: $^{\uparrow_\sigma}: L_2^B \to L_1^A$ y $^{\downarrow^\sigma}: L_1^A \to L_2^B$, definidas, para cada $a \in A$ y $b \in B$, como

$$\begin{aligned} g^{\uparrow_\sigma}(a) &= inf\{R(a,b') \swarrow^{\sigma(b')} g(b') \mid b' \in B\} \\ f^{\downarrow^\sigma}(b) &= inf\{R(a',b) \nwarrow_{\sigma(b)} f(a') \mid a' \in A\} \end{aligned}$$

Como se ha comentado, estas aplicaciones forman una conexión de Galois.

Un *concepto multi-adjunto* se define de forma usual a la clásica, es decir, es un par (g, f) tal que $g \in L_2^B$, $f \in L_1^A$, y que satisface $g^\uparrow = f$ y $f^\downarrow = g$. Al subconjunto difuso de objetos g se le llama extensión del concepto y a f intensión del concepto.

Finalmente, se define el retículo de conceptos multi-adjunto.

Definición 5 *El* retículo de conceptos multi-adjunto *asociado a un marco* $(L_1, L_2, P, \&_1, , \ldots, \&_n)$ *y un contexto* (A, B, R, σ) *es el conjunto*

$$\mathcal{M} = \{(g, f) \mid g \in L_2^B, f \in L_1^A \text{ y } g^\uparrow = f, f^\downarrow = g\}$$

con el siguiente orden $(g_1, f_1) \preceq (g_2, f_2)$ *si y solo si* $g_1 \preceq_1 g_2$ *o, equivalentemente,* $f_2 \preceq_2 f_1$*.*

En [17] se prueban todos los resultados que se han comentado anteriormente, incluso se expresa cómo se obtienen el supremo y el ínfimo de cualquier subconjunto de conceptos.

$$\begin{aligned} \bigwedge_{i\in I}(g_i, f_i) &= (\bigwedge_{i\in I} g_i, (\bigvee_{i\in I} f_i)^{\downarrow\uparrow} \\ \bigvee_{i\in I}(g_i, f_i) &= ((\bigvee_{i\in I} g_i)^{\uparrow\downarrow}, \bigwedge_{i\in I} f_i) \end{aligned}$$

Este hecho nos podría servir para calcular todos los conceptos, partiendo del menor concepto e ir construyendo los siguientes conceptos ayudándonos de las igualdades anteriores, pero este proceso es poco eficiente como lo son otros que se podrían aplicar.

Antes de terminar esta sección debemos recordar una propiedad importante que satisfacen los operadores de una conexión de Galois, la cual será muy importante para poder usar el algoritmo introducido en [5] y que se pretende implementar.

Esta propiedad hace referencia a los operadores de clausura:

Definición 6 *Dado un poset* $(P, \leq)$*, la aplicación* $c: P \to P$ *se dice que es un* operador de clausura *si es creciente,* $x \leq c(x)$ *y* $c(c(x)) = c(x)$*, para todo* $x \in P$*.*

Concretamente, dice que, dados dos posets $(P_1, \leq_1)$ y $(P_2, \leq_2)$, y una conexión de Galois entre ellos $(^\uparrow, ^\downarrow)$, se verifica que los operadores $^{\uparrow\downarrow}: P_2 \to P_2$ y $^{\downarrow\uparrow}: P_1 \to P_1$ son operadores de clausura. Por lo tanto, la composición de las aplicaciones anteriores $^{\uparrow_\sigma}: P_2 \to P_1$ y $^{\downarrow^\sigma}: P_1 \to P_2$, forman operadores de clausura, es decir, $^{\uparrow_\sigma\downarrow^\sigma}: L_2^B \to L_2^B$ y $^{\downarrow^\sigma\uparrow_\sigma}: L_1^A \to L_1^A$ son operadores de clausura.

3. ALGORITMO EN *Python*

Para el caso clásico, en el que se consideran conjuntos crips, existen distintos algoritmos que obtienen todos los conceptos y su relación entre ellos, es decir, el retículo completo de conceptos, de forma rápida y eficiente, como por ejemplo los dados en [4, 11]. Sin embargo, no hay muchos algoritmos eficientes para el caso difuso.

3.1. Descripción del algoritmo a implementar

Entre los algoritmos desarrollados para el caso difuso, el más destacable es el presentado en [5], en el que se ha introducido un rápido algoritmo para obtener los puntos fijos de un operador de clausura, operador que se obtiene, como se comentó anteriormente, al componer los operadores de la conexión de Galois utilizados para obtener el retículo de conceptos multi-adjunto.

Concretamente, define tres módulos que detallamos a continuación: NEIGHBORS(B, C), que obtiene los vecinos superiores de un conjunto difuso B a partir de un operador de clausura C. GENERATEFROM(B) es una función recursiva que, a partir de un conjunto B, obtiene todos los puntos fijos del operador de clausura C mayores que B, sus vecinos inferiores y superiores. Finalmente, LATTICE(C, Y) aplica el módulo anterior al menor punto fijo del operador de clausura C, obteniendo como resultado el conjunto total de puntos fijos de C junto con sus vecinos inferiores y superiores, es decir, toda la información necesaria para construir el retículo completo de puntos fijos.

Debemos destacar que el algoritmo necesita un sólo soporte, que sea finito y lineal, esto se traduce en nuestro entorno a que el marco multi-adjunto considerado tenga como retículos L_1, L_2 y P a un intervalo granulado del intervalo unidad $[0, 1]$.

A su favor está el hecho de que se pueden usar distintos triples adjuntos, pues la composición de los operadores de la conexión de Galois del enfoque multi-adjunto forman operadores de clausura aunque se estén usando distintos triples adjuntos.

3.2. Lenguaje y herramientas

Python es un lenguaje de programación interpretado de propósito general que ha visto incrementada drásticamente su popularidad en la última década.

Se ha utilizado *Python* para la implementación realizada por su versatilidad y rapidez de desarrollo, así como por su orientación a objetos. Además, existe una gran variedad de alternativas para adaptar el algoritmo desarrollado hacia la computación de alto rendimiento. Por ejemplo, se pueden usar bibliotecas orientadas al cálculo matricial, como *Scipy*/*Numpy*, la programación de módulos específicos en *C* o *C++*, que podrían llegar a comprender la totalidad del programa, haciendo así posible el alcanzar tiempos de cómputo óptimos.

Esta preocupación por la velocidad de proceso es especialmente importante en el ámbito del análisis formal de conceptos, donde el número de operaciones crece drásticamente cuando aumenta el cardinal del conjunto de objetos y atributos.

En la implementación del algoritmo se ha hecho un importante esfuerzo para maximizar el uso de la aritmética de enteros y eliminando el uso de operaciones en punto flotante, lo que debe suponer una disminución en el tiempo de cálculo.

Para ello, la clave ha estado en el desarrollo de nuestra propia biblioteca de conjuntos difusos (en vez de optar por las distintas implementaciones existentes en *Python*) que, estando directamente orientada a conjuntos de valores de verdad uniformemente espaciados, nos permite representarlo mediante rangos finitos de números enteros. La comprobación de las mejoras propuestas con respecto a otras implementaciones se desarrollará como trabajo futuro.

Para las operaciones matriciales se ha utilizado la biblioteca *Scipy* y, además, se han experimentado algunas variantes para la paralelización de los algoritmos, lo cual es un requisito ineludible si realmente deseamos profundizar en la computación de alto rendimiento en el marco del análisis formal de conceptos.

3.3. Descripción detallada de la implementación realizada

Como se comentó en la introducción, para la implementación del algoritmo se han aprovechado las características de orientación a objetos de Python, lo cual fomenta la reutilización y la adaptación del código, agilizando su desarrollo y mantenimiento y dejando abierto el camino para la programación de variantes más complejas.

Así, la implementación se basa en las siguientes clases:

`FuzzySet.` Conjunto difuso sobre un universo de valores de verdad equiespaciados, $x_k = x_0 + h \cdot k$, $k = 0, ..., n$. En nuestra implementación, un conjunto difuso queda unívocamente definido por un rango de números enteros, $k_1, k_2, \ldots, k_r$ $(0 \leq r \leq n)$ más h y así, de forma interna, evitamos el uso de números con representación en coma flotante. Los objetos de esta clase implementan métodos habituales en conjuntos difusos (unión, intersección, etc.) y, además, se define el método `increase(k)`, que incrementa (en h) el valor de verdad asociado al índice `k`. Esta función será usada en el método `neighbors` de la clase `FCA_lattice`, descrita más adelante.

`FuzzyContext.` Contexto formal que relaciona dos conjunto difusos: objetos y atributos. Esta relación, R, se define a través de una matriz, implementada mediante el módulo Python "*NumPy*". Los métodos fundamentales de la clase `FuzzyContext` son `derivedObjects(f)` y `derivedAttributes(g)` que devuelven listas de enteros que representan a la extension e intensión del subconjunto difuso de atributos `f` y al subconjunto difuso de objetos `g`. Usando éstos se definen los operadores de clausura `attributes_closure(B)` y `objects_closure(A)`.

`Concept.` Par (g, f), donde g es un subconjunto difuso de objetos y f es un subconjunto difuso de atributos, de forma que `derivedObjects(f)` es igual a `derivedAttributes(g)`. Sobre esta clase definimos operadores $<$, $\leq$, etc. que definen una relación de orden adecuada.

`FCA_lattice.` Clase principal, que se construye a partir de un objeto de tipo `FuzzyContext` y utiliza las clases antes descritas como soporte para programar los distintos módulos que componen el algoritmo ([5]): `neighbors`, `generateFrom` y `lattice`.

3.4. Ejemplos

A continuación, se van a presentar algunos ejemplos en los que se ha utilizado la implementación realizada. El primero de éstos se toma de [5], y se comprueba que se tienen los mismos resultados.

Ejemplo 1 *Se considera el marco multi-adjunto* $(L, \preceq, \&_L)$*, en el que* $L = \{0, 0{,}5, 1\}$ *y* $\&_L$ *es el conjuntor de Łukasiewicz definido sobre* L*. Sobre este marco, el contexto viene dado por la tupla* (A, B, R)*, donde* $A = \{y_1, y_2, y_3, y_4, y_5\}$*,* $B = \{x_1, x_2, x_3\}$ *y* R*:* $A \times B \to L$ *viene dada por la Tabla 2. En este ejemplo, como existen* $3^3 = 27$ *subconjuntos difusos de objetos y* $3^5 = 243$ *subconjuntos de atributos difusos, como mucho pueden*

Cuadro 1: Relación R del Ejemplo 1.

R	x_1	x_2	x_3
y_1	1	1	0
y_2	0,5	1	0
y_3	0,5	1	0,5
y_4	1	1	0,5
y_5	1	0,5	1

existir 27 *conceptos, que sería si todos los subconjuntos de objetos fueran intensiones de conceptos, sin embargo son muchos menos los conceptos que se obtienen.*

En este caso, el retículo de conceptos $(\mathcal{M}_L, \preceq)$*, asociado al marco y contexto fijados anteriormente, tiene 10 conceptos que se enumeran a continuación.*

$$C_0 = \langle \{0{,}5/x_1, 1{,}0/x_2\}, \{1{,}0/y_1, 1{,}0/y_2, 1{,}0/y_3, 1{,}0/y_4, 0{,}5/y_5\}\rangle$$
$$C_1 = \langle \{1{,}0/x_1, 1{,}0/x_2\}, \{1{,}0/y_1, 0{,}5/y_2, 0{,}5/y_3, 1{,}0/y_4, 0{,}5/y_5\}\rangle$$
$$C_2 = \langle \{0{,}5/x_1, 0{,}5/x_2, 0{,}5/x_3\}, \{0{,}5/y_1, 0{,}5/y_2, 1{,}0/y_3, 1{,}0/y_4, 1{,}0/y_5\}\rangle$$
$$C_3 = \langle \{0{,}5/x_1, 0{,}5/x_2\}, \{1{,}0/y_1, 1{,}0/y_2, 1{,}0/y_3, 1{,}0/y_4, 1{,}0/y_5\}\rangle$$
$$C_4 = \langle \{1{,}0/x_1, 1{,}0/x_2, 0{,}5/x_3\}, \{0{,}5/y_1, 0{,}5/y_2, 0{,}5/y_3, 1{,}0/y_4, 0{,}5/y_5\}\rangle$$
$$C_5 = \langle \{1{,}0/x_1, 0{,}5/x_2, 1{,}0/x_3\}, \{0{,}5/y_3, 0{,}5/y_4, 1{,}0/y_5\}\rangle$$
$$C_6 = \langle \{1{,}0/x_1, 0{,}5/x_2, 0{,}5/x_3\}, \{0{,}5/y_1, 0{,}5/y_2, 0{,}5/y_3, 1{,}0/y_4, 1{,}0/y_5\}\rangle$$
$$C_7 = \langle \{0{,}5/x_1, 1{,}0/x_2, 0{,}5/x_3\}, \{0{,}5/y_1, 0{,}5/y_2, 1{,}0/y_3, 1{,}0/y_4, 0{,}5/y_5\}\rangle$$
$$C_8 = \langle \{1{,}0/x_1, 0{,}5/x_2\}, \{1{,}0/y_1, 0{,}5/y_2, 0{,}5/y_3, 1{,}0/y_4, 1{,}0/y_5\}\rangle$$
$$C_9 = \langle \{1{,}0/x_1, 1{,}0/x_2, 1{,}0/x_3\}, \{0{,}5/y_3, 0{,}5/y_4, 0{,}5/y_5\}\rangle$$

El diagrama de Hasse de este retículo aparece en la Figura 1.

En el siguiente ejemplo se considera una matriz asociada a la relación de orden mayor y se considera una granulación del intervalo $[0, 1]$ en 10 subintervalos, es decir, $[0, 1]_{10} = \{0, 0{,}1, 0{,}2, \ldots, 0{,}9, 1\}$

Ejemplo 2 *Sea* $(L, \preceq, \&_L)$ *el marco multi-adjunto en el que* $L = \{0, 0{,}1, 0{,}2, \ldots, 0{,}9, 1\}$ *y* $\&_L$ *es el conjuntor de Łukasiewicz definido sobre* L*. Fijado este marco, el contexto viene dado por la tupla* (A, B, R)*, donde* $A = \{y_1, y_2, y_3, y_4, y_5\}$*,* $B = \{x_1, x_2, x_3, x_4\}$ *y* $R\colon A \times B \to L$ *se define según la Tabla 2. En este ejemplo, hay considerablemente más subconjuntos difusos de objetos,* $11^4 = 14\,641$*, así como subconjuntos difusos de atributos,* $11^5 = 161\,051$*, por lo tanto, obtenemos un aumento notable en el número de atributos.*

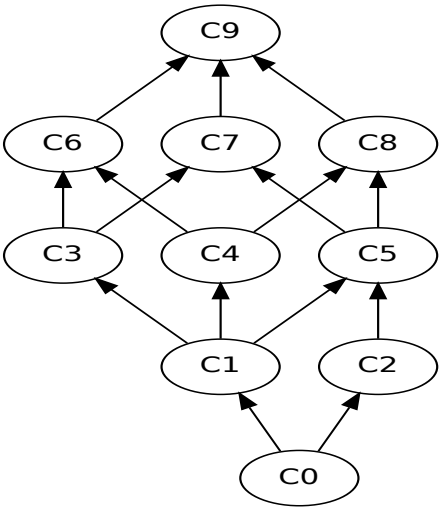


Figura 1: Diagrama de Hasse de $(\mathcal{M}_L, \preceq)$

Cuadro 2: Relación R del Ejemplo 1.

R	x_1	x_2	x_3	x_4
y_1	0,3	0,2	0,7	0,8
y_2	0,5	0,4	0,4	0,9
y_3	0,7	0,3	0,2	0,7
y_4	0,9	0,6	0,1	0,2
y_5	0,8	0,5	0,6	0,4

Concretamente, se obtienen 384 *conceptos, que representan un* $2{,}6\,\%$ *del total de subconjuntos de objetos difusos.*

El diagrama de Hasse del retículo correspondiente se ha representado en la Figura 2, aunque para una observación óptima se necesitaría representar en una escala mayor.

Si en los ejemplos anteriores se considera la conjunción de Gödel en lugar de la de Łukasiewicz, el número de conceptos es menor, en el contexto del Ejemplo 1 salen 8 conceptos, y en el contexto del Ejemplo 1 salen 117. Como trabajo futuro se estudiará la relación que pueda existir entre estos retículos, cambiando únicamente los conjuntores utilizados.

4. CONCLUSIONES Y TRABAJO FUTURO

Se ha realizado una implementación del algoritmo dado en [5] utilizando el lenguaje de programación *Python*, para esto se han utilizado estrategias de optimización como maximizar el uso de la aritmética de enteros, eliminando el uso de operaciones en punto flotante y disminuyendo así el tiempo de cálculo. Además, se

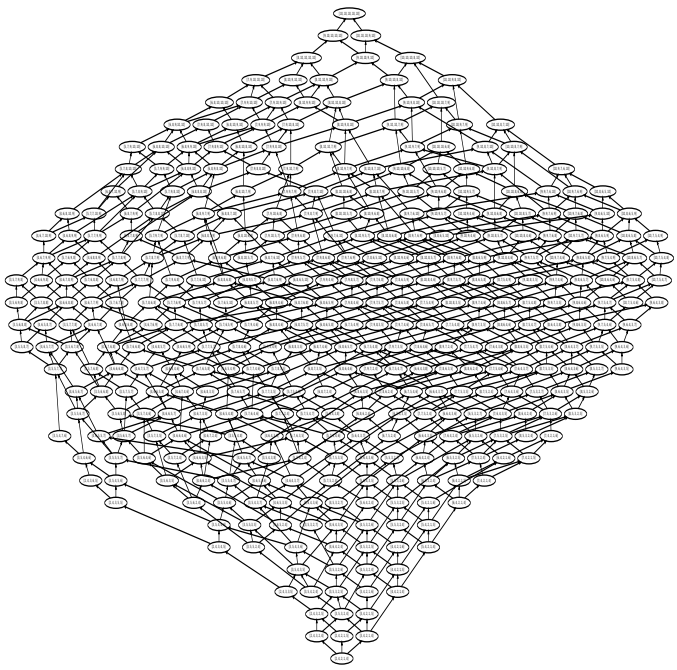


Figura 2: Diagrama de Hasse de $(\mathcal{M}, \preceq)$

ha utilizado la biblioteca *Scipy* para las operaciones matriciales y se han experimentado algunas variantes para la paralelización de los algoritmos.

De cara al futuro, nos proponemos implementar otros algoritmos, así como mejorar el existente, y comparar el código que hemos desarrollado con otras implementaciones. Además, se analizarán las zonas críticas de cara al tiempo de cálculo para mejorarlas mediante su implementación como bibliotecas C/C++ y su paralelización.

Agradecimientos

Partially supported by Junta de Andalucía grant P09-FQM-5233, and by the EU (FEDER), and the Spanish Science and Education Ministry (MEC) under grant TIN2009-14562-C05-03.

Referencias

[1] A. Abdel-Hamid and N. Morsi. Associatively tied implicacions. *Fuzzy Sets and Systems*, 136(3):291–311, 2003.

[2] R. Bělohlávek. Lattice generated by binary fuzzy relations (extended abstract). In *4th Intl Conf on Fuzzy Sets Theory and Applications*, page 11, 1998.

[3] R. Bělohlávek. Lattices of fixed points of fuzzy Galois connections. *Mathematical Logic Quartely*, 47(1):111–116, 2001.

[4] R. Bělohlávek, B. D. Baets, J. Outrata, and V. Vychodil. Lindig's algorithm for concept lattices over graded attributes. *Lecture Notes in Computer Science*, 4617:156–167, 2007.

[5] R. Bělohlávek, B. D. Baets, J. Outrata, and V. Vychodil. Computing the lattice of all fixpoints of a fuzzy closure operator. *IEEE Transactions on Fuzzy Systems*, 18(3):546–557, 2010.

[6] A. Burusco and R. Fuentes-González. Construction of the L-fuzzy concept lattice. *Fuzzy Sets and Systems*, 97(1):109–114, 1998.

[7] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.

[8] G. Georgescu and A. Popescu. Concept lattices and similarity in non-commutative fuzzy logic. *Fundamenta Informaticae*, 53(1):23–54, 2002.

[9] G. Georgescu and A. Popescu. Non-dual fuzzy connections. *Arch. Math. Log.*, 43(8):1009–1039, 2004.

[10] S. Krajči. A generalized concept lattice. *Logic Journal of IGPL*, 13(5):543–550, 2005.

[11] C. Lindig. Fast concept analysis. In G. Stumme, editor, *Working with Conceptual Structures-Contributions to ICCS 2000*, pages 152–161, 2000.

[12] J. Medina. Overcoming non-commutativity in multi-adjoint concept lattices. *Lect. Notes in Computer Science*, 5517:279–286, 2009.

[13] J. Medina. Multi-adjoint viewpoint on property/object-oriented concept lattices. In *III Simposio sobre Lógica Fuzzy y Soft Computing (LFSC 2010)*, pages 398–403, 2010.

[14] J. Medina. Towards multi-adjoint property-oriented concept lattices. *Lect. Notes in Artificial Intelligence*, 6401:159–166, 2010.

[15] J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.

[16] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Relating generalized concept lattices with concept lattices for non-commutative conjunctors. *Applied Mathematics Letters*, 21(12):1296–1300, 2008.

[17] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets and Systems*, 160(2):130–144, 2009.

[18] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. On multi-adjoint concept lattices: denition and representation theorem. *Lect. Notes in Artificial Intelligence*, 4390:197–209, 2007.

[19] N. N. Morsi and A. A. Fahmy. On generalized modus ponens with multiple rules and a residuated implication. *Fuzzy Sets and Systems*, 129(2):267–274, 2002.

[20] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival, editor, *Ordered Sets*, pages 445–470. Reidel, 1982.

[21] W.-X. Zhang, J.-M. Ma, and S.-Q. Fan. Variable threshold concept lattices. *Information Sciences*, 177(22):4883–4892, 2007.

Cardinality of IF-Sets

Patricia Díaz[1], **Tania Iglesias**[1], **Vladimír Janiš**[2], **Susana Montes**[1]

[1] *University of Oviedo, Spain*, {diazdpatricia,iglesiasctania,montes}@uniovi.es
[2] *Faculty of Sciences, Matej Bel University, Banská Bystrica, Slovak Republic*, vladimir.janis@umb.sk

Abstract

An axiomatic approach to scalar cardinalities of intuitionistic fuzzy sets is presented. It is based on a system of four simple postulates. A characterization theorem for those cardinalities is given. Moreover, some interesting examples and properties are provided.

Keywords: Intuitionistic fuzzy set, cardinality, inclusion-exclusion principle.

1 Introduction

Cardinality is a very important characteristic of a crisp set. For this reason, several authors (see review and references in [7, 15, 17]) have considered a similar characteristic also for fuzzy sets. There was rather theoretical motivation for dealing with the notion of the cardinality of fuzzy sets. Nevertheless, measuring the cardinality of fuzzy sets had also many applications, especially in the case of finite fuzzy sets. For instance, communication with databases and information/intelligent systems, modeling the meaning of imprecise quantifiers in natural language statements, computing with words, the computational theory of perceptions, decision-making in a fuzzy environment, probabilities of fuzzy events, metrical analysis of grey images, etc. (see, e.g., [7, 10, 14, 21]). Following these approaches we get a single number, an interval or a fuzzy set as a cardinality of fuzzy sets. In many applications one prefers a simple scalar approximation of the cardinality of fuzzy sets. The complete axiomatic theory of scalar cardinality of fuzzy sets can be found in Wygralak's book ([18]). The axiomatic theory of fuzzy cardinality was introduced by Casasnovas and Torrens ([5]).

However, there are situations when the cardinality of a fuzzy set does not provide the required information. Suppose for example the following (simplified but clear) case:

Example 1. *A supermarket owner decides between two types of a "shopping music", denoted by A and B. Six randomly chosen customers are asked to express their grade of acceptance on a scale from 0 to 1. Suppose the result* $(0.2;0.3;0.3;0.7;0.3;0.2)$ *for A and the result* $(0.8;1;0.2;0.1;0;0.1)$ *for B. If we consider these results as a pair of fuzzy sets representing the people that accept one or other option, then it is quite natural to compare sums of the values in both cases (which is, in fact one of cardinalities for fuzzy sets). In this example the sum is 2 for the alternative A and 2.2 for B, thus B should be preferred.*

On the other hand, suppose an additional question to the same set of customers: Express the grade of your negative attitude to a given type of music and suppose the following responses: $(0;0;0;0;0;0)$ *for A and* $(0.1;0;0;0.9;1;0.9)$ *for B. Now if we compare the set of people that either prefer certain option or cannot decide about their attitude, we see that the sums of respective values are 6 for A, but only 3.1 for B.*

The aim of the previous example is ti show that in practical situations we may estimate the grade of the membership and the grade of nonmembership to a certain vaguely defined collection only with some degree of hesitation, which cannot be always ignored.

We introduce here an axiomatic definition of the scalar cardinality for IF sets as a generalization of cardinality of fuzzy sets (see, for instance, [9]). In particular, we will consider the class of cardinalities fulfilling the inclusion-exclusion principle. For this class, we will provide a characterization, study its properties and present some examples.

This communication is organized as follows: Section 2 is devoted to introduce the most important concepts on the theory of intuitionistic fuzzy sets and the cardinality of fuzzy sets. Then, in Section 3, we define and characterize the cardinality in a intuitionistic fuzzy framework. In Section 4 we present some properties and examples of this concept. Section 5 brings some concluding remarks and open problems.

2 Basic concepts

2.1 Intuitionistic Fuzzy Sets

Atanassov intuitionistic fuzzy sets were introduced as a generalization of fuzzy sets that are useful to model situations in which a lack of information is presented (see [1, 2, 3]). While fuzzy sets allow us a freedom degree (the membership degree of every point to the set), IF-sets allow us two freedom degrees: a membership degree and a non-membership degree. Given a finite universe Ω, an IF-sets A on Ω is defined by:

$$A = \{(w, \mu_A(w), \nu_A(w)) \mid w \in \Omega\},$$

where μ_A and ν_A are two functions

$$\mu_A, \nu_A : \Omega \to [0,1]$$

that fulfill the natural restriction:

$$0 \leq \mu_A(w) + \nu_A(w) \leq 1.$$

Thus, $\mu_A(w)$ denotes the membership degree of the element w to the sets A and $\nu_A(w)$ denotes the non-membership degree of the element w to A. In this way, we can define another function $\pi_A : \Omega \to [0,1]$, called the intuitionistic fuzzy index or the hesitation index, that is defined by:

$$\pi_A(w) = 1 - \mu_A(w) - \nu_A(w),$$

that expresses the lack of knowledge of whether w belongs to A or not. From now on, $IFSs(\Omega)$ denotes all the IF-sets that can be defined on Ω, and $FS(\Omega)$ denotes all the fuzzy sets that can be defined on Ω. We have to note that, when the IF-index π_A always takes the value 0, then

$$\mu_A(w) = 1 - \nu_A(w), \text{ for every } w \in \Omega.$$

Consequently, A is in particular a fuzzy set. Thus, every fuzzy set is an IF-set. We conclude that $FS(\Omega) \subseteq IFSs(\Omega)$. Moreover we know that any crisp set is a fuzzy set, and therefore, an IF-set. In particular, referential Ω and the empty set are elements of $IFS(\Omega)$, since they can be represented as IF-sets as follows $\Omega = \{(w,1,0) \mid w \in \Omega\}$ and $\emptyset = \{(w,0,1) \mid w \in \Omega\}$.

Since intuitionistic fuzzy sets allow two degrees of freedom into a set description, and fuzzy sets only allow one, this generalization gives us an additional possibility to represent the lack of information what leads when we try to describe many real problems. Thus, they are very convenient to model situations where human answers are present: *yes*, *no* or *does not apply*. A good example of these kind of situations is voting, since human voters can be divided into three groups: *vote for*, *vote against* or *abstain* (see for example [13]).

In fuzzy sets theory many operations between fuzzy sets can be applied, like unions, intersections, complements, and the inclusion relation can be defined. These operations have been generalized in the literature for IF-sets.

Definition 2. *Given two fuzzy sets $A, B \in IFSs(\Omega)$, the following operations are defined.*

- *Union of A and B:*

$$A \cup B = \{(w, \mu_{A\cup B}(w), \nu_{A\cup B}(w)) \mid w \in \Omega\},$$

  where $\mu_{A\cup B}(w) = \max\{\mu_A(w), \mu_B(w)\}$ *and* $\nu_{A\cup B}(w) = \min\{\nu_A(w, \nu_B(w)\}$.

- *Intersection of A and B:*

$$A \cap B = \{(w, \mu_{A\cap B}(w), \nu_{A\cap B}(w)) \mid w \in \Omega\},$$

  where $\mu_{A\cap B}(w) = \min\{\mu_A(w), \mu_B(w)\}$ *and* $\nu_{A\cap B}(w) = \max\{\nu_A(w), \nu_B(w)\}$.

- *Complement of A:*

$$A^c = \{(w, \nu_A(w), \mu_A(w)) \mid w \in \Omega\}.$$

- *A is a subset of B (denoted by $A \subseteq B$) if and only if for every $w \in \Omega$ it holds that:*

$$\mu_A(w) \leq \mu_B(w) \text{ and } \nu_A(w) \geq \nu_B(w).$$

We have considered the initial definitions given by Atanassov ([1, 2]), since they have been the most usually considered in the literature. However, let us recall that some generalizations based on t-norms and t-conorms (see [8]) have already been presented (see for example [6]). Moreover, we have to recall that, when A and B are fuzzy sets, the previous definitions coincide with the operations between fuzzy sets.

2.2 Cardinality of a fuzzy set

Classically, the cardinality of a crisp set is a measure of the "number of elements of the set". For example, the set $A = \{2,4,6\}$ contains 3 elements, and therefore A has a cardinality of 3. The importance of this characteristic has as a consequence that there exist several approaches to the notion of cardinality of a fuzzy set, as we commented in the introduction. Sometimes, a natural or real number is needed as a simple evaluation of the cardinality. An axiomatic definition of scalar cardinalities of fuzzy sets was proposed in [16]. It brings together the most commonly used scalar cardinalities of fuzzy sets. In order to present this definition, we need to fix the following notation: let $x \in [0,1]$ and $a \in \Omega$, x/a denotes the fuzzy set with membership function:

$$\mu_{x/a}(w) = \begin{cases} x & \text{if } w = a, \\ 0 & \text{otherwise.} \end{cases}$$

With this notation, the definition of scalar cardinality of a fuzzy set is obtained by means of quite natural postulates.

Definition 3. *A function $sc : FS(\Omega) \to [0,\infty)$ is called a scalar cardinality if the following conditions are fulfilled:*

1. *Singleton monotonicity: for any $x,y \in [0,1]$ with $x \leq y$ and any $w,w' \in \Omega$, it holds that $sc(x/w) \leq sc(y/w')$.*

2. *Finite additivity: for any finite family $\{A_i\}_{i\in I} \subset FS(\Omega)$ such that $A_i \cap A_j = \emptyset$ if $i \neq j$, it holds that:*

$$sc\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} sc(A_i).$$

3. *Coincidence: for any $w \in \Omega$, it holds that $sc(1/w) = 1$.*

These requirements are quite natural. Condition 1 says that scalar cardinalities of singletons should be increasing with respect to their unique positive membership degrees, whereas Condition 2 shows how to calculate scalar cardinalities of more complex fuzzy sets. Finally, Condition 3 says the scalar cardinality of a crisp set coincides with the ordinary cardinality of that sets and together with Condition 2, it implies that for any crisp set A, $sc(A) = |\{w \in \Omega \mid w \in A\}|$.

In the following theorem we can see a simple characterization of scalar cardinalities of fuzzy sets, which will be useful in next Section.

Theorem 4. *[16] Let Ω be a finite universe. A map $SC : FS(\Omega) \to [0,\infty)$ is a scalar cardinality on $FS(\Omega)$ iff*

$$sc(A) = \sum_{w\in\Omega} f(\mu_A(w))$$

for any $A \in FS(\Omega)$, where $f : [0,1] \to [0,1]$ is a function such that

1. *$f(0) = 0$, $f(1) = 1$,*
2. *f is increasing.*

3 Cardinality of IF-sets

In this section we are going to introduce and study the notion of cardinality for IF-sets.

Definition 5. *Let Ω be a finite universe. A map $card : IFSs(\Omega) \to [0,\infty)$ is called cardinality if:*

1. *card is a scalar cardinality of fuzzy sets if we restrict its domain to $FS(\Omega)$.*
2. *$card(A) \leq card(B)$ if $A \subseteq B$ with $A,B \in IFSs(\Omega)$.*

Let us remark that the previous definitions is intuitive. On one hand, it is obvious that the new definition should coincide with the definition considered for fuzzy sets, since any fuzzy set is a particular case of an IF-set. On the other hand, it seems reasonable that, if a set is included in other one, then its cardinality is lower than or equal to.

In combinatorics, the inclusion-exclusion principle (also known as the sieve principle) is an equation relating the sizes of two sets and their union. It states that if A and B are two (finite) sets, then

$$card(A \cup B) = card(A) + card(B) - card(A \cap B)$$

The meaning of the statement is that the number of elements in the union of the two sets is the sum of the elements in each set, respectively, minus the number of elements that are in both. As we commented in the Introduction, our purpose is to characterize the cardinality maps for $IFSs(\Omega)$ fulfilling this principle.

Theorem 6. *Let Ω be a finite universe and let Δ be the subset of the unit square defined by $\Delta = \{(x,y) \in [0,1]^2 \mid x+y \leq 1\}$. A map $card : IFSs(\Omega) \to [0,1]$ is a cardinality fulfilling the inclusion-exclusion principle iff there exists a map $\phi : \Delta \to [0,1]$ with*

1. *$\phi(0,1) = 0$,*
2. *$\phi(1,0) = 1$,*
3. *ϕ is increasing in the first component and decreasing in the second one,*
4. *$\phi(x,t) + \phi(z,y) = \phi(x,y) + \phi(z,t)), \forall (x,y),(z,t) \in \Delta$ with $x \vee y \leq z \wedge t$ or $x \wedge y \leq z \wedge t \leq x \vee y \leq z \vee t$,*

such that

$$card(A) = \sum_{w\in\Omega} \phi(\mu_A(w), \nu_A(w))$$

for any $A \in IFSs(\Omega)$.

As a consequence of this characterization, we can see that the greater the membership degree of w to A, the greater its cardinality is, and the greater the non-membership degree of w to A, the lower its cardinality is.

Now we should study is if the inclusion-exclusion principle is fulfilled by all the cardinality functions for IF-sets. Unfortunately, the answer is negative. To see that, it is enough to consider the following counterexample.

Example 7. *Let Ψ be a map defined by:*

$$\Psi(\alpha,\beta) = \alpha(1-\beta)$$

on a universe $\Omega = \{w\}$. It is immediate that Ψ is a cardinality on $IFSs(\Omega)$. However, if we consider the IF-sets A and B defined by:

$$A = \{(x,0.2,0.4)\},$$
$$B = \{(x,0.3,0.5)\},$$

we obtain that $card(A) = 0.12$ and $card(B) = 0.15$. Moreover, it we compute their union and intersection, we obtain the following IF-sets:

$$A \cap B = \{(x,0.2,0.5)\},$$
$$A \cup B = \{(x,0.3,0.4)\},$$

for which $card(A\cap B)=0.1$ *and* $card(A\cup B)=0.18$. *Then, we obtain that:*

$$\begin{aligned} card(A\cup B) &= 0.18 \neq 0.17 = 0.12+0.15-0.1 \\ &= card(A)+card(B)-card(A\cap B), \end{aligned}$$

and therefore Ψ *is an example of a cardinality not-fulfilling the inclusion-exclusion principle.*

Thus, there are cardinalities that does not satisfy such principle. The question we should answer now is if there are cardinalities satisfying this principle. We give the answer in the following result.

Theorem 8. *Let* Ω *be a finite universe. If there exists two maps* $f,g:[0,1]\to[0,1]\,|\,f(0)=g(1)=0, f(1)+g(0)=1$, *being* f *an increasing function and* g *a decreasing function, the map* $card:IFSs(\Omega)\to[0,\infty)$ *defined by*

$$card(A)=\sum_{x\in X} f(\mu_A(x))+g(\nu_A(x))$$

for any $A\in IFSs(\Omega)$, *is a cardinality fulfilling the inclusion-exclusion principle.*

In the following result we present a way of build cardinalities from another cardinalities.

Proposition 9. *Let* $card_1,\ldots,card_n$ *be n cardinalities for IF-sets. Let us also consider a function* $f:[0,1]^n\to[0,1]$ *satisfying the following properties:*

- *If* $\vec{0}=(0,\overset{n}{\ldots},0)$, *then* $f(\vec{0})=0$.
- *If* $\vec{1}=(1,\overset{n}{\ldots},1)$, *then* $f(\vec{1})=1$.
- *If we consider* $\vec{a}=(a_1,\ldots,a_n)$ *and* $\vec{b}=(b_1,\ldots,b_n)$, *satisfying* $a_i+b_i\leq 1$ *and* $a_i\leq b_i$ *for* $i=1,\ldots,n$, *it holds that* $f(\vec{a})\leq f(\vec{b})$.

Then, the function card defined by:

$$card(A)=$$

$$\sum_{w\in\Omega} f(card_1(A_{\mu_A(w)\nu_A(w)}),\ldots,card_n(A_{\mu_A(w)\nu_A(w)}))$$

for every $A\in IFSs(\Omega)$ *is another cardinality for IF-sets.*

This proposition shows us a way of building cardinalities from another cardinalities. In addition, if every cardinality $card_i$ satisfies the inclusion-exclusion principle, it is possible that *card* also satisfies it.

Corollary 10. *In the conditions of the previous result, if every* $card_i$ *satisfies the inclusion-exclusion principle and*

$$f(\vec{a})=\sum_{i=1}^{n}\omega_i a_i,$$

where $0\leq\omega_i\leq 1$ *and* $\sum_{i=1}^{n}\omega_i=1$, *then card also satisfies the inclusion-exclusion principle.*

Some properties of cardinalities are presented in the following proposition.

Proposition 11. *Let* Ω *be a finite universe and let card be a cardinality on* $IFSs(\Omega)$ *fulfilling the inclusion-exclusion principle. The following properties are satisfied:*

1. *Given A and B in* $IFSs(\Omega)$, *if there exists a bijection* $h:\Omega\to\Omega$ *such that*

   $$\mu_A(w)=\mu_B(h(w)) \text{ and } \nu_A(w)=\nu_B(h(w)),\forall w\in\Omega$$

   then $card(A)=card(B)$.

2. $|core(A)|\leq card(A)\leq|supp(A)|,\forall A\in IFSs(\Omega)$, *where* $core(A)=\{w\in\Omega|\mu_A(w)=1 \text{ and } \nu_A(w)=0\}$ *and* $supp(A)=\{w\in\Omega|\mu_A(w)>0\}$.

3. $card(\cup_i A_i)\leq\sum_i card(A_i)$.

4 Examples of cardinality of IF-sets

In this section we show several examples of cardinalities. We also study if they satisfy or not the inclusion-exclusion principle. For this aim, we will use some of the result presented in the previous section.

Example 12.

- *Let us consider the function* $card_1$ *defined by:*

  $$card_1(A)=|\Omega|-\sum_{w\in\Omega}\nu_A(w)$$

  for any $A\in IFSs(\Omega)$. *From Theorem 6 it is a cardinality for IF-sets with the inclusion-exclusion principle, since in that case* $\phi(x,y)=1-y$, $\forall(x,y)\in\Delta$.

  This cardinality is called the biggest cardinality or max-sigma-count (see [19, 20, 12]).

- *Let us now consider the function* $card_2$ *defined by:*

  $$card_2(A)=\sum_{w\in\Omega}\mu_A(w),$$

  for any $A\in IFSs(\Omega)$. *In that case* $\phi(x,y)=x$, $\forall(x,y)\in\Delta$ *and therefore, from Theorem 6,* $card_2$ *is a cardinality for IF-sets fulfilling the inclusion-exclusion principle.*

  This cardinality is knows as the least ("sure") cardinality or min-sigma-count (see [19, 20, 12]).

- *Let us now consider* $a,b\in[0,1]$ *such that* $a+b=1$. *Then, using Proposition 9, the function*

  $$card(A)=a\cdot card_1(A)+b\cdot card_2(A)$$

  for every $A\in IFSs(\Omega)$ *is a cardinality, and using Corollary 10, it also holds the inclusion-exclusion principle.*

In this last example we connect the theory of cardinalities for IF-sets with the theory of the measures of comparison of IF-sets.

Example 13. *Let us consider an IF-divergence for IF-sets D_{IF}. Let us recall that an IF-divergence is a function*

$$D_{\mathrm{IF}} : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$$

that satisfies the following properties:

- $D_{\mathrm{IF}}(A,A) = 0$ *for every* $A \in IFSs(\Omega)$.
- D_{IF} *is a symmetric function.*
- $D_{\mathrm{IF}}(A \cap C, B \cap C) \leq D_{\mathrm{IF}}(A,B)$ *for every* $A,B,C \in IFSs(\Omega)$.
- $D_{\mathrm{IF}}(A \cup C, B \cup C) \leq D_{\mathrm{IF}}(A,B)$ *for every* $A,B,C \in IFSs(\Omega)$.

These types of functions are used to measure the difference between IF-sets (see [11]). Let us assume that the IF-divergence satisfies the following property: $D_{\mathrm{IF}}(\emptyset, X) = 1$. *Then, the functions* $card_3$ *defined by:*

$$card_3(A) = D_{\mathrm{IF}}(A^c, \Omega)$$

and

$$card_4(A) = D_{\mathrm{IF}}(A, \emptyset)$$

for every $A \in IFSs(\Omega)$ *are cardinalities for IF-sets. However, in general they do not fulfil the inclusion-exclusion principle.*

5 Conclusions

The presented axiomatic approach generates a wide family of possible cardinalities of IF-sets, which have as a particular case the scalar cardinalities for fuzzy sets. A characterization of this family is showed and several properties and examples are considered.

As a future work, we would like to be able to extend this communication in several directions: by considering a general t-norm and t-conorm for defining the intersection and union of IF-sets; by measuring the cardinality by means of interval-valued set and by working on non-countable universes.

Acknowledgements

The research in this paper is partly supported by the Scientific Technical Facilities of University of Oviedo, by the Agency of the Slovak Ministry of Education for the Structural Funds of the EU, under project ITMS:26220120007, grant 1/0297/11 by Slovak grant agency VEGA and the Spanish Ministry of Science and Innovation grant MTM2010-17844.

References

[1] K. Atanassov: Intuitionistic Fuzzy Sets. In *proceedings of VII ITKR*, Sofia, 1983.

[2] K. Atanassov: Intuitionistic Fuzzy Sets. *Fuzzy Sets and Systems*, 20, pp. 87–96, 1986.

[3] K. Atanassov: *Intuitionistic Fuzzy Sets: Theory and Applications*, Physica-Verlang, Wyrzburg, 1999.

[4] L. C. Atanoassova: Remark on the cardinality of the intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 75, pp. 399–400, 1995.

[5] M. Casasnovas, J. Torrens: An axiomatic approach to fuzzy cardinalities of finite fuzzy sets. *Fuzzy Sets and Systems* 133, pp. 193–209, 2003.

[6] G. Deschrijver, E. Kerre: A generalization of operators on intuitionistic fuzzy sets using triangular norms and conorms. *Notes on IFS* 1, pp. 19–27, 2002.

[7] D. Dubois, H. Prade: Fuzzy cardinality and the modeling of imprecise quanti/cation. *Fuzzy Sets and Systems* 16, pp. 199–230, 1985.

[8] E. P. Klement, R. Mesiar, E. Pap: *Triangular Norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[9] P. Král: Scalar Cardinalities of If Sets. *Acta Univ. M. Belii Math.* 12, pp. 67–77, 2005.

[10] Y. Liu, E.E. Kerre: An overview of fuzzy quantifiers. Part I: interpretation. *Fuzzy Sets and Systems* 95, pp. 1–21, 1998.

[11] I. Montes, V. Janiš, S. Montes: An axiomatic definition of divergence for intuitionistic fuzzy sets. In: *Proceedings of the 2011 EUSFLAT conference*, Aix-Les-Bains (French), 2011.

[12] E. Szmidt, J. Kacprzyk: Entropy for intuitionistic fuzzy sets. *Fuzzy sets and Systems*, 118, pp. 467–477, 2001.

[13] E. Szmidt, J. Kacprzyk: Analysis of Similarity Measures for Atanassov's Intuitionistic Fuzzy Sets. In J. P. Carvalho, D. Dubois, U. Kaymak and J. M. C. Sousa, editors, proceedings of the *International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference* (IFSA-EUSFLAT 2009), 1416–1421, 2009.

[14] M. Wygralak: *Vaguely Defined Objects. Representations, Fuzzy Sets and Nonclassical Cardinality Theory*. Kluwer Academic Publishers, Dordrecht, 1996.

[15] M. Wygralak: Questions of cardinality of finite fuzzy sets. *Fuzzy Sets and Systems* 102, pp. 185–210, 1999.

[16] M. Wygralak: An axiomatic approach to scalar cardinalities of fuzzy sets. *Fuzzy Sets and Systems* 110, pp. 175–179, 2000.

[17] M. Wygralak: Fuzzy sets with triangular norms and their cardinality theory. *Fuzzy Sets and Systems* 124, pp. 1–24, 2001.

[18] M. Wygralak: *Cardinalities of Fuzzy Sets*, Studies in Fuzziness and Soft Computing 118 Springer-Verlag, Berlin, Heidelberg, 2003.

[19] L.A. Zadeh: A computational approach to fuzzy quantifiers in natural languages. *Comput. Math. Appl.* 9 (1), pp. 149–184, 1983.

[20] L.A. Zadeh: The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems* 11, pp. 199–227, 1983.

[21] L.A. Zadeh: From computing with numbers to computing with words - From manipulation of measurements to manipulation of perceptions. *IEEE Trans. Circuits Systems - I* 45, pp. 105–119, 1999.

T-convexity for lattice-valued fuzzy sets

Tania Iglesias[1], **Ignacio Montes**[1], **Vladimir Janiš**[2], **Susana Montes**[1]

[1]*University of Oviedo, Spain*, {iglesiasctania,imontes,montes}@uniovi.es
[2]*Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, Slovak Republic*, vladimir.janis@umb.sk

Abstract

The concept of quasiconvexity for fuzzy sets is studied, as it enables to be extended also for fuzzy sets with values in bounded lattices. We develop the study of aggregation operators preserving quasiconvexity taking into account a given triangular norm representing the intersection of fuzzy sets or the conjunction in the underlying model of fuzzy logic. A necessary condition for an aggregation operator to preserve such type of a generalized convexity is provided.

Keywords: Quasiconvexity, convexity, t-norm, aggregation operator.

1 Introduction

There is no doubt that convexity is one of the most important geometrical aspects in various mathematical areas, mostly in optimization theory (see for example [5, 9, 12]). There is a lot of various versions of convexity and its generalizations, see e.g. [10]. Such attitude can be found also in the framework of fuzzy set theory in [1, 11], where applications in fuzzy optimization are studied as well.

Our starting point is the notion of quasiconvexity. As we will explain, in certain aspects it reflects better the nature of the convexity idea and can be used also for lattice valued fuzzy sets ([4]).

The practical applications of fuzzy models use various types of fuzzy logic connectives. These are based in the first place on the chosen triangular norm. Therefore we formulate a weak form of quasiconvexity based on a given triangular norm. The main goal of this contribution is to formulate conditions, under which an aggregation of weakly quasiconvex sets remains weakly quasiconvex.

This communication is organized as follows: in the second section we present basic notions about convexity and quasiconvexity of fuzzy sets, t-norms and aggregation operators. The generalization of quasiconvesity to T-convexity is given in Section 3. The main results are in Section 4 followed by some examples illustrating these results. Finally, in the last section, we present some concluding remarks and we comment some open problems.

2 Preliminaries

We introduce the basic notions that will be used in the rest of the work. We are going to work with fuzzy sets that are defined on $\mathbb{R}^n$. These fuzzy sets will be considered L-valued, where $(L,\wedge,\vee,0_L,1_L)$ is a bounded lattice, in the sense proposed by Goguen ([3]). In particular, $[0,1]$ is a bounded lattice, and consequently usual fuzzy sets are also L-valued fuzzy sets.

In the rest of this section, we will introduce some notions about convexity and quasiconvexity of L-valued fuzzy sets. Also we introduce t-norms and aggregation operators for L-valued fuzzy sets.

2.1 Convexity and quasiconvexity of fuzzy sets

If we consider a [0,1]-valued fuzzy set $\mu:\mathbb{R}^n\to[0,1]$, it is said to be convex if it satisfies the following property:

$$\mu(\lambda x+(1-\lambda)y)\geq\lambda\mu(x)+(1-\lambda)\mu(y)$$

for every $x,y\in\text{supp}(\mu)$ and $\lambda\in[0,1]$, where $\text{supp}(\mu)$ denotes the support of μ, that is, $\text{supp}(\mu)=\{x\in\mathbb{R}^n\mid\mu(x)>0\}$. This definition of convexity for real valued fuzzy set is a simple generalization of the convexity for crisp sets on $\mathbb{R}^n$ in the sense that the hypograph of a convex fuzzy set is a usual convex set in $\mathbb{R}^{n+1}$. However, this definition is very restrictive as it requires a linear structure as the image set, hence it is not applicable for L-valued fuzzy sets. Moreover, convexity in such sense is not equivalent to the convexity of all the cuts (level sets) of a given fuzzy set. For these reasons the notion of quasiconvexity was introduced ([1]).

Definition 1. *A [0,1]-valued fuzzy set $\mu : \mathbb{R}^n \to [0,1]$ is quasiconvex if the following inequality is satisfied for every $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$:*

$$\mu(\lambda x + (1-\lambda)y) \geq \min\{\mu(x), \mu(y)\}.$$

Obviously, every convex fuzzy set is also quasiconvex.

Let us now consider a bounded lattice $(L, \wedge, \vee, 0_L, 1_L)$ and an L-valued fuzzy set $\mu : \mathbb{R}^n \to L$. The notion of quasiconvexity can be defined for this type of fuzzy sets.

Definition 2. *An L-valued fuzzy set μ is quasiconvex if*

$$\mu(\lambda x + (1-\lambda)y) \geq \mu(x) \wedge \mu(y)$$

for every $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Recall that for $\alpha \in L$ the α-cut of a fuzzy set μ is the set

$$(\mu)_\alpha = \{x \in \mathbb{R}^n \mid \mu(x) \geq \alpha\}.$$

It is easy to show (see [4]) that a fuzzy set is quasiconvex if and only if all its α-cuts are convex.

Let us remark that, when $L = [0,1]$ we have that $\mu(x) \wedge \mu(y) = \min(\mu(x), \mu(y))$, and consequently the definition of quasiconvexity for lattice valued fuzzy sets is coherent with the quasiconvexity for $[0,1]$ valued fuzzy sets.

2.2 Triangular norms and Aggregation Operators

In this section we are going to introduce two different binary operators, t-norms and aggregation operators. First of all, we introduce their usual definitions, and then we give their definitions for lattices.

A t-norm is a binary operator $T : [0,1] \times [0,1] \to [0,1]$ that is commutative, associative, non-decreasing and its neutral element is 1. They are very useful in the fuzzy set theory, since they allow us to make intersections between fuzzy sets and represent the conjunction in fuzzy logic. Detailed study on t-norms can be found in [6]. The three most important t-norms are the minimum operator $T_M(\alpha, \beta) = \min(\alpha, \beta)$, the algebraic product $T_P(\alpha, \beta) = \alpha \cdot \beta$ and the Łukasiewicz t-norm $T_L(\alpha, \beta) = \max(\alpha + \beta - 1, 0)$. The minimum operator is the greatest t-norm; the smallest t-norm is the drastic product defined by

$$T_D(\alpha, \beta) = \begin{cases} \min(\alpha, \beta) & \text{if } \max(\alpha, \beta) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The above t-norms can be ordered (usual ordering of functions) as follows: $T_D \leq T_L \leq T_P \leq T_M$.

On the other hand, an aggregation operator ([7]) is a non-decreasing binary operator $A : [0,1] \times [0,1] \to [0,1]$ such that $A(0,0) = 0$ and $A(1,1) = 1$. They can also be defined on a more general way (see [2]), however, we will use only binary operators.

Both t-norms and aggregation operators can be defined in a more general settings, if we consider bounded lattices $(L, \wedge, \vee, 0_L, 1_L)$.

Definition 3. *A t-norm on L is a binary operator $T : L \times L \to L$ that satisfies the following properties:*

- *$T(\alpha, T(\beta, \gamma)) = T(T(\alpha, \beta), \gamma)$ for every $\alpha, \beta, \gamma \in L$.*
- *$T(\alpha, \beta) = T(\beta, \alpha)$ for every $\alpha, \beta \in L$.*
- *If $\alpha, \beta, \gamma \in L$ such that $\alpha \leq \beta$, then $T(\alpha, \gamma) \leq T(\beta, \gamma)$.*
- *$T(\alpha_L, 1) = \alpha$ for every $\alpha \in L$.*

Definition 4. *An aggregation operator A on L is a binary operator $A : L \times L \to L$ such that*

- *$A(0_L, 0_L) = 0_L$ and $A(1_L, 1_L) = 1_L$.*
- *$A(a_1, a_2) \leq A(b_1, b_2)$ for $a_1 \leq b_1$ and $a_2 \leq b_2$.*

Note that t-norms are particular cases of aggregation operators.

3 T-convexity

We have already seen in the previous section that the notion of convexity for fuzzy sets is very restrictive. For this reason, quasiconvexity was introduced. However, this notion can also be very strong in several situations, especially in a framework of a fuzzy logic model in which a t-norm other than minimum is used. Consequently, here we present a less restrictive definition, called T-convexity.

The new concept we present here is defined by means of a t-norm.

Definition 5. *Let μ be a L-valued fuzzy set and let T be a t-norm. Then μ is T-convex if*

$$\mu(\lambda x + (1-\lambda)y) \geq T(\mu(x), \mu(y))$$

for every $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Some simple facts on T-convexity are obvious. First, if we consider the minimum operator, $T_M(x, y) = x \wedge y$, then T_M-convexity of the L-valued fuzzy set μ is equivalent to its quasiconvexity. Second, T-convexity is less restrictive than quasiconvexity. In fact, if a L-valued fuzzy set is T-convex, then it is also T^*-convex for every t-norm T^* that satisfies $T^* \leq T$. Then, in particular, if the L-valued fuzzy set is quasiconvex, it is T-convex with respect to every t-norm T.

Example 6. *Let $n = 1$ be and let L be the unit interval ($L = [0,1]$).*

- *The fuzzy set*

$$\mu_1(x) = \begin{cases} x/3 & \text{if } 0 \le x \le 3 \\ (6-x)/3 & \text{if } 3 < x \le 6 \\ 0 & \text{otherwise} \end{cases}$$

is T_M-convex and therefore, T_P-convex, T_L-convex and T_D-convex.

- *However, the fuzzy set*

$$\mu_2(x) = \begin{cases} 0.5 & \text{if } 0 \le x \le 2 \text{ or } 4 < x \le 6 \\ 0.4 & \text{if } 2 < x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

is T_P-convex, and therefore T_L-convex and T_D-convex, but it is not T_M-convex (for instance, $T_M(\mu_2(1),\mu_2(5)) = T_M(0.5,0.5) = 0.5 > \mu_2(3)$).

- *Since $T_P(0.5,0.5) = 0.25 > 0$, we have that μ_3 defined by*

$$\mu_3(x) = \begin{cases} 0.5 & \text{if } 0 \le x \le 2 \text{ or } 4 < x \le 6 \\ 0 & \text{otherwise} \end{cases}$$

is an example of T_L-convex fuzzy set, which is no T_P-convex.

- *Finally, μ_4 is a T_D-convex fuzzy set, but it is not T_L-convex, where*

$$\mu_4(x) = \begin{cases} 0.9 & \text{if } 0 \le x \le 2 \text{ or } 4 < x \le 6 \\ 0 & \text{otherwise.} \end{cases}$$

4 Properties of the T-convexity

In this section we are going to investigate some properties of T-convexity, mainly in connection to their aggregation. Such considerations are extensions of the well known fact that the intersection of two convex sets is convex, where the intersection is replaced by an arbitrary aggregation operator. Thus, we could consider any t-norm for defining the intersection, since any t-norm is a particular case of aggregation operator.

Let us consider two fuzzy sets μ and ν. We are going to aggregate them by means of an aggregation operator. In this sense, if μ and ν are two L-valued fuzzy sets, and A is an aggregation operator, we define the L-valued fuzzy set $A(\mu,\nu)$ by:

$$A(\mu,\nu)(x) = A(\mu(x),\nu(x)), \forall x \in \mathbb{R}^n.$$

The main result is formulated in the following proposition:

Proposition 7. *Let A be an aggregation operator, let T be a t-norm and two L-valued T-convex fuzzy sets μ and ν be defined on $\mathbb{R}^n$. If $A(\mu,\nu)$ is T-convex, then*

$$A(T(\alpha,\gamma),T(\beta,\delta)) \ge T(A(\alpha,\beta),A(\gamma,\delta)).$$

Any aggregation operator A commuting with a t-norm T (see [8]) fulfills the previous condition. In particular, since any t-norm commutes with itself, the above condition is fulfilled when $A = T$, that is,

Corollary 8. *The T-intersection of two T-convex fuzzy sets is always a T-convex fuzzy set, for any t-norm T.*

However, the converse implication of Proposition 7 does not hold in general, as it will be shown on the following example.

Example 9. *For T_D (the drastic product), let us consider the usual aggregation operator $A(\alpha,\beta) = \frac{1}{2}(\alpha+\beta)$ (arithmetic mean).*

Note that with this t-norm the class of weakly quasiconvex fuzzy sets consists on all the normal quasiconvex fuzzy sets (attaining the value 1 at least at one point) and of all the other fuzzy sets. Clearly A preserves weak quasiconvexity, however, for the choice $\alpha = \beta = 1, \gamma = \delta = \frac{1}{2}$ the inequality from the Proposition 7 does not hold.

As a consequence of Proposition 7 we obtain again a result that can be found in [4].

Corollary 10. *Let A be an aggregation operator and two L-valued fuzzy sets μ and ν defined on $\mathbb{R}^n$. Then, the following statements are equivalent:*

- *The fuzzy sets $A(\mu,\nu)$ is quasiconvex.*
- *$A(\alpha\wedge\gamma,\beta\wedge\gamma) = A(\alpha,\beta)\wedge A(\gamma,\delta)$.*

In the following examples we will show that the implication from the Proposition 7, that is in fact an equivalence for the minimum, cannot be inverted for other T-norms, not even in case of real valued operators with the most common triangular norms.

Example 11. *Let us showan example with the product t-norm T_P and another example with the Łukasiewics operator T_L.*

- *For the product t-norm $T_P(x,y) = xy$, let us consider the aggregation operator*

$$A(\alpha,\beta) = \begin{cases} \frac{1}{2}\alpha\beta & \text{if } \alpha\beta < 1 \\ 1 & \text{if } \alpha = \beta = 1 \end{cases}$$

To verify that A preserves T_P-convexity take $x < y < z$ and T_P-convex fuzzy sets μ,ν. We have to show that

$$A(\mu,\nu)(y) \ge A(\mu,\nu)(x)A(\mu,\nu)(z).$$

In case both terms on the right-hand side are 1, from the T_P convexity of μ and ν we obtain that the left-hand side of the inequality is also equal to 1.

In all other cases the smallest value of the left-hand side is

$$\frac{1}{2}\mu(y)\nu(y)$$

and the largest value of the right-hand side is

$$\frac{1}{2}\mu(x)\nu(x)\mu(z)\nu(z)$$

and again from the T_P-convexity of μ and ν we have the required inequality. Thus A preserves T_P-convexity. However, the choice $\alpha=\beta=1, \gamma=\delta=\frac{1}{2}$ shows, that the inequality from the Proposition 7 does not hold.

- *For the Łukasiewics t-norm $T_L(x,y)=\max\{0,x+y-1\}$, let us take the aggregation operator $A(\alpha,\beta)=\alpha\beta$.*

  A routine computation shows that A preserves T_L-convexity. The choice $\alpha=\beta=\frac{2}{3}, \gamma=\delta=\frac{1}{2}$ leads again to the violation of the inequality from the Proposition 7.

Finally, if we consider symmetric aggregation operators (the ones that satisfy $A(\alpha,\beta)=A(\beta,\alpha)$ for every $\alpha,\beta\in L$), we obtain some interesting inequalities for the class of symmetric aggregation operators preserving T-convexity.

Proposition 12. *Let $A: L\times L\to L$ be a symmetric aggregation operator on a bounded lattice $(L,\wedge,\vee,0_L,1_L)$, let T be a t-norm on L and let μ and ν be two L-valued T-convex fuzzy sets. If the L-valued fuzzy set $A(\mu,\nu)$ is T-convex, then:*

$$A(\alpha,\beta)\geq A(T(\alpha,\beta),T(\alpha,\beta))\geq T(A(\alpha,\alpha),A(\beta,\beta))$$

for every $\alpha,\beta\in L$.

As in Proposition 7, the converse does not hold in this case as well. For that it is enough to consider again the t-norm and aggregation operator of Example 9, with $\alpha=1$ and $\beta=1/2$.

As consequence of this result, we obtain the following corollary that have been already found in [4].

Corollary 13. *If A is a symmetric aggregation operator defined on a bounded lattice $(L,\wedge,\vee,0_L,1_L)$, and μ and ν are two L-valued quasiconvex fuzzy sets the following statements are equivalent.*

- *A preserves quasiconvexity, that is, $A(\mu,\nu)$ is quasiconvex.*
- *For every $\alpha,\beta\in L$ the following equalities hold:*

$$A(\alpha,\beta)=A(\alpha,\alpha)\wedge A(\beta,\beta)=A(\alpha\wedge\beta,\alpha\wedge\beta).$$

Thus we have seen, that for the quasiconvexity, we have obtained a necessary and sufficient condition for A preserving quasiconvexity. However, this condition is also necessary when we refer to T-convexity (with respect to a t-norm different than $T(x,y)=x\wedge y$).

5 Conclusions

We have introduced a version of convexity for lattice-valued fuzzy sets connected to a given triangular norm and studied preserving of such convexity under aggregation. In Proposition 7 we have shown a necessary condition for an aggregation operator to preserve this type of convexity.

The complete description of the class of all aggregation operators that preserve T-convexity for a given triangular norm T remains open, unless T is the minimum t-norm, which was described in [4].

Preserving quasiconvexity is closely connected with commuting (see [8]), in fact the class of aggregation operators preserving it is exactly the class of aggregation operators commuting with T_M. From this point of view it would certainly be interesting to study connections between aggregation operators preserving T-convexity and those commuting with T.

In the future, we would like to be able to study in detail the behaviour of T-convexity for strict and nilpotent t-norms and, in particular, T_P-convexity and T_L-convexity.

Acknowledgements

The research in this paper is partly supported by the Scientific Technical Facilities of University of Oviedo, by the Science and Education Ministry FPU grant AP2009-1034, by the Agency of the Slovak Ministry of Education for the Structural Funds of the EU, under project ITMS:26220120007, grant 1/0297/11 by Slovak grant agency VEGA, grant 1/0297/11 and the Spanish Ministry of Science and Innovation grant MTM2010-17844.

References

[1] E. Ammar, J. Metz: On the fuzzy convexity and parametric fuzzy optimization. *Fuzzy Sets and Systems* 49, pp. 135–141, 1992.

[2] T. Calvo, A. Kolesárová, M. Komorníková, R. Mesiar: *Aggregation operators, New Trends and Applications*. Physica-Verlag, Heidelberg, 2002.

[3] J.A. Goguen: *L*-fuzzy Sets. *J. Math. Anal. Appl.* 18, pp. 145–174, 1967.

[4] V. Janiš, P. Král', M. Renčová: Aggregation operators preserving quasiconvexity. *Information Science*. Submitted.

[5] K. Maity, M. Maiti: Numerical Approach of Multiobtective Optimal Control Problems in Imprecise Enviroment. *Fuzzy Optimization and Decision Making* 4, pp. 313–330, 2005.

[6] E. P. Klement, R. Mesiar, E. Pap: *Triangular norms*. Academic Publishers, Boston-London-Dordrecht, 2000.

[7] R. Mesiar, M. Komorníková: Aggregation Operators. IN: *Proceeding of the XI Conference on applied Mathematics PRIM' 96, Herceg D., Surla K. (eds.)*, Institute of Mathematics, Novi Sad, pp. 193–211, 1997.

[8] S. Saminger-Platz, R. Mesiar, D. Dubois: Aggregation operators and commuting. *IEEE Transactions on Fuzzy Systems* 15, pp. 1032–1045, 2007.

[9] J. Ramík: Duality in Fuzzy Linear Programming: Some New Concepts and Results. *Fuzzy Optimization and Decision Making* 4, pp. 25–39, 2005.

[10] I. Singer: Abstract convex analysis. Canadian Mathematical Society series of monographs and advanced texts. New York: John Wiley & Sons, Inc.. pp. xxii+491, 1997.

[11] Y. R. Syau, E. S. Lee: Fuzzy Convexity with Application to Fuzzy Decision Making. In: *Proceedings of the 42th IEEE Conference on Decision and Control*, Maui (EEUU), pp. 5221–5226, 2003.

[12] H-C. Wu: The Rarush-Kuhn-Tucker optimality conditions for multi-objective programming problems with fuzzy-valued objective functions. *Fuzzy Optimization and Decision Making* 8, pp. 1–28, 2009.

SHORT NOTE ON A DEDUCTIVE SCHEME IN FUZZY LOGIC

Laura Menéndez-Arduengo[1] **Itziar García-Honrado**[2] **Enric Trillas**[1]

[1] European Centre for Soft Computing
33600 Mieres (Asturias)
{laura.menendez, enric.trillas}@softcomputing.es
[2] Departamento de Estadística, I.O. y Didáctica de las matemáticas
Universidad de Oviedo.
garciaitziar@uniovi.es

Abstract

The paper is devoted to the study of a particular deductive scheme, involving conditionals and intersections. Given two three premises p, q and r if p and q implies r and it is added the evidence of p, it must be deduced that q implies r. The study is done in different algebraic structures with diverse translations of conditionals functions, focussing in the structure of Standard Algebras of fuzzy sets.

Key words: deductive scheme, standard algebras of fuzzy sets, implication.

1 Introduction

Classical deductive schemes in logic, are not trivially translated into fuzzy logic. That is, because instead of working with the properties of a Boolean Algebra, in which classical logic is developed, the stronger structure where fuzzy sets could be defined is the De Morgan Algebra, $([0,1]^X, min, max, 1-id)$. In the context of fuzzy logic, there are many different ways of representing the basic operation. Therefore, it is important to choose the best suitable representation of each operation, to make sure we are going to represent the meaning of a concrete problem. This is framed in the problem of design ([5]). Along this study, we are going to analyze different forms of translating intersection, union, negation and implication for fuzzy sets in order to assure the following deductive scheme,

$$\frac{\begin{array}{l} p \cdot q \longrightarrow r \\ p \end{array}}{q \to r.}$$

2 Basic Concepts

This study deals with different algebraic structures. In this section some basic notions regarding preordered sets are briefly recalled (see [1]):

Definition 2.1 (Poset) *Given a set L, a binary relation $\leq \subseteq L \times L$ is a* partial order *on L provided reflexivity, antisymmetry and transitivity hold. The pair $(L, \leq)$ is said to be a* partially ordered set *or a* poset.

Definition 2.2 *Let $(L, \leq)$ be a poset. Then:*

- *An element 0 in L is said to be a* first element *of the preset if $0 \leq a$ for all $a \in L$.*
- *An element 1 in L is said to be a* last element *of the preset if $a \leq 1$ for all $a \in L$.*
- *A* lattice *is a poset endowed with both an inf-operation, $\cdot$, and a sup-operation, $+$. If the lattice has first and last element, then it is a* bounded lattice.
- *A unary operation $' : L \to L$ is said to be a* negation *if it verifies the two following conditions:*
  1. *If $a \leq b$, then $b' \leq a'$ for all $a, b \in L$.*
  2. *If the preset has a unique first element 0 and unique last element 1, then $0' = 1$ and $1' = 0$.*
- *A bounded lattice $(L, \cdot, +; 0, 1)$ is an* ortholattice *once it is endowed with a negation $'$ verifying the* non-contradiction *law, $a \cdot a' = 0$ for any $a \in L$, and the* involutive *law, $(a')' = a$ for any $a \in L$. Recall that ortholattices verify other well-known properties, such as the* excluded-middle *law, $a + a' = 1$ for any $a \in L$, or the* De Morgan *laws, $(a \cdot b)' = a' + b'$ and $(a + b)' = a' \cdot b'$ for any $a, b \in L$. An* orthomodular lattice *is an ortholattice verifying the so-called* orthomodular law*: for all $a, b \in L$, if $a \leq b$, $b = a + a' \cdot b$, or, equivalently, $a = b \cdot (a + b')$.*
- *A* Boolean algebra *is a distributive ortholattice, i.e., an ortholattice verifying the distributive laws $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ and $a + (b \cdot c) = (a+b) \cdot (a+c)$ for all $a, b, c \in L$.*
  A De Morgan algebra *is a bounded distributive*

lattice endowed with a negation verifying the involutive law and the De Morgan laws.

In fuzzy logic, usually the translation of intersection is doing by a continuous t-norm T, the translation of union by a continuous t-conorm S, and the negation is translated by an strong negation N (an involutive negation). Along this paper we consider that $([0,1]^X, T, S, N)$ is a standard algebra of fuzzy set. From these operators different ways of representing the classical $\rightarrow$, by means of a function J [4], are obtained

- R-implications: $J_T(a,b) = sup\{z \in [0,1]; T(a,z) \leq b\}$
- S-implications: $J(a,b) = S(N(a),b)$
- Q-operators: $J(a,b) = S(N(a), T(a,b))$
- D-operators: $J(a,b) = S(b, T(N(a), N(b)))$
- ML-operators: $J(a,b) = T(A(a),b)$

Those of them allowing to do *forward reasoning*, that is $T(J(a,b),a) \leq b$, are called T-conditionals. The ones of them that translate the relation $a \rightarrow b = 1 \Leftrightarrow a \leq b$, in Boolean algebra, are $a \rightarrow b = a' + b$, but in orthomodularlattices, are $a \rightarrow b = a' + a{\cdot}b$. In De Morgan algebras there is not any conditional verifying this relation, and in the algebras of fuzzy sets it is only the residuated implications $J_T(a,b) = sup\{z \in [0,1] : T(a,z) \leq b\}$. Notice that if there are two t-norms $T_1 \leq T_2$, then $J_{T_2} \leq J_{T_1}$.

In Boolean algebras S-implications, Q-operators and D-operators are coincidental since

$a' + a \cdot b = (a' + a) \cdot (a' + b) = a' + b$

$b + a' \cdot b' = (b + a') \cdot (b + b') = b + a' = a' + b$

3 Posing the problem

The paper deals with the deductive scheme:

$$\frac{\begin{array}{l} p \cdot q \longrightarrow r \\ p \end{array}}{q \rightarrow r.}$$

that, it is written as a functional equation as, $p{\cdot}(p{\cdot}q \rightarrow r) \leq q \rightarrow r$.

In any structure, using the conjunctive conditional ($a \rightarrow b \equiv a \cdot b$) the scheme is verified, since it is $p \cdot q \rightarrow r \equiv p \cdot q \cdot r$, and it is obtained $p \cdot q \cdot r \cdot p = p{\cdot}q{\cdot}r \leq p{\cdot}r \equiv q \rightarrow r$. So, the case of ML-operators is also valid, since they are conjunctive conditionals for fuzzy sets.

In classical logic, i.e., in the structure of a Boolean Algebra (see [3]), the scheme is easily proved translating the implication by its classical expression $a \rightarrow b \equiv a' + b$, and since it is $p \cdot q \rightarrow r \equiv (p \cdot q)' + r = p' + q' + r \Rightarrow p \cdot (p' + q' + r) = 0 + p \cdot q' + p \cdot r = p \cdot (q' + r) \leq q' + r \equiv q \rightarrow r$. Therefore the scheme holds.

But, the scheme does not always hold, for instance, in **Orthomodular lattices,** particularly, in the orthomodular lattice of a *Chinese Lantern* (Fig.1). We have found a counterexample taking the common representation of the implication $a \rightarrow b \equiv a' + a \cdot b$, well known as *Sasaki's operator*. It's enough to take $p = a$, $q = b$ and $r = 0$. With this elements we obtain that,

- $p \cdot q = a \cdot b = 0 \Rightarrow p \cdot q \rightarrow r \equiv 0' + 0 \cdot r = 1 + 0 = 1$
- $q \rightarrow r \equiv b \rightarrow 1 \equiv b' + b \cdot 0 = b'$

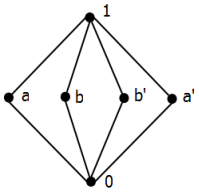


Figura 1: Chinese Lantern

so, the scheme stands that $p \cdot (p \cdot q \rightarrow r) = a \cdot 1 = a \leq b' = q \rightarrow r$ but the elements a and b' are not order-comparable in the *Chinese Lantern*. Therefore, the deductive scheme it is not valid in any orthomodular lattices.

4 Standard Algebras of Fuzzy Sets

In this section, we are going to study what happens for fuzzy sets using different conditional function. The deductive scheme is written in general as

$$T_0(J(T(p,q),r),p) \leq J(q,r) \qquad (1)$$

4.1 R-implications

Dealing with an R-implication $J_T(a,b)$ as a translation of J.

In that case,

Teorema 4.1 *The functional inequality, $T_0(J_T(T_1(p,q),r),p) \leq J_T(q,r)$ where T_0, T_1, T are continuous t-norms, and J_T the residuated implication associated to the t-norm T, holds if and only if $T_0 \leq T \leq T_1$.*

`Proof.` If $T_0 \leq T \leq T_1$, then $T_0(J_T(T_1(p,q),r),p) \leq T(J_T(T_1(p,q),r),p) \leq T(J_T(T(p,q),r),p) = T(J_T(p,J_T(q,r)),p) \leq J_T(q,r)$.

Reciprocally, to proof that $T_0 \leq T$, it is enough to take $q = 1$, in that case, the inequality becomes, $T_0(p, J_T(p,r)) \leq r$, for all $p, r \in [0,1]$, which is equivalent to $J_T(p,q) \leq J_{T_0}(p,q)$. So, $T_0 \leq T$.

To proof that $T \leq T_1$, it will be shown that it is impossible to find any r, such that $T_1(p,q) < r < T(p,q)$. Assuming its existence it will be reached a contradiction, from

- $T_0(p, J_T(T_1(p,q),r)) = T_0(p,1) = p \leq J_T(q,r)$, since $T_1(p,q) < r$ and the verification of the inequality.
- $p > J_T(q,r)$, since $r < T(p,q)$ and the residuation condition.

Therefore, $T \leq T_1$. $\square$

In order to analyze the verification in different cases, the main cases are studied.

4.1.1 Gödel implication

Gödel implication is the residuated implication of the t-norm of min.
In this particular case, the scheme

$$T_0(J_{min}(T(p,q),r),p) \leq J_{min}(q,r) \qquad (2)$$

holds for every T_0 with $T = \min$, but is not verified when it's taken T equal to W and *prod*, since Theorem 4.1, and the chain: $W \leq prod \leq min$.

Taking $J_{\min}(a,b)$ as the conditional, the deductive scheme is verified in the standard algebra $([0,1]^X, min, max, 1-id)$, and also if dealing with different representation for intersections, for instance, dealing with two t-norms, $T_0 = W$ and $T = min$. Similarly, it is easy to find examples where $T = W$, or $T = prod$ that make the inequality false,

- Taking $T = W$ in (4), the scheme is equivalent to $T_0(p, sup\{z \in [0,1] : \min(\max(0, p+q-1), z) \leq r\}) \leq sup\{z \in [0,1] : \min(q,z) \leq r\}$, it's not valid, for example, with $p = 0.7$, $q = 0.3$ and $r = 0.2$ it is obtained, $T_0(0.4,1) = 0.4 \not\leq 0.2$.
- Taking $T = prod$ in (4), the scheme is equivalent to $T_0(p, sup\{z \in [0,1] : \min(p \cdot q, z) \leq r\}) \leq sup\{z \in [0,1] : \min(q,z) \leq r\}$, it's not valid, for example, with $p, q = 0.6$ and $r = 0.4$ it is obtained, $T_0(0.6,1) = 0.6 \not\leq 0.4$.

4.1.2 Menger-Goguen implication

Menger-Goguen implication is the residuated implication of the t-norm of *prod*.
In this case, the scheme

$$T_0(J_{prod}(T(p,q),r),p) \leq J_{prod}(q,r) \qquad (3)$$

holds for any T_0, T in the conditions of Theorem 4.1, that are, $T_0 \leq prod \leq T$. So, T_0 could be product or Łukasiewicz, and T could be the minimum t-norm or product.

Taking $J_{prod}(a,b)$ as the conditional, the deductive scheme is verified in the standard algebra $([0,1]^X, prod, prod^*, 1-id)$, and also if dealing with different representation for intersections, for instance, dealing with two t-norms, $T_0 = W$ and $T = min$. Similarly, it is easy to find examples where $T = W$, or $T = prod$ make the inequality false,

- Taking $T = W$ in (3), the scheme is equivalent to $T_0(p, sup\{z \in [0,1] : \max(0, p+q-1) \cdot z \leq r)\}) \leq sup\{z \in [0,1] : q \cdot z \leq r\}$, with, for example, $r = 0.1 \leq p = 0.4 \leq q = 0.3$, it is obtained $T_0(0.4,1) = 0.4 \not\leq \frac{0.1}{0.3}$.
- Taking $T = prod$ in (3), the scheme is equivalent to $T_0(p, sup\{z \in [0,1] : min(p,q) \cdot z \leq r)\}) \leq sup\{z \in [0,1] : q \cdot z \leq r\}$. In this case a counterexample is found with $T_0 = \min$, for example taking $r = 0.5 \leq 0.75 = 0.75 \leq q = 1$, then $T_0(J_{prod}(min(p,q),r),p) = T_0(\frac{3}{4}, \frac{2}{3}) = \frac{2}{3} \not\leq 0.5 = J_{prod}(q,r)$.

Therefore, the scheme is valid with $T_0 = prod$.

The scheme is also valid with $T_0 = W$, since, $W \leq prod$. Regarding the t-norm T, the scheme is valid with $T = prod$ or $T = min$.

4.1.3 The Residuated implication of Łukasiewicz

In this particular case, the scheme

$$T_0(J_W(T(p,q),r),p) \leq J_W(q,r) \qquad (4)$$

holds for every T with $T_0 = W$, but is not verified when it's taken T_0 equal to *prod* or *min*, since the Theorem 4.1, and the chain: $W \leq prod \leq min$.

Taking $J_W(a,b)$ as the conditional, the deductive scheme is verified in the standard algebra $([0,1]^X, W, W^*, 1-id)$, and also if dealing with different representation for intersections, for instance, dealing with two t-norms, $T_0 = W$ and $T = min$. Similarly, it is easy to find examples where $T_0 = \min$, or $T_0 = prod$ that make the inequality false,

- Taking $T_0 = \min$ in (4), the scheme is equivalent to $\min(p, sup\{z \in [0,1] : W(p,q) \cdot z \leq r)\}) \leq sup\{z \in [0,1] : W(q,z) \leq r\}$. In this case a counterexample is found with $r = 0.5$, $p = 0.7$, $q = 0.9$, then $\min(J_W(min(p,q),r),p) = \min(1, 0.8, 0.7) = 0.7 \not\leq 0.6 = \min(1, 0.6) = J_W(q,r)$.

4.2 S-implications

Using an S-implication $J(a,b) = S(N(a),b)$ as a translation of J, the scheme is translated as,

$$\frac{S(N(T(p,q)),r) \quad p}{S(N(q),r)}$$

or written shortly

$$T_0(S(N(T(p,q)),r),p) \leq S(N(q),r) \qquad (5)$$

In this case a necessary condition is found,

Teorema 4.2 *For any continuous t-norms T_0 and T, any continuous t-conorm S, and any strong negation N, if (5) is verified, implies that $T_0 = W_\varphi$, and $N \leq N_\varphi$.*

Proof. Let's take $q = 1$ and $r = 0$, then $T_0(S(N(T(p,q)),r),p) = T_0(S(N(p),0),p) = T_0(N(p),p)$ and $S(N(q),r) = S(0,0) = 0$. So the inequality it's true only if $T_0(N(p),p) \leq 0 \Rightarrow T_0(N(p),p) = 0$ what it is verified only when $T_0 = W_\varphi$. Then, the inequality became $\varphi^{-1}(max(0, \varphi(N(p)) + \varphi(p) - 1)) \leq 0$, that is verified when $\varphi(N(p)) + \varphi(p) - 1) \leq 0$,therefore it is $N(p) \leq \varphi^{-1}(1 - \varphi(p)) = N_\varphi(p)$. So, $N \leq N_\varphi$. □

To continue with the study of this scheme, we are going to particularize the analysis of S-implications with different pairs of (S,T), with S a t-conorm and T a t-norm, using the negation $N = 1 - id \Rightarrow \varphi = id$ and $T_0 = W_\varphi = W$.

4.2.1 Kleene-Dienes implication

With the implication $J(a,b) = S(N(a),b) = max(1 - a, b)$ the deductive scheme satisfies that

Teorema 4.3 *If $T_0 = W$, taking $S = \max$ and $T = \min$ the inequality (5) holds.*

Proof. Let's check in which cases $W(p, max(1 - min(p,q), r)) \leq max(1 - q, r)$ is satisfied.
$W(p, max(1 - min(p,q), r) \equiv \max(0, p + \max(1 - \min(p,q), r) - 1) = \max(0, \max(p - \min(p,q), p + r - 1))$

- if $p \leq q$: $\max(0, \max(0, p + r - 1)) = \max(0, p + r - 1) \leq \max(0, q + r - 1)$
- if $p > q$: $\max(0, \max(p - q, p + r - 1))$

If $p \leq q$ and $p + r - 1 \geq 0 \Rightarrow q + r - 1 \geq 0$, then $q + r - 1 = \max(0, q + r - 1) \leq \max(1 - q, r) = r$ is always true. Otherwise it is always verified.

If $p > q$, $\max(p - q, p + r - 1) = p - q$ if $p - q \geq p + r - 1$ (i.) and $p + r - 1$ if $p + r - 1 \geq p - q$ (ii.). So, if
$\begin{cases} (i.) \text{ and } 0 \leq p - q \text{ then } p - q \leq 1 - q = \max(1 - q, r) \\ (ii.) \text{ and } 1 - p \leq r \text{ then } p + r - 1 \leq r = \max(1 - q, r) \end{cases}$
in this cases for any p, q and r the scheme is valid. Otherwise the inequality is always verified. □

4.2.2 Reichenbach implication

We are going to study the implication $J(a,b) = S(N(a),b) = prod^*(1 - a, b)$

Teorema 4.4 *If $T_0 = W$, taking $S = prod^*$ and $T = prod$, the inequality (5) holds.*

Proof. $W(p, 1 - p \cdot q + r - (1 - p \cdot q) \cdot r) = \max(0, 1 - p \cdot q + p \cdot q \cdot r + p - 1) = \max(0, p \cdot (1 - q \cdot (1 - r))) = p \cdot (1 - q \cdot (1 - r))$. With $p \cdot (1 - q \cdot (1 - r)) \geq 0$ since $p, q, r \in [0,1]$. The second part of the inequality $prod^*(1 - q, r) = 1 - q \cdot (1 - r)$. So the inequality stands, that $p \cdot (1 - q \cdot (1 - r)) \leq 1 - q \cdot (1 - r)$ what is true. □

4.2.3 Using Łukasiewicz pair (W, W^*)

In this case, the implication is rewritten as $J(a,b) = S(N(a),b) = W^*(1 - a, b)$, and it is coincidental with J_W, studied in section 4.1.3.

4.3 Q-operators

$J(a,b) = S(N(a), T(a,b))$, then the deductive scheme is translated as,

$$\frac{S(N(T(p,q)), T(T(p,q),r)) \quad p}{S(N(q), T(q,r))}$$

or written shortly

$$T_0(S(N(T(p,q)), T(T(p,q),r)), p) \leq S(N(q), T(q,r)) \quad (6)$$

Teorema 4.5 *For any continuous t-norms T_0 and T, any continuous t-conorm S, and any strong negation N, if the deductive scheme (6) is verified, that implies that $T_0 = W_\varphi$, and $N \leq N_\varphi$.*

Proof. Taking $q = 1, r = 0$, $T_0(S(N(T(p,1)), T(T(p,1),0)), p) \leq S(N(1), T(1,0))$. With $T_0(S(N(T(p,1)), T(T(p,1),0)), p) = T_0(S(N(p),0),p) = T_0(N(p),p)$ and $S(N(1), T(1,0)) = S(0,0) = 0$. Then, the inequality $T_0(N(p),p) \leq 0 \Leftrightarrow T_0(N(p),p) = 0 \Rightarrow T = W_\varphi$. And $N \leq N_\varphi$ is identically to the proof of Theorem 4.2. □

To study the scheme using R-implications, we are going to analyze different pairs of (S,T), with S a t-conorm and T a t-norm, taking $N = 1 - id \Rightarrow T_0 = W_\varphi = W$.

4.3.1 Kleene-Dienes implication

With $J(a,b) = \max(1 - a, \min(a,b))$, i.e. $S = \max$, $T = \min$, the scheme (6) is written shortly as

$$W(\max(1 - \min(p,q), \min(\min(p,q), r)), p) \leq \max(1 - q, \min(q,r)) \qquad (7)$$

Teorema 4.6 *If $T_0 = W$, taking $S = \max, T = \min$, the deductive scheme (7) holds.*

Proof. The inequality is equivalent to $\max(0, \max(1 - \min(p,q), \min(\min(p,q), r)) + p - 1) \leq \max(1 - q,$

$\min(q,r))$,
• $\max(1-\min(p,q),\min(\min(p,q),r))+p-1 = \max($
$p-\min(p,q),\min(p,q,r)+p-1) =$
$\diamond\ p-\min(p,q) = \begin{cases} 0 & \text{if } p\le q \\ p-q & \text{if } p\ge q \end{cases}$
$\diamond\ \min(p,q,r)-p+1 = \begin{cases} 2p-1 & \text{if } p\le q,r \\ p+q-1 & \text{if } q\le p,r \\ r+q-1 & \text{if } r\le p,q \end{cases}$

$$= \begin{cases} \max(0,2p-1) & \text{if } p\le q,r \quad (a.) \\ \max(0,r+p-1) & \text{if } r\le p\le q \quad (b.) \\ \max(p-q,p+q-1) & \text{if } q\le p,r \quad (c.) \\ \max(p-q,r+q-1) & \text{if } r\le q\le p \quad (d.) \end{cases}$$

$$\text{(a.)- } = \begin{cases} 2p-1 & \text{if } \frac{1}{2}\le p \quad (a.1) \\ 0 & \text{if } p\le\frac{1}{2} \quad (a.2) \end{cases}$$
$$\text{(b.)- } = \begin{cases} r+p-1 & \text{if } r+p\ge 1 \quad (b.1) \\ 0 & \text{if } r+p\le 1 \quad (b.2) \end{cases}$$
$$\text{(c.)- } = \begin{cases} p-q & \text{if } q\le\frac{1}{2} \quad (c.1) \\ p+q-1 & \text{if } q\ge\frac{1}{2} \quad (c.2) \end{cases}$$
$$\text{(d.)- } = \begin{cases} p-q & \text{if } r+p\le 1 \quad (d.1) \\ r+q-1 & \text{if } r+p\ge 1 \quad (d.2) \end{cases}$$

For this eight cases $\max(0,\max(1-\min(p,q),\min(\min($ $p,q),r))+p-1)$ is equal to the value of $\max(1-\min($ $p,q),\min(\min(p,q),r))+p-1)$, otherwise it's zero and the scheme is trivially valid.

Let's study the second part in each case,
• $\max(1-q,\min(q,r)) =$
(a.1)-
$$= \begin{cases} \max(1-q,q)=q & 1-q\le\frac{1}{2}\le q \\ \max(1-q,r)=r & q\ge\frac{1}{2}\Rightarrow 1-q\le\frac{1}{2}\le r \end{cases}$$
(a.2)- $= \max(1-q,\min(q,r))$
(b.1)and(b.2)- $= \max(1-q,r)$
(c.1)- $= \max(1-q,q) = 1-q$ since in this case, $q\le\frac{1}{2}$
(c.2)- $= \max(1-q,q) = q$ since in this case, $q\ge\frac{1}{2}$
(d.1) and (d.2)- $= \max(1-q,r)$
Combining correctly the former cases it's easy to verify that the inequality holds in each one, so the deductive scheme is valid with $J(a,b) = \max(1-a,\min(a,b))$. □

4.3.2 Reichenbach implication

With $J(a,b) = prod^*(1-a,a\cdot b) = 1-a+a\cdot b-(1-a)\cdot a\cdot b$,i.e. $S = prod^*$, $T = prod$, the deductive scheme (6) is written shortly as

$$W(prod^*(1-prod(p,q),r),p) \le prod^*(1-q,r) \quad (8)$$

Teorema 4.7 *If $T_0 = W$, taking $S = prod^*, T = prod$, the deductive scheme (8) holds.*

`Proof.` The scheme is equivalent to $\max(0,1-p\cdot q+p\cdot q\cdot r-(1-p\cdot q)\cdot p\cdot q\cdot r+p-1) = \max(0,p-p\cdot q+p\cdot q\cdot r-(1-p\cdot q)\cdot p\cdot q\cdot r) = p(1-q\cdot(1-p\cdot q\cdot r))$ since $p(1-p\cdot q(1-p\cdot q\cdot r))\ge 0$ always. Then, the inequality stands that, $p(1-q\cdot(1-p\cdot q\cdot r))\le 1-q(1-q\cdot r)$, what it's true because $p,q,r\in[0,1]$ means that $p\cdot q\cdot r\le p\cdot r \Rightarrow q\cdot(1-p\cdot q\cdot r)\le q\cdot(1-q\cdot r)\Rightarrow p(1-q(1-p\cdot q\cdot r))\le (1-q(1-p\cdot q\cdot r))\le 1-q\cdot(1-q\cdot r)$. □

4.3.3 Using Łukasiewicz pair (W,W^*)

With $J(a,b) = W^*(1-a,W(a,b))$ the deductive scheme (6) is written shortly as

$$W(W^*(1-W(p,q),W(W(p,q),r)),p) \le W^*(1-q,W(q,r)) \quad (9)$$

Teorema 4.8 *If $T_0 = W$, taking $S = W^*, T = W$, the deductive scheme (9) holds.*

`Proof.` The scheme is equivalent to $\max(0,\min(1,$ $1-\max(0,p+q-1)+\max(0,\max(0,p+q-1)+r-1))$ $+p-1)\le\min(1,1-q+\max(0,q+r-1))$.
Let's calculate each component of the inequality,
$\max(0,\max(0,p+q-1)+r-1) =$
$$\begin{cases} \max(0,r-1)=0 & \text{if } p+q-1\le 0 \\ \max(0,p+q-1+r-1) & \text{if } p+q-1\ge 0 \end{cases}$$
• $\max(0,p+q-1+r-1) =$
$$\begin{cases} 0 & \text{if } p+q-1+r-1\le 0 \\ p+q-1+r-1 & \text{if } p+q-1+r-1\ge 0 \end{cases}$$
$1-\max(0,p+q-1)+\max(0,\max(0,p+q-1)+r-1) =$
$$\begin{cases} 1 & \text{if } p+q-1\le 0 \\ 2-(p+q) & \text{if } 0\le p+q-1\le 1-r \\ r & \text{if } 1-r\le p+q-1 \end{cases}$$
$\min(1,1-\max(0,p+q-1)+\max(0,\max(0,p+q-1)+r-1)) =$
$$\begin{cases} 1 & \text{if } p+q-1\le 0 \\ 2-(p+q) & \text{if } 0\le p+q-1\le 1-r \\ r & \text{if } 1-r\le p+q-1 \end{cases}$$
$\max(0,\min(1,1-\max(0,p+q-1)+\max(0,\max(0,p+$ $q-1)+r-1))+p-1) =$
$$\begin{cases} p & \text{if } p+q-1\le 0 \\ 1-q & \text{if } 0\le p+q-1\le 1-r \\ r+p-1 & \text{if } 1-r\le p+q-1 \end{cases}$$
So, finally the first part of the inequality is equals to
$$\begin{cases} p & \text{if } p+q-1\le 0 \\ 1-q & \text{if } 0\le p+q-1\le 1-r \\ r+p-1 & \text{if } 1-r\le p+q-1 \end{cases}$$
We have that the second part of the inequality,
$$\max(0,q+r-1) = \begin{cases} 0 & \text{if } q+r-1\le 0 \\ q+r-1 & \text{if } q+r-1\ge 0 \end{cases}$$
$1-q+\max(0,q+r-1) =$
$$\begin{cases} 1-q & \text{if } q+r-1\le 0 \\ 1-q+q+r-1=r & \text{if } q+r-1\ge 0 \end{cases}$$
Then, the second part is $\min(1,1-q+\max(0,q+$ $r-1)) = \begin{cases} \min(1,1-q)=1-q & \text{if } q+r-1\le 0 \\ \min(1,r)=r & \text{if } q+r-1\ge 0 \end{cases}$
Now, comparing the former cases, it's easy to prove that the inequality is valid. □

4.4 D-operators

$J(a,b) = S(b,T(N(a),N(b)))$, then the scheme is translated as,

$$\frac{\begin{array}{l} S(T(p,q),T(N(T(p,q)),N(r))) \\ p \end{array}}{S(q,T(N(q),N(r)))}$$

or written shortly,

$$T_0(S(r,T(N(T(p,q)),N(r))),p) \leq$$

$$S(r,T(N(q),N(r))) \qquad (10)$$

Teorema 4.9 *For any continuous t-norms T_0 and T, any continuous t-conorm S, and any strong negation N, if the deductive scheme (10) is verified, implies that $T_0 = W_\varphi$, and $N \leq N_\varphi$.*

`Proof.` Taking $q = 1, r = 0$, $T_0(S(0,T(N(T(p,1)),N(0))),p) \leq S(0,T(N(1),N(0)))$.
With $T_0(S(0,T(N(p),1)),p) = T_0(S(0,N(p)),p) = T_0(N(p),p)$ and $S(0,T(0,1)) = S(0,0) = 0$ Then, the inequality $T_0(N(p),p) \leq 0 \Leftrightarrow T_0(N(p),p) = 0 \Rightarrow T = W_\varphi$. And $N \leq N_\varphi$ is proved as in the proof of Theorem 4.2. □

When we take the implication as a D-operator doing the same analysis it is obtained analogous results to the Q-operators.

5 Conclusions

It is studied the deductive scheme,

$$\frac{\begin{array}{l} p \cdot q \to r \\ p \end{array}}{q \to r.}$$

The scheme is always valid in Boolean algebras, but not in orthomodular lattices or De Morgan algebras.

Dealing with fuzzy sets the scheme is written as $T_0(J(T(p,q),r),p) \leq J(q,r)$, and the following solutions are found,

R-implications		
	T	T_0
J_{min}	min	all
J_{prod}	min, *prod*	W, *prod*
J_W	all	W

S-implications		
$T_0 = W_\varphi, N \leq N_\varphi$		
S	T	T_0
max	min	W
$prod^*$	*prod*	W

Q-operators and D-operators		
$T_0 = W_\varphi, N \leq N_\varphi$		
S	T	T_0
max	min	
W^*	W	W
$prod^*$	*prod*	

Given a system of n fuzzy rules ([2]),
If 'x_1 is P_{1_i}' and 'x_2 is P_{2_i}', then 'y is Q_i' , with $1 \leq i \leq n$,
suppose that there are a continuous t-norm T, and a fuzzy T-conditional J translating, respectively, the linguistic connectives and, and If/then , verifying the inequality,

$T(a,J(T(a,b),c)) \leq J(b,c)$, for all a,b,c in [0,1].

In a case like this, and provided it is known that the statements 'x_1 is P_{1_i}' do hold for all i between 1 and n, the above system of fuzzy rules can be reduced to the simpler system,

If 'x_2 is P_{2_i}', then y is Q_i , with $1 \leq i \leq n$.

Acknowledgements

The authors are indebted with the anonymous reviewers for the analysis done to the first version of this paper. This paper is partially supported by the Foundation for the Advancement of Soft Computing (Asturias, Spain), and CICYT (Spain) under grant TIN2008-06890-C02-01.

Referencias

[1] G. Birkhoff. *Lattice Theory.* Amer. Math. Society, Colloq. Publs, 3th. edition, (7th. print.), 1993.

[2] G. J. Klir, B. Yuan, Fuzzy sets and fuzzy logic, Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[3] I. García-Honrado E. Trillas. ¿Hacia un replanteamiento del cálculo proposicional clásico? *accepted in Ágora*, 2011.

[4] E. Trillas and I. García-Honrado. A reflection on the design of fuzzy conditionals. Technical Report FSC-2010-12, European Centre for Soft Computing, 2010. Submitted to Experimentation and Theory/ A Homage to A. Mamdani.

[5] E. Trillas and S. Guadarrama. Fuzzy representations need a careful design. *International Journal of General Systems*, 39(3):329–346, 2010.

ALGUNOS MODELOS COMPUTACIONALES DE RAZONAMIENTO CONJETURAL

Itziar García-Honrado[1], **David P. Pancho**[2], **Laura Menéndez-Arduengo**[2]

[1] *Departamento de Estadística, I. O. y Didáctica de las Matemáticas, Universidad de Oviedo, España.*
[2]*European Centre for Soft Computing, Mieres, 33600, España*
garciaitziar@uniovi.es, {david.perez, laura.menendez}@softcomputing.es

Resumen

Esta contribución es un primer intento de automatizar la computación de modelos de Conjeturas, Hipótesis y Consecuencias (CHC) introducidos por Enric Trillas. Estos modelos formalizan el razonamiento de sentido común u ordinario, consiguiendo reorganizar el Universo de Discurso una vez conocido un conjunto de premisas. Así se obtiene información no contradictoria de las premisas, el conjunto de las conjeturas, que a su vez se divide en el de Consecuencias (formalizadas a partir de un operador de consecuencias en sentido de Tarski), Hipótesis y Especulaciones.

La propuesta aquí realizada computa los modelos CHC en Universos de Discurso cualesquiera en los que puede existir una relación de preorden entre sus elementos y/o en aquellos que pueden representarse mediante un ortoretículo.

Palabras Clave: Conjeturas, Operadores de Consecuencia, Preordenes, Orto-retículos, Razonamiento Ordinario

1. INTRODUCCIÓN

En Computación se está tratando de modelizar formalmente el razonamiento ordinario, el cuál no sigue las pautas del razonamiento matemático-deductivo. Por ello, no es posible crear modelos con operadores monótonos, ni tampoco es posible utilizar los modelos de razonamiento no monótonos existentes en Inteligencia Artificial.

En el razonamiento ordinario es posible distinguir cuatro tipos de razonamiento, deductivo, abductivo, especulativo y por analogía. Nos centraremos en los tres primeros tipos.

La creación del modelo se hará a partir de una base de conocimiento dada. La relación entre los elementos de esa base de conocimiento, a priori no ha de ser ni reflexiva, ni transitiva, no obstante, si se cumplen estas propiedades, la relación a posteriori se mantiene; en otro caso, se crea una nueva relación manteniendo la esencia de la original e imponiendo dichas propiedades.

Después, para crear el modelo, se ha de fijar un operador de consecuencias en el sentido de Tarski. Así, se podrá modelar la abducción y la especulación, basándose en la deducción y englobándolo todo en el concepto de conjetura.

El modelo utilizado se denomina modelo CHC [3] y para su utilización se han de tener:

1. Una base de conocimiento, L, con la información disponible de cada problema en concreto. La relación entre los elementos $R \subseteq L \times L$ se representa mediante una matriz $A \subseteq L \times L$, que transformaremos en otro matriz A^* transitiva y reflexiva, representante de un preorden, con la cuál construimos un ortoretículo $l(A^*)$.

2. Un conjunto de premisas, aplicamos el modelo CHC sobre la base de conocimiento creada.

La estructura más débil en la que se puede definir el modelo es un preorden, ya que un operador de consecuencias C define una relación R como $aRb \Leftrightarrow b \in C(\{a\})$. Para razonar siempre es útil añadir la negación (*lo que NO es*) a los elementos de la base de conocimiento, así podemos definir el modelo en un ortoretículo.

En resumidas cuentas, lo que trataremos es de, 1) mostrar un proceso para transformar cualquier relación en un preorden y éste en un ortoretículo; y 2) construir un programa que nos permita obtener directamente todas las conjeturas, consecuencias, hipótesis y especulaciones con dos modelos CHC diferentes.

2. CONCEPTOS BÁSICOS

2.1. Modelado de la base de conocimiento

El modelo matemático utilizado para construir la base de conocimiento es un preorden o un ortoretículo [1].

Definamos algunos conceptos a los que nos referiremos con posterioridad.

- Ínfimo: operación $inf : \mathscr{P}(L) \rightarrow L$ tal que $inf(S)$ es un ínfimo de S.
- Negación: operación unaria $' : L \rightarrow L$ tal que verifica las siguientes condiciones,
  1. Si $a \leq b$, entonces $b' \leq a' \; \forall a,b \in L$
  2. Si en el conjunto preordenado (preset) los elementos primero y último son únicos, entonces $0' = 1'$ y $1' = 0'$.

Dado que toda relación se puede representar a través de un grafo, para verificar si se trata de un preorden (relación reflexiva y transitiva) es suficiente comprobar si la matriz de adyacencia, A, tiene unos en la diagonal principal y si el producto min-max de A por A coincide con A.

2.2. Conjeturar a partir de consecuencias

La estructura más débil en la que se ha definido un modelo CHC es un preorden dotado de una negación [8]. Si esto no es suficiente, se le añade un ínfimo y un supremo, con lo cual pasaríamos a tener un poset (conjunto parcialmente ordenadao). El ínfimo del conjunto L se llama primer elemento, 0, y el supremo se llama último elemento, 1. Aquí únicamente trataremos con estructuras algebraicas finitas, un conjunto L finito.

Dado L conjunto y $\mathfrak{F} \subset \mathscr{P}(L)$, diremos que $(L,\mathfrak{F},C)$ es una estructura de consecuencias, con $C : \mathfrak{F} \rightarrow \mathfrak{F}$, si se verifica que 1) $P \subset C(P)$, para todo $P \in \mathfrak{F}$ (C es *extensivo*) 2) si $P \subset Q$, entonces $C(P) \subset C(Q)$, para todo $P,Q \in \mathfrak{F}$ (C es *monótono*) y 3) $C(C(P)) = C(P)$, ó $C^2 = C$, para todo $P \in \mathfrak{F}$ (C es una *clausura*).

Luego, C es un operador de consecuencias [2] (en el sentido de Tarski) para $\mathfrak{F}$ en L.

En este se utilizan dos operadores de consecuencias diferentes,

- $C_\wedge(P) = \{q \in L; p_\wedge \leq q\}$, con $inf(P) := p_\wedge$.
- $C_\leq(P) = \{q \in L; \exists p \in P : p \leq q\}$.

En [8] se muestran los espacios $\mathfrak{F}$ donde se puede definir una estructura de consecuencias. Los elementos del ortoretículo construido pertenecen a todos esos espacios, luego únicamente hemos de tener en cuenta que el modelo no tiene sentido si $p_\wedge = 0$ en el operador de consecuencias $C_\wedge$.

2.2.1. Elementos de la base de conocimiento, L.

Consecuencias, $C(P)$, son aquellos elementos que se deducen del conjunto de las premisas.
Conjeturas, son aquellos elementos que no contradicen la información o, en su defecto, el elemento resumen de la información.

Dado un operador de consecuencias existente, C, (véase en [4]) podemos definir las conjeturas de P como aquellos elementos cuya negación no se encuentra en $C(P)$, esto es, $Conj_C(P) = \{q \in L; q' \notin C(P)\}$.
Hipótesis, se definen como

$$Hyp_C(P) = \{h \in Conj_C(P) - C(P); \{h\} \in \mathfrak{F},\ P \subset C(h)\}.$$

Especulaciones, son aquellos elementos que no son consecuencias ni hipótesis, por lo tanto, $Sp_C(P) = Conj_C(P) - [C(P) \cup Hyp_C(P)] = \{q \in L, q \notin C(P), q' \notin C(P), q \notin Hyp_C(P)\}$.

Refutaciones, son aquellos elementos que contradicen la información proporcionada por las premisas, i.e., son aquellos elementos que no se pueden conjeturar, $R(P) = L - Conj_C(P)$.

Así, obtenemos una partición del conjunto L,

$$Conj_C(P) = C(P) \cup Hyp_C(P) \cup Sp_C(P).$$

Entre las definiciones anteriores, el conjunto de las especulaciones es el más desconocido, aún siendo el de mayor interés, por su carácter no-monótono. De hecho, en retículos ortomodulares, las especulaciones tienen gran relevancia, ya que nos permiten hallar consecuencias o hipótesis, a partir de una caracterización expuesta en [9],

3. MODELOS COMPUTACIONALES

Como se ha mostrado en el apartado anterior, la estructura más débil en la cual se pueden definir los modelos de conjeturas es un preorden [8]. Por lo tanto, en este artículo se propone aplicar el modelo a cualquier grafo sin ciclos de longitud mayor que uno, forzándolo a ser un preorden, ésto es, transitivo y reflexivo, siempre que sea semánticamente coherente. En caso de obtener un grafo totalmente conexo, se descarta el ejemplo para aplicar el modelo CHC, ya que todos los elementos se obtendrían como consecuencia de cualquier otro elemento y no resultarían útiles los conjuntos obtenidos.

Representaremos el grafo por su matriz de adyacencia y construiremos un grafo reflexivo y transitivo a partir del

inicial. Para ello se hará que la matriz de adyacencia del nuevo grafo tenga unos la diagonal principal y se calculará la clausura transitiva de ésta, que es la matriz más pequeña que es transitiva y cuyo grafo asociado contiene al inicial.

Evidentemente, si el grafo inicial es reflexivo y transitivo, el proceso no lo modificará. Si no lo es, antes de hacer esta construcción debemos reparar en el significado de la relación que se nos da en un principio y estudiar si admite el ser considerada reflexiva y transitiva. Por ejemplo, si el grafo representa la relación "*existencia de un vuelo directo*", no podríamos considerarla transitiva, ni reflexiva, si la forzamos a ser transitiva y reflexiva, la semántica de de la relación se debilitaría, pudiendo redefinirse la relación como "*existencia de una conexión aérea, admitiendo trasbordos*". De este modo, el modelo no representaría fielmente la semántica de la relación inicial.

Todas estas restricciones nos hacen que tengamos las siguientes hipótesis de trabajo:

- La relación original R no debe producir un grafo conexo[1]. Estas relaciones producen una situación no informativa.
- El grafo producido por la relación no ha de tener ciclos de longitud mayor a 1 en su estructura.
- Las relaciones han de admitir de forma natural la reflexividad y transitividad. La semántica de la nueva relación no es exactamente la de la original, al relajarse de alguna forma la semántica de ésta.

Teniendo un preorden, nos encontramos en condiciones de poder construir el conjunto de consecuencias bajo el operador $C_{\leq}$.

3.1. CONJETURAS EN PREÓRDENES

La mínima condición para poder manejar el operador $C_{\leq}$ es calcular las consecuencias sobre un preorden. En el caso de querer calcular el conjunto de conjeturas, además de trabajar con un preorden se necesita manejar una negación, $'$, de los elementos, se dirá que un elemento q es conjeturable a partir de un conjunto de premisas P, siempre que $q' \notin C_{\leq}(P)$.

Así pues si en el grafo inicial aparecen los elementos y sus negados, se podrían calcular las conjeturas y clasificarlas en consecuencias, hipótesis y especulaciones.

Observar que si hay relaciones entre un elemento y su negado el modelo se puede aplicar pero no se obtendría una estructura de consecuencias consistente. Así podrían obtenerse como consecuencias un elemento y su negado, lo que resulta contradictorio. Por lo tanto, es deseable que en el Universo de Discurso no existan relaciones entre un elemento y su negado.

3.1.1. Operador de consecuencias $C_{\leq}$ en preórdenes

Las *conjeturas* son la negación de los elementos que tienen un 0 en la fila de cada premisa, P. Aquellos elementos que verifiquen esta condición para todas las premisas son conjeturas.

$$Conj_{C_{\leq}}(P) = \bigcap_{i=1}^{n} Conj_{C_{\leq}}\{p_i\}$$

Para este caso se obtienen las *consecuencias* individuales de cada premisa, para posteriormente obtener la unión de estas como las consecuencias totales del grupo de premisas dadas.

$$C_{\leq}(P) = \bigcup_{i=1}^{n} C\{p_i\}$$

Como se vió en la sección 2, aquellas conjeturas que no son consecuencias han de ser *especulaciones* o *hipótesis*. Así mismo, las hipótesis han de tener como consecuencias todas las premisas, con lo que dividimos a este grupo siguiendo este criterio.

3.2. CONJETURAS EN ORTORRETÍCULOS

Se propone restringir los preórdenes utilizados en la sección 3.1 para que no contengan relaciones entre elementos positivos y negativos ($a \not R b' \ \forall a, b \in L$). Con esta limitación, y las indicadas en la sección 3, podremos construir un ortoretículo asociado con el que poder utilizar los operadores de consecuencias $C_{\leq}$ y $C_{\wedge}$ [8, 6].

A continuación se desglosa el proceso de construcción del citado ortoretículo, añadiendo ciertos elementos artificialmente que no modifican la semántica de la relación original.

- Añadir elementos negados y sus correspondientes relaciones: La principal propiedad de la negación es que aRb implica $b'Ra'$ [7]. Con ello, añadimos los elementos negativos y las relaciones derivadas de los originales.
- Relacionar todos los elementos: Se añaden dos elementos *primero* (0) y *último* (1) cumpliendo las siguientes condiciones $0Ra \ \forall a \in L$ y $aR1 \ \forall a \in L$.

3.2.1. Operador de consecuencias $C_{\leq}$ en ortoretículos

En este caso, al igual que se realizó en la sección 3.1.1, se utiliza el operador de consecuencias $C_{\leq}$, para obtener las conjeturas en ortoretículos.

[1] Un grafo es conexo si para cualquier par de vértices a y b existe al menos una trayectoria (una sucesión de vértices adyacentes que no repita vértices) que los una.

3.2.2. Operador de consecuencias $C_\wedge$ en ortoretículos

Para utilizar el operador de consecuencias $C_\wedge$ ha de calcularse el ínfimo de las premisas. Puede darse el caso de que este ínfimo sea el elemento primero, con lo que tendríamos como consecuencia todos los elementos del ortoretículo. En este caso no se calcularán las conjeturas, ya que en la relación original no habría ínfimo de las premisas.

Una vez se ha calculado el ínfimo, procedemos a calcular las *conjeturas*, la negación de los elementos que tienen un 0 en la fila del ínfimo de las premisas, $inf(P)$, en la matriz de adyacencia. Las *consecuencias* pueden calcularse buscando los elementos que tienen un 1 en la fila del $inf(P)$. Las *hipótesis* y *especulaciones* se obtienen utilizando aquellas conjeturas que no son consecuencia. Con ellas, se comprueba cuando de estos elementos se llega al $inf(P)$. En caso de que se llegue será una hipótesis, en caso contrario tendremos una especulación.

3.3. PROCESO DE CONSTRUCCIÓN

El proceso de construcción de preórdenes y ortoretículos, sobre los que calcular conjeturas, es el siguiente:

1. Obtener la matriz de adyacencia, A, relacionada a la relación original, R.
2. Obtener una nueva relación, si no fuera reflexiva y transitiva, derivada de A que sea reflexiva y transitiva, obteniendo A^*. La reflexividad se obtiene colocando la diagonal principal de A a unos y la transitividad haciendo el cierre transitivo.

   Esta nueva relación es un preorden y ya se pueden obtener las conjeturas a través de ella como se mostró en la sección 3.1.
3. Añadir elementos negados y sus correspondientes relaciones: Partiendo de A^* se crea una nueva matriz,

$$\left(\begin{array}{ccc|ccc} & & & 0 & \dots & 0 \\ & A^* & & 0 & \dots & 0 \\ & & & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & & & \\ 0 & \dots & 0 & & (A^*)' & \\ 0 & \dots & 0 & & & \end{array}\right)$$

4. Finalmente, el *primer* y *último* elementos son añadidos a la matriz (primer y última fila/columna respectivamente).

$$l(A^*) = \left(\begin{array}{c|ccc|ccc|c} 1 & 1 & 1 & & \dots & & & 1 \\ \hline 1 & & & & 0 & \dots & 0 & 1 \\ 1 & & A^* & & 0 & \dots & 0 & 1 \\ 1 & & & & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & & & & 1 \\ 1 & 0 & \dots & 0 & & (A^*)' & & 1 \\ 1 & 0 & \dots & 0 & & & & 1 \\ \hline 0 & 0 & 0 & & \dots & & & 1 \end{array}\right)$$

   Esta relación creada es un ortoretículo y sobre ella se podrán obtener conjeturas utilizando los operadores de consecuencias $C_\leq$ y $C_\wedge$.

3.4. DIFERENCIAS Y APLICACIONES DE LOS DOS MODELOS

Partiendo de un grafo sin ciclos, siempre que los nodos que aparezcan sean elementos en positivo y sus negaciones, transformándolo en un preorden se obtendrá un modelo CHC para el operador $C_\leq$.

En múltiples ocasiones se muestran esquemas mentales las relaciones entre elementos positivos. Es por ello que en la sección 3.2 se construye un modelo añadiendo de forma artificial las negaciones de los elementos y las relaciones entre ellos, obteniéndose un orto-retículo, estructura matemática en las que los modelos de consecuencias CHC se han estudiado más en profundidad. Por lo tanto, no sólo podemos utilizar el operador de consecuencias $C_\leq$, sino que también se modelará bajo el $C_\wedge$, el mayor operador de consecuencias en orto-retículos.

Las diferencias entre el modelo en preordenes y ortoretículos construido a partir de un grafo, está en cómo se define la negación.

- Si el grafo inicial no tiene la negación de ninguno de sus nodos, se podría aplicar sólo el modelo de conjeturas en orto-retículos.
- Si el grafo inicial tiene como nodos todas las negación de sus elementos, podríamos aplicar sólo el modelo de conjeturas en preórdenes.
- Si el grafo inicial sólo tiene algún nodo con su respectiva negación, se podría aplicar el modelo de conjeturas en preórdenes, añadiendo las negaciones de los demás elementos.

Estos dos últimos casos requieren un estudio más profundo, ya que dependiendo de los espacios $\mathfrak{F}$ en los que se defina el conjunto de premisas, P, podríamos obtener consecuencias inconsistentes. Esto es, que tanto un elemento como su negado pertenezcan a $C_\leq(P)$, pero si, por ejemplo $\mathfrak{F} = \{P \subset L; \forall p_1, p_2 \in P, p_1 \not\leq p_2'\}$ se puede asegurar la consistencia del operador $C_\leq$ [8].

4. EJEMPLOS

Este ejemplo, obtenido de [5], va a servir para ilustrar el proceso de construcción y explicar la forma de cálculo de conjeturas en cada caso. Se puede observar la relación original en la figura 1, donde las premisas se han destacado en color amarillo.

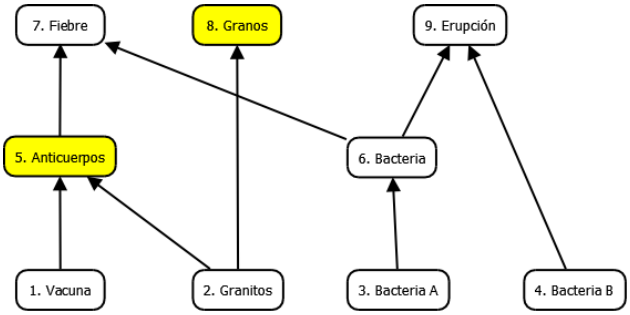


Figura 1: Grafo vinculado a la relación

De esta relación se obtiene la matriz A, cuyas filas y columnas corresponden a los números indicados en el grafo.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La matriz A tiene un uno en la posición (i, j) si existe un arco que una estos dos nodos, y un cero en caso contrario.

A continuación se crea una nueva relación, derivada de la anterior, que sea reflexiva y transitiva, A^*. Para ello se fija la diagonal principal a unos y se hace el cierre transitivo.

$$A^* = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Llegados a este punto, en el que ya tenemos un preorden, se pueden calcular las conjeturas con el operador $C_{\leq}$. Para ello, nos hemos de fijar en los elementos relacionados con las premisas: {Anticuerpo (5), Granos (8)}. Las consecuencias serán los elementos a los que se pueda llegar desde las premisas, esto es, los unos en las filas 5 y 8 de la matriz, en este caso {Anticuerpo (5), Fiebre (7), Granos (8)}.

Para lograr un orto-retículo hemos de continuar con el proceso de construcción. El siguiente paso es añadir los elementos negativos y las relaciones entre ellos. Para ello, realizamos la traspuesta de la matriz A^*, y la colocamos en su posición para relacionar a éstos.

$$\left(\begin{array}{ccccccccc|ccccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

Lo siguiente es añadir el primer y último elementos. El elemento *primero* está relacionado con todos, mientras que todos están relacionados con el *último*.

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1
0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Una vez se añaden estos elementos sabemos que tenemos un orto-retículo por construcción, con lo que podemos obtener las conjeturas con el operador de consecuencias $C_{\leq}$ y $C_{\wedge}$.

Para el operador $C_{\leq}$ calculamos las *conjeturas* como la negación de aquellas posiciones de las filas de premisas que tengan 0. Por ello, {Vacuna (2), Granitos (3), BacteriaA (4), BacteriaB (5), Anticuerpo (6), Bacteria (7), Fiebre (8), Granos (9), Erupción (10), Vacuna' (11), Granitos' (12), BacteriaA' (13), BacteriaB' (14), Bacteria' (16), Erupción' (19), *elemento último* (20)} son las conjeturas. El resto son refutaciones. Las consecuencias se obtienen buscando los elementos a los que se pueden llegar desde las premisas, tienen un 1 en alguna de las filas de las premisas {Anticuerpo (6), Fiebre (8), Granos (9), *elemento último* (20)}. Llegados a este punto, y para calcular *hipótesis* y *especulaciones* se realiza la diferencia de conjuntos, observando aquellas conjeturas que no son consecuencias. Así, {Vacuna (2), Granitos (3), BacteriaA (4), BacteriaB (5), Bacteria (7), Erupción (10), Vacuna' (11), Granitos' (12), BacteriaA' (13), BacteriaB' (14), Bacteria' (16), Erupción' (19)} son los candidatos. Para distinguirlos, se observa si desde

ellos se puede llegar a las premisas, teniendo así una hipótesis, o no, siendo así una especulación. Con esto, tenemos como hipótesis {Granitos (3)} y como especulaciones {Vacuna (2), BacteriaA (4), BacteriaB (5), Bacteria (7), Erupción (10), Vacuna' (11), Granitos' (12), BacteriaA' (13), BacteriaB' (14), Bacteria' (16), Erupción' (19)}.

En el caso del operador $C_\wedge$, primeramente calculamos el ínfimo (en este caso Granitos (2)), y utilizando este elemento como premisa, procedemos de forma análoga al caso anterior.

Con estos tres casos obtenemos los siguientes conjuntos de conjeturas:

	Preorden $C_\leq$	Ortoretículo $C_\leq$	Ortoretículo $C_\wedge$
Consecuencias	Anticuerpo, Fiebre, Vacuna	Anticuerpo, Fiebre, Granos	Granitos, Anticuerpo, Fiebre, Granos
Hipótesis	-	Granitos	$\emptyset$
Especulaciones	-	Vacuna, BacteriaA, BacteriaB, Bacteria, Erupción, Vacuna', Granitos', BacteriaA', BacteriaB', Bacteria', Erupción'	Vacuna, BacteriaA, BacteriaB, Bacteria, Erupción, Vacuna', BacteriaA', BacteriaB', Bacteria', Erupción'

5. CONCLUSIONES Y TRABAJO FUTURO

Se ha tratado de automatizar, por vez primera, el cálculo de conjeturas. Desde un estructura muy débil, un grafo, se ha propuesto la transformación a preorden y ortoretículo donde es posible realizar el cálculo de conjeturas.

La propuesta tiene fuertes restricciones, como la imposibilidad de que el grafo original contenga ciclos, que han de ser salvadas buscando nuevas estructuras donde representar la información. No poder relacionar elementos positivos y negativos es otro gran handicap a la hora de representar situaciones de las bases de conocimiento reales.

Se ha desarrollado un software que permite el cálculo de conjeturas con las distintas estructuras, preorden y ortoretículo, y operadores de consecuencias propuestos, $C_\leq$ y $C_\wedge$. Esta herramienta, aún automatizando el cálculo de conjeturas, tiene muchas limitaciones a la hora de incorporar nuevo conocimiento, ya que ha de rehacerse toda la base de conocimiento en caso de que algo de lo presente cambie.

La formalización propuesta aquí, aunque establecida en términos precisos, abre la puerta al trabajo con lógica fuzzy, al permitir trabajar en preórdenes como es el caso de el conjunto de los conjuntos fuzzy ($[0,1]^X$) con el orden puntual ($\mu \leq \sigma$ si y sólo si $\mu(x) \leq \sigma(x)$ para todo $x \in X$, siendo $\mu, \sigma \in [0,1]^X$.

Agradecimientos

Este artículo está parcialmente financiado por la Fundación para el Progreso del Soft Computing (España) y el MICINN bajo el proyecto TIN2008-06890-C02-01.
Los autores quieren expresar su agradecimento a los dos revisores anónimos.

Referencias

[1] G. Birkhoff. *Lattice Theory*. Amer. Math. Society, Colloq. Publs, 3th. edition, (7th. print.), 1993.

[2] J. L. Castro and E. Trillas. Sobre preordenes y operadores de consecuencias de Tarski. *Theoria*, 4(11):419–425, 1989. (In Spanish).

[3] E. Castiñeira E. Trillas, S. Cubillo. On conjectures in orthocomplemented lattices. *Artificial Intelligence*, 117:255–275, 2000.

[4] M. Monserrat E. Trillas, M. Mas. Conjecturing from consequences. *International Journal of General Systems*, 38:567–578, 2009.

[5] I Garcia-Honrado, E. Trillas, S. Guadarrama, and E. Renedo. Evaluating premises, partial consequences and partial hypotheses. In *Proceedings IFSA-EUSFLAT*, pages 897–902, 2009.

[6] A. Fernández Pineda, E. Trillas, and C. Vaucheret. Additional comments on conjectures, hypotheses and consequences in orthocomplemented lattices. In J. A. Campbell and E. Roanes Lozano, editors, *Artificial Intelligence and Symbolic Computation. Lecture Notes in Artificial Intelligence.*, pages 107–114. Springer. Berlin, 2000.

[7] E. Trillas. Sobre funciones de negación en la teoría de conjuntos difusos. *Stochastica*, 3(1):47–59, 1979. (In Spanish). Reprinted (English version) in *Advances of Fuzzy Logic* (S. Barro et al. (eds.)):31-43. Press of the Universidad de Santiago de Compostela, 1998.

[8] E. Trillas, I. García-Honrado, and A. Pradera. Consequences and conjectures in preordered sets. *Information Sciences*, 180(19):3573–3588, 2010.

[9] E. Trillas, A. Pradera, and A. Álvarez. On the reducibility of Hypotheses and Consequences. *Information Sciences*, 179(23):3957–3963, 2009.

UN MÉTODO PARA CALCULAR OPENINGS TRANSITIVOS

Dionís Boixader[1], **Jordi Recasens**[1]

[1]*Secció Matemàtiques i Informàtica*
ETS Arquitectura del Vallès
Universitat Politècnica de Catalunya
Pere Serra 1-15
08190 Sant Cugat del Vallès
Barcelona
{dionis.boixader,j.recasens}@upc.edu

Resumen

Así como la clausura transitiva de una relación R reflexiva y simétrica es única y se conocen varios algoritmos para obtenerla, puede haber una gran cantidad de openings transitivos (operadores de indistinguibilidad maximales entre los menores o iguales que R). Este trabajo presenta un método para calcular openings transitivos de una relación reflexiva y simétrica R respecto a una t-norma arquimediana continua.

Palabras Clave: Similitud, operador de indistinguibilidad, t-norma finito valorada, clausura transitiva, opening transitivo.

1 INTRODUCCIÓN

La clausura transitiva de una relación reflexiva y simétrica R es un operador de indistinguibilidad mayor o igual que R. En este caso es posible obtener la mejor aproximación superior porque el ínfimo de operadores de indistinguibilidad es también un operador de indistinguibilidad. Si lo que se desea es una aproximación inferior, la situación es más complicada ya que el supremo de operadores de indistinguibilidad no es un operador de indistinguibilidad en general. Lo que se puede hallar son operadores de indistinguibilidad maximales entre los que son menores o iguales que R. Estos operadores se llaman openings transitivos y no son únicos en general, sino que puede haber incluso un número no finito de ellos incluso en universos finitos.

En [5] se propone un algoritmo para hallar openings transitivos de una relación reflexiva y simétrica dada, pero las relaciones obtenidas no son simétricas en general y en el proceso de simetrización se puede perder la maximalidad. Se han propuesto métodos heurísticos para hallar operadores de indistinguibilidad menores o iguales que una relación reflexiva y simétrica dada ([3]), pero que no aseguran su maximalidad.

La t-norma mínimo es una excepción dada el comportamiento especial de los min-operadores de indistinguibilidad (similitudes). En este caso hay varios algoritmos para calcular algunos openings transitivos de una relación reflexiva y simétrica dada. Un método clásico es el complete linkage. Otros algoritmos se pueden encontrar en [4] [6].

En este trabajo se expondrá un método para hallar todos los openings transitivos de una L-relación reflexiva y simétrica en un conjunto finito X. Este resultado se usará para obtener openings transitivos respecto a la t-norma de Łukasiewicz que se pueden generalizar a cualquier t-norma arquimediana continua. El método es efectivo cuando se combina con algoritmos heurísticos como los propuestos en [3] (véase el Ejemplo 3.8).

2 PRELIMINARES

Esta sección contiene definiciones y resultados sobre t-normas finito valoradas que se usarán en el trabajo. Las demostraciones y más información sobre t-normas finito valoradas pueden hallarse en [11]. También se recuerdan las definiciones de operador de indistinguibilidad, proximidad y clausura transitiva.

Los operadores definidos en una cadena finita L son de gran interés, especialmente porque usualmente se razona usando términos lingüísticos o etiquetas que están totalmente ordenadas. Por ejemplo, el tamaño de un objeto se puede granularizar en *muy pequeño, pequeño, normal, grande, muy grande*. Si un operador T está definido en este conjunto, entonces será posible combinar estas etiquetas para obtener, por ejemplo, $T(\textit{normal}, \textit{muy grande})$. Los cálculos se simplifican enormemente si se combinan las etiquetas de este modo ya que no es necesario asignarles valores numéricos o identificarlas con un intervalo o un subconjunto difuso.

También son útiles las cadenas finitas en casos en que los valores son discretos por naturaleza o discretización. En

un estudio de mercado, se puede pasar un test a determinados clientes para que describan su grado de satisfacción sobre un cierto servicio en el que se le pida evaluar usando números enteros entre 0 y 5 o asignando etiquetas variando entre *totalmente insatisfecho/a* y *totalmente satisfecho/a*.

En este contexto, diversos autores han estudiado t-normas y t-conormas definidas en cadenas finitas ([10], [11]) obteniendo resultados teóricos muy interesantes.

Sea L un conjunto finito totalmente ordenado con mínimo e y máximo u.

Definición 2.1. *Una operación binaria $T : L \times L \to L$ es una t-norma si, y sólo si, para todo $x,y,z \in L$*

1. $T(x,y) = T(y,x)$
2. $T(T(x,y),z) = T(x,T(y,z))$
3. $T(x,y) \leq T(x,z)$ cuando $y \leq z$
4. $T(x,u) = x$.

El conjunto de t-normas en una cadena finita depende sólo de su cardinal. Por esta razón consideraremos sólo las cadenas $L = \{0 = \frac{0}{n}, \frac{1}{n}, \frac{2}{n}, ..., \frac{n}{n} = 1\}$.

Ejemplo 2.2.

1. *La t-norma Mínimo T en L está definida por $T(i,j) = \min\{i,j\}$.*
2. *La t-norma de Łukasiewicz T en L está definida por $T(i,j) = \max\{i+j-1,0\}$.*

Definición 2.3.

- *Una aplicación $f : L \to L$ es suave si, y sólo si,*

$$0 \leq f(i+1) - f(i) \leq \frac{1}{n} \text{ para todo } i \in L,\ i < 1.$$

- *Una aplicación $F : L \times L \to L$ es suave si, y sólo si, es suave respecto a ambas variables.*

La suavidad de las t-normas en cadenas finitas es equivalente a la continuidad de las t-normas definidas en $[0,1]$.

Definición 2.4. *Una t-norma T en L es divisible si, y sólo si, para todo $i,j \in L$ con $i \leq j$ existe $k \in L$ tal que*

$$i = T(j,k).$$

Suavidad y divisibilidad son conceptos equivalentes para t-normas.

Proposición 2.5. *Una t-norma en L es suave si, y sólo si, es divisible.*

La siguiente proposición caracteriza a todas las t-normas suaves en L como ciertas sumas ordinales de copias de la t-norma de Łukasiewicz.

Proposición 2.6. *Una t-norma T en L es suave si, y sólo si, existe $J = \{0 = i_0 < i_1 < ... < i_m = 1\} \subseteq L$ tal que*

$$T(i,j) = \begin{cases} \max\{i_k, i+j-i_{k+1}\} & \text{si } i,j \in [i_k, i_{k+1}] \\ & \text{para algún } i_k \in J \\ \min\{i,j\} & \text{en caso contrario.} \end{cases}$$

Una tal T se llama suma ordinal y se representa mediante $T =< 0 = i_0, i_1, ...i_m = 1 >$.

Los operadores de indistinguibilidad fuzzifican los conceptos de igualdad y relación de equivalencia clásicas. Se han estudiado en distintos contextos, básicamente valorados en el intervalo unidad y respecto a una t-norma continua por la izquierda si bien se han estudiado también en estructuras más generales tales como los GL-monoides.

Definición 2.7.

- *Una relación borrosa en X es una aplicación $X \times X \to [0,1]$*
- *Una L-relación en X es una aplicación $X \times X \to L$.*

Definición 2.8. *Dada una t-norma T, un T-operador de indistinguibilidad (L-operador de indistinguibilidad) E en un conjunto X es una relación borrosa (una L-relatción) en X que para todo $x,y,z \in X$ satisface*

1. $E(x,x) = 1$ *(Reflexividad)*
2. $E(x,y) = E(y,x)$ *(Simetría)*
3. $T(E(x,y),E(y,z)) \leq E(x,z)$ *(T-transitividad).*

Definición 2.9. *Una relación borrosa (L-relación) R en un conjunto X es una relación de proximidad si, y sólo si, es reflexiva y simétrica.*

Definición 2.10. *Sea T una t-norma en L (en $[0,1]$) y R una relación de proximidad en un conjunto X. La clausura T-transitiva de R es el T-operador de indistinguibilidad (L-operador de indistinguibilidad) E en X que satisface*

1. $R \leq E$.
2. *Si E' es un T-operador de indistinguibilidad (L-operador de indistinguibilidad) en X tal que $R \leq E' \leq E$, entonces $E' = E$.*

3 OPENINGS TRANSITIVOS

En esta sección se expondrá un método para calcular todos los openings transitivos de una L-relación reflexiva y simétrica en un conjunto finito X. Este resultado se usará más tarde para hallar openings transitivos de una relación reflexiva y simétrica (valorada en el intervalo unidad) respecto a la t-norma de Łukasiewicz. El método es muy efectivo cuando se combina con algoritmos heurísticos como los propuestos en [3] (véase Ejemplo 3.8).

Definición 3.1. *Sea R una relación de proximidad en un conjunto X y T una t-norma en L. Un T-operador de indistinguibilidad (L-operador de indistinguibilidad) $\underline{R}$ en X es un opening T-transitivo (opening L-transitivo) de R si, y sólo si,*

- $\underline{R} \leq R$
- *Si E es otro T-operador de indistinguibilidad (L-operador de indistinguibilidad) en X que satisface $E \leq R$, entonces $E \leq \underline{R}$.*

Proposición 3.2. *Sea R una relación de proximidad en un conjunto finito $X = \{r_1, r_2, ..., r_s\}$ de cardinal s y T una t-norma. S es un T-operador de indistinguibilidad (L-operador de indistinguibilidad) menor o igual que R si, y sólo si, sus valores satisfacen el siguiente sistema de inecuaciones:*

$$\begin{aligned} 0 \leq S(r_i,r_j) &\leq R(r_i,r_j) \\ &\text{for all } i,j = 1,2,...,s. \\ T(S(r_i,r_j),S(r_j,r_k)) &\leq S(r_i,r_k) \\ &\text{para todo } i,j,k = 1,2,...,s. \\ S(r_i,r_j) &= S(r_j,r_i) \\ &\text{para todo } i,j = 1,2,...,s. \end{aligned}$$

Proof. Trivial. □

Ejemplo 3.3. *Consideremos la L-relación reflexiva y simétrica R en $X = \{a,b,c\}$ con $L = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$.*

$$R = \begin{pmatrix} 1 & \frac{2}{3} & 0 \\ \frac{2}{3} & 1 & \frac{2}{3} \\ 0 & \frac{2}{3} & 1 \end{pmatrix}.$$

Una L-relación S en X con matriz

$$S = \begin{pmatrix} 1 & p & q \\ p & 1 & r \\ q & r & 1 \end{pmatrix}$$

es un L-operador de indistinguibilidad menor o igual que R si, y sólo si,

$$\begin{aligned} 0 \leq p &\leq \frac{2}{3} \\ 0 \leq q &\leq 0 \\ 0 \leq r &\leq \frac{2}{3} \\ T(p,q) &\leq r \\ T(p,r) &\leq q \\ T(q,p) &\leq r \\ T(q,r) &\leq p \\ T(r,p) &\leq q \\ T(r,q) &\leq p. \end{aligned}$$

Si T es la t-norma de Łukasiewicz, entonces el sistema tiene 8 *soluciones:*

$$\begin{aligned} p = 0, \frac{1}{3}, \quad q = 0, \quad r = 0, \frac{1}{3}, \frac{2}{3} \\ p = \frac{2}{3}, \quad q = 0, \quad r = 0, \frac{1}{3}. \end{aligned}$$

Entre ellas hay 2 *openings L-transitivos de R:*

$$\begin{pmatrix} 1 & \frac{1}{3} & 0 \\ \frac{1}{3} & 1 & \frac{2}{3} \\ 0 & \frac{2}{3} & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & \frac{2}{3} & 0 \\ \frac{2}{3} & 1 & \frac{1}{3} \\ 0 & \frac{1}{3} & 1 \end{pmatrix}.$$

Este ejemplo muestra cómo se pueden obtener los openings transitivos de una L-relación reflexiva y simétrica. El método es muy costoso y en general necesita de muchos cálculos. Afortunadamente, se pueden hallar mejores cotas inferiores para los valores de las matrices. Por ejemplo, es bien conocido ([15]) que el ínfimo de los L-operadores de indistinguibilidad generados por las columnas de una L-relación reflexiva y simétrica R es siempre menor o igual que R. En el ejemplo anterior, este ínfimo es

$$\begin{pmatrix} 1 & \frac{1}{3} & 0 \\ \frac{1}{3} & 1 & \frac{1}{3} \\ 0 & \frac{1}{3} & 1 \end{pmatrix}$$

y, por lo tanto, las tres primeras inecuaciones se pueden reemplazar por

$$\begin{aligned} \frac{1}{3} &\leq p \leq \frac{2}{3} \\ 0 &\leq q \leq 0 \\ \frac{1}{3} &\leq r \leq \frac{2}{3}. \end{aligned}$$

Se pueden hallar mejores cotas inferiores usando métodos heurísticos (cf [3], por ejemplo) para calcular aproximaciones inferiores de relaciones borrosas reflexivas y simétricas mediante relaciones T-transitivas. Esto queda ilustrado en el Ejemplo 3.8.

Los resultados anteriores se pueden usar para obtener openings transitivos de relaciones borrosas reflexivas y simétricas valoradas en el intervalo unidad tal y como se demostrará en el Teorema 3.7.

Definición 3.4. *Dado $\alpha \in [0,1]$, sea $\lfloor \alpha \rfloor$ el mayor valor de L que satisface $\lfloor \alpha \rfloor \leq \alpha$.*

Lema 3.5. *Sea R una L-relación reflexiva y simétrica en un conjunto finito X. Si $\underline{R}$ es un L-opening transitivo de R con respecto a la t-norma de Łukasiewicz T en L y S es un opening T-transitivo de R como relación borrosa (valuada en $[0,1]$) con $\underline{R} \leq S$, entonces $\underline{R}(x,y) = \lfloor S(x,y) \rfloor$ para todo $x,y \in X$.*

Proof. $\lfloor S \rfloor$ es un L-operador de indistinguibilidad mayor o igual que $\underline{R}$. Puesto que $\underline{R}$ es un opening transitivo de R, $\underline{R} = \lfloor S \rfloor$. □

Lema 3.6. *Sea $L = \{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$, R una L-relación reflexiva y simétrica en un conjunto finito X de cardinal s y $\underline{R}$ un L-opening transitivo de R respecto a al t-norma de Łukasiewicz T. Dados $a, b \in X$ tales que $\underline{R}(a,b) < R(a,b)$ y $0 < \varepsilon < \frac{1}{n}$, consideremos la siguiente relación borrosa R':*

$$R'(x,y) = \begin{cases} \underline{R}(a,b) + \varepsilon & \text{si } (x,y) = (a,b) \\ & \text{o } (y,x) = (a,b) \\ \underline{R}(x,y) & \text{en caso contrario.} \end{cases}$$

Entonces la clausura T-transitiva $\overline{R'}$ de R' no es un opening transitivo de R en $[0,1]$.

Proof. Consideremos la L-relación R'' en X.

$$R''(x,y) = \begin{cases} \underline{R}(a,b) + \frac{1}{n} & \text{si } (x,y) = (a,b) \\ & \text{o } (y,x) = (a,b) \\ \underline{R}(x,y) & \text{en caso contrario.} \end{cases}$$

Dado que $\underline{R}$ es un opening transitivo de R en L, la clausura transitiva $\overline{R''}$ de R'' no es menor o igual que R. Por tanto existen $x, y \in X$ con $\overline{R''}(x,y) > R(x,y)$.

Como que X es finito de cardinal s,

$$\begin{aligned} &\overline{R''}(x,y) \\ &\quad = \max_{z_1,...,z_{s-1}} T(R''(x,z_1), R''(z_1,z_2), ..., R''(z_{s-1},y)) \end{aligned}$$

y este máximo se alcanza cuando contiene $R''(a,b)$. Por tanto existen $z_1, z_2, ..., z_{s-3} \in X$ tales que

$$\begin{aligned} &\overline{R''}(x,y) \\ &\quad = T(R''(x,z_1), R''(z_1,z_2), ..., R''(z_i,a), R''(a,b), \\ &\quad R''(b,z_{i+1}), R''(z_{s-3},y)) \\ &\quad = T(\underline{R}(x,z_1), \underline{R}(z_1,z_2), ..., \underline{R}(z_i,a), \underline{R}(a,b) \\ &\quad + \frac{1}{n}, \underline{R}(b,z_{i+1}), \underline{R}(z_{s-3},y)). \end{aligned}$$

$$\begin{aligned} &T(\underline{R}(x,a), \underline{R}(a,b) + \frac{1}{n}, \underline{R}(b,y)) \\ &\quad \geq T(R''(x,z_1), R''(z_1,z_2), ..., R''(z_i,a), R''(a,b), \\ &\quad R''(b,z_{i+1}), R''(z_{s-3},y)) \\ &\quad = \overline{R''}(x,y) > R(x,y) \end{aligned}$$

y por consiguiente

$$T(\underline{R}(x,a), \underline{R}(a,b), \underline{R}(b,y)) = \underline{R}(x,y) = R(x,y).$$

Así,

$$R'(x,y) \geq T(\underline{R}(x,a), \underline{R}(a,b) + \varepsilon, \underline{R}(b,y)) > R(x,y).$$

y en consecuencia $\overline{R'}$ no es un opening transitivo de R.

□

Teorema 3.7. *Sea R una L-relación reflexiva y simétrica en un conjunto finito X de cardinal s y $\underline{R}$ un L-opening transitivo de R respecto a la t-norma de Łukasiewicz. Entonces $\underline{R}$ también es un T-opening transitivo de R como relación borrosa (valuada en $[0,1]$).*

Proof. Consideremos una relación borrosa S en X tal que

$$\underline{R} < S.$$

Entonces existen $a, b \in X$ con $\underline{R}(a,b) < S(a,b)$. La relación borrosa R' del anterior Lema 3.6 con $\varepsilon \leq S(a,b) - \underline{R}(a,b)$ satisface

$$\underline{R} < R' \leq S.$$

Puesto que la clausura transitiva $\overline{R'}$ de R' es mayor que R y la aplicación que asigna la clausura transitiva a una relación reflexiva y simétrica es no decreciente, a fortiori S es mayor que R. Por consiguiente no hay T-operadores de indistinguibilidad entre $\underline{R}$ y R. □

El siguiente ejemplo muestra la efectividad del método propuesto para calcular openings transitivos cuando se combina con algoritmos de tipo heurístico.

Ejemplo 3.8. *Consideremos la relación borrosa reflexiva y simétrica R en un conjunto de cardinal* 5 *con matriz*

$$R = \begin{pmatrix} 1.0 & 0.5 & 0.5 & 0.7 & 0.6 \\ 0.5 & 1.0 & 0.8 & 0.7 & 0.5 \\ 0.5 & 0.8 & 1.0 & 0.9 & 0.8 \\ 0.7 & 0.7 & 0.9 & 1.0 & 0.9 \\ 0.6 & 0.5 & 0.8 & 0.9 & 1.0 \end{pmatrix}.$$

En [3] se usa un algoritmo para calcular un T-operador de indistinguibilidad S (T la t-norma de Łukasiewicz) que no es un opening transitivo de R, pero que es menor que R y "cercano" a un opening transitivo:

$$S = \begin{pmatrix} 1.0 & 0.5 & 0.5 & 0.6 & 0.6 \\ 0.5 & 1.0 & 0.8 & 0.6 & 0.5 \\ 0.5 & 0.8 & 1.0 & 0.8 & 0.7 \\ 0.6 & 0.6 & 0.8 & 1.0 & 0.9 \\ 0.6 & 0.5 & 0.7 & 0.9 & 1.0 \end{pmatrix}.$$

R (y S) pueden también considerarse L-relaciones, donde $L = \{0, \frac{1}{10}, \frac{2}{10}, ..., 1\}$. Hay sólo 2 *$T$-operadores de indistinguibilidad intermedios valuados en L y que se pueden obtener de la misma manera que tras el Ejemplo 3.3:*

$$\begin{pmatrix} 1.0 & 0.5 & 0.5 & 0.6 & 0.6 \\ 0.5 & 1.0 & 0.8 & 0.6 & 0.5 \\ 0.5 & 0.8 & 1.0 & 0.8 & 0.7 \\ 0.6 & 0.6 & 0.8 & 1.0 & 0.9 \\ 0.6 & 0.5 & 0.7 & 0.9 & 1.0 \end{pmatrix}$$

$$\begin{pmatrix} 1.0 & 0.5 & 0.5 & 0.7 & 0.6 \\ 0.5 & 1.0 & 0.8 & 0.6 & 0.5 \\ 0.5 & 0.8 & 1.0 & 0.8 & 0.7 \\ 0.6 & 0.6 & 0.8 & 1.0 & 0.9 \\ 0.6 & 0.5 & 0.7 & 0.9 & 1.0 \end{pmatrix}.$$

La segunda matriz es un opening transitivo de R considerada como una L-relación y, por lo tanto, también como una relación borrosa (valorada en $[0,1]$*).*

4 CONCLUSIONES

Se ha propuesto un método para hallar todos los openings transitivos de una L-relación reflexiva y simétrica que se ha usado más tarde para hallar openings transitivos de una relación reflexiva y simétrica (valorada en $[0,1]$).

El algoritmo para calcular los openings transitivos de una relación reflexiva y simétrica respecto a la t-norma Ł de Łukasiewicz se puede usar para otras t-normas arquimedianas continuas. En efecto, si T es una t-norma arquimediana continua con generador aditivo t, $\alpha \in (0,1]$ y E es un T-operador de indistinguibilidad en un conjunto finito X, entonces la relación borrosa E' definida por $E'(x,y) = \min(1-\frac{1}{\alpha}t \circ E(x,y),0)$ para todo $x,y \in X$ es un Ł-operador de indistinguibilidad en X [2]. De este modo se puede transferir un opening transitivo respecto a la t-norma de Łukasiewicz a un opening transitivo respecto a otra t-norma arquimediana continua.

AGRADECIMIENTOS.

Este trabajo ha sido parcialmente financiado por el proyecto número TIN2009-07235.

Referencias

[1] D. Boixader, J. Jacas, J. Recasens (2000). Fuzzy Equivalence Relations: Advanced Material. In Dubois, Prade Eds. *Fundamentals of Fuzzy Sets*, Kluwer, 261-290.

[2] D. Boixader, J. Recasens (2011). Indistinguishability operators with respect to different t-norms. Submitted.

[3] B. De Baets, H. De Meyer (2003). Transitive approximation of fuzzy relations by alternating closures and openings. Soft Computing 7 210-219.

[4] P. Dawyndt, H. De Meyer, B. De Baets (2005). The complete linkage clustering algorithm revisited. Soft Computing 9 85-392.

[5] J.C. Fodor, M. Roubens (1995). Structure of transitive valued binary relations. Math. Soc. Sci. 30, 71-94.

[6] L. Garmendia, R. González del Campo, V. López, J. Recasens (2009). An Algorithm to Compute the Transitive Closure, a Transitive Approximation and a Transitive Opening of a Fuzzy Proximity. Mathware & Soft Computing 16, 175-191.

[7] J. Jacas (1990). Similarity relations - the calculation of minimal generating families. *Fuzzy Sets and Systems* 35, 151-162.

[8] F. Klawonn, J. L. Castro (1995). Similarity in Fuzzy Reasoning. Mathware & Soft Computing 3, 197-228.

[9] U. Höhle (1992) M-valued sets and sheaves over integral, commutative cl-monoids. In: S.E. Rodabaughg et al Eds. *Applications of Category Theory to Fuzzy Subsets*, Kluwer, 33-72.

[10] M. Mas, M. Monserrat and J. Torrens (2003) QL-Implications on a finite chain. Proc. Eusflat 2003. Zittau. 281-284.

[11] G. Mayor, J. Torrens (2005) Triangular norms on discrete settings. In: Klement, E.P., Mesiar, R. (Eds.), Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms, Elsevier , Amsterdam. 189-230.

[12] E. P. Klement, R. Mesiar, E. Pap (2000). *Triangular norms*. Kluwer. Dordrecht.

[13] J. Recasens. (2011) Indistinguishability Operators. Modelling Fuzzy Equalities and Fuzzy Equivalence Relations. Studies in Fuzziness and Soft Computing 260. Springer.

[14] B. Schweizer, A. Sklar (1983) *Probabilistic Metric Spaces*. North-Holland. Amsterdam.

[15] L.Valverde (1985). On the structure of F-indistinguishability operators, *Fuzzy Sets and Systems* 17, 313-328.

[16] L.A.Zadeh (1971). Similarity relations and fuzzy orderings, *Information Science* 3, 177-200.

AUTOCONTRADICCIÓN EN EL CONJUNTO DE LAS FUNCIONES DE $[0,1]^{[0,1]}$ NORMALES Y CONVEXAS

Pablo Hernández[1], **Susana Cubillo**[2], **Carmen Torres-Blanc**[2]

[1]*Departamento de Matemática y Física, Universidad Nacional Experimental del Táchira, Avenida Universidad, San Cristóbal, Táchira, Venezuela*, phernandezv@unet.edu.ve

[2]*Departamento de Matemática Aplicada, Universidad Politécnica de Madrid, 28660 Boadilla del Monte-Madrid, España*, {scubillo,ctorres}@fi.upm.es

Resumen

Las propiedades N-autocontradicción y autocontradicción, han sido suficientemente estudiadas en los conjuntos borrosos ordinarios y en los conjuntos borrosos intuicionistas de Atanassov. En el presente artículo se inicia el estudio de las mencionadas propiedades, dentro del marco de los conjuntos borrosos de tipo 2 cuyos grados de pertenencia son funciones normales y convexas (**L**). En este sentido, aquí se extienden los conceptos de N-autocontradicción y autocontradicción al conjunto **L**, y se establecen algunos criterios para verificar tales propiedades.

Palabras Clave: Conjuntos borrosos de tipo 2, funciones normales y convexas, N-autocontradicción, autocontradicción.

1. INTRODUCCIÓN

Los conjuntos borrosos de tipo 2 (T2FSs) fueron introducidos por L.A. Zadeh en 1975 [17], como una extensión de los conjuntos borrosos ordinarios (FSs, o conjuntos borrosos de tipo 1). Mientras que en estos últimos el grado de pertenencia de un elemento al conjunto viene determinado por un valor en el intervalo $[0,1]$, en el caso de los T2FSs el grado de pertenencia de un elemento es un conjunto borroso en [0,1], es decir, un T2FS queda determinado por una función de pertenencia $\mu : X \to [0,1]^{[0,1]}$, ver [9, 10, 11]. Por ello, para poder trabajar con dichos conjuntos, es necesario hacerlo previamente en el espacio de funciones $[0,1]^{[0,1]}$. Particularmente en este trabajo nos restringiremos al espacio **L**, de las funciones de $[0,1]^{[0,1]}$ normales y convexas.

Trillas et al. en [13, 14], iniciaron el estudio de la autocontradicción y contradicción de conjuntos en el contexto de la lógica borrosa, no sólo como un tema de interés teórico, sino también práctico, en el sentido de analizar las posibles consecuencias contradictorias en los procesos de inferencia borrosa. En artículos posteriores (ver [1, 4], por ejemplo): Se abordó con más profundidad la autocontradicción y contradicción en los FSs; se propusieron formas de medir el grado en que estas propiedades aparecen; se desarrolló una axiomática para las medidas de autocontradicción, contradicción entre conjuntos, autocontradicción respecto a una negación fuerte N (N-autocontradicción) y contradicción entre conjuntos respecto a una negación fuerte N (N-contradicción).

Así mismo, en [5] se aborda la contradicción en el marco de los conjuntos borrosos intuicionistas de Atanassov (AIFSs), introduciendo el concepto de un AIFS contradictorio, y la contradicción entre dos AIFSs. Luego, en diferentes artículos (ver [2, 3, 6, 12], por ejemplo): Se estudian los grados de autocontradicción y de N-autocontradicción de un AIFS; se proporcionan algunas fórmulas para medir dichos grados; se desarrollan las axiomáticas para medir dichas propiedades; se trata con más profundidad el tema de la contradicción entre AIFSs; se presentan algunas fórmulas para medir tal propiedad; y se proporciona un modelo axiomático para las medidas de contradicción entre AIFSs. En el presente trabajo se pretende iniciar con un estudio análogo al desarrollado anteriormente, especialmente en [2, 5, 13, 14], sobre la N-autocontradicción y autocontradicción, pero en los conjuntos de tipo 2, con grados de pertenencia en **L**.

El artículo se organiza como sigue: En la sección 2 recordamos algunas definiciones, operaciones y propiedades sobre los T2FSs, especialmente acerca del subálgebra **L**, presentadas en la literatura reciente (ver por ejemplo, [16]). En la sección 3 presentamos las definiciones y propiedades sobre la autocontradicción y N-contradicción en conjuntos borrosos, establecidas en artículos anteriores (por ejemplo, [14]); y extendemos algunas de estas definiciones al subálgebra **L**, el cual representa nuestro objeto de estudio. Por último, en la sección 4 determinamos algunos criterios para verificar dichas propiedades en **L**.

2. SOBRE LOS T2FSs

De aquí en adelante denotaremos por X un conjunto no vacío que representará el universo de discurso. Además, el símbolo $\leq$ denotará la relación de orden en el retículo de los números reales.

Definición 1. ([10, 11]) Un conjunto borroso de tipo 2, A, esta caracterizado por una *función borrosa de pertencia* μ_A, como

$$\mu_A : X \to [0,1]^{[0,1]},$$

donde $\mu_A(x)$ es un *grado borroso*, y es un conjunto borroso en el intervalo $[0,1]$. $\mu_A(x)$ es representado por

$$\mu_A(x) = f_x$$

donde

$$f_x : [0,1] \to [0,1].$$

Las operaciones teóricas en los T2FSs fueron deducidas a partir del principio de extensión de Zadeh [9, 10, 11, 17]. Luego, en [9] establecieron el "Teorema de Representación", mediante el cual se ratifican las fórmulas para dichas operaciones. Como los grados de verdad de los T2FSs están en $[0,1]^{[0,1]}$, entonces trabajaremos en este marco.

Definición 2. ([8, 16]) Sobre $[0,1]^{[0,1]}$ se definen las operaciones $\sqcup$ (unión), $\sqcap$ (intersección), N (complemento, o negación), $\bar{0}$ y $\bar{1}$ de la siguiente manera:

$$(f \sqcup g)(x) = f(x) \sqcup g(x) = sup\{f(y) \wedge g(z) : y \vee z = x\},$$

$$(f \sqcap g)(x) = f(x) \sqcap g(x) = sup\{f(y) \wedge g(z) : y \wedge z = x\},$$

$(N(f))(x) = \neg f(x) = sup\{f(y) : n(y) = x\} = f(n(x)) = f(1-x)$,

$$\bar{0}(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0 \end{cases}, \ \bar{1}(x) = \begin{cases} 1 & \text{si } x = 1 \\ 0 & \text{si } x \neq 1. \end{cases}$$

$\vee, \wedge$ y n son las operaciones máximo, mínimo, y la negación estandar ($n = 1 - id$), respectivamente, en el retículo $[0,1]$.

$\mathbf{M}$= $([0,1]^{[0,1]}, \sqcup, \sqcap, N, \bar{0}, \bar{1})$ es el álgebra básica de los valores de verdad de los T2FSs, [8, 16]. $\mathbf{M}$ no tiene estructura de retículo, ya que generalmente no se cumple la ley de absorción [8, 16]. Sin embargo, cada una de las operaciones $\sqcup$ y $\sqcap$ tienen las propiedades requeridas para definir un orden parcial, [16].

Definición 3. ([11, 16]) $f \sqsubseteq g$ si $f \sqcap g = f$; $f \preceq g$ si $f \sqcup g = g$.

En general, estos dos órdenes no coinciden, [11, 16]. Por otra parte, $\bar{0}$ y $\bar{1}$ son los elementos neutros de las operaciones $\sqcup$ y $\sqcap$, respectivamente, es decir, $\bar{0} \sqcup f = f$ y $\bar{1} \sqcap f = f$, para todo $f \in \mathbf{M}$, ver [16]. Además, $\bar{0}$ es el menor elemento (cota inferior) del orden parcial $\preceq$, ya que $\bar{0} \preceq f, \ \forall f \in \mathbf{M}$; y $\bar{1}$ es el mayor elemento (cota superior) del orden parcial $\sqsubseteq$, porque $f \sqsubseteq \bar{1}, \ \forall f \in \mathbf{M}$, ver [16].

Para facilitar las operaciones en el álgebra $\mathbf{M}$, se presentan las siguientes operaciones auxiliares.

Definición 4. ([8, 16]) Para $f \in \mathbf{M}$, se definen $f^L, f^R \in \mathbf{M}$ como

$$f^L(x) = sup\{f(y) : y \leq x\}, \ f^R(x) = sup\{f(y) : y \geq x\}.$$

En este estudio nos centraremos en el conjunto de las funciones de $[0,1]^{[0,1]}$ normales y convexas.

Definición 5. ([8, 11, 16]) Una función $f \in \mathbf{M}$ es normal si $sup\{f(x) : x \in [0,1]\} = 1$, o equivalentemente, si $f^L \vee f^R = 1$.

Definición 6. ([8, 11, 16]) Una función $f \in \mathbf{M}$ es convexa, si para cualquier $x \leq y \leq z$, se tiene que $f(y) \geq f(x) \wedge f(z)$, o equivalentemente, si $f = f^L \wedge f^R$.

El espacio $\mathbf{L}$ de las funciones de $[0,1]^{[0,1]}$ normales y convexas, es un subálgebra de $\mathbf{M}$; y es un retículo acotado y completo, con una negación que satisface las leyes de Morgan, ver [8, 11, 16]. En $\mathbf{L}$, los órdenes parciales de $\mathbf{M}$ (ver definición 3) son equivalentes, es decir $\sqsubseteq \equiv \preceq$. Por otra parte, tenemos la siguiente caracterización, que nos ayudará a establecer nuevos resultados:

Teorema 1. ([8]) Sean $f, g \in \boldsymbol{L}$. $f \sqsubseteq g$ si y sólo si $g^L \leq f^L$ y $f^R \leq g^R$.

Si f es normal, evidentemente $f^L(x) = 1$, o $f^R(x) = 1$, para todo $x \in [0,1]$; además, $f^L(1) = 1$ y $f^R(0) = 1$, ver [8, 16]. Si f es convexa, se satisface $f(x) = f^L(x)$, o $f(x) = f^R(x)$, para todo $x \in [0,1]$, ver [8]. En virtud de lo anterior, se puede afirmar que si f es normal y convexa, entonces en cada $x \in [0,1]$, ocurren únicamente dos posibilidades:

$$f^L(x) = f(x) \leq 1, \ f^R(x) = 1$$

o

$$f^R(x) = f(x) \leq 1, \ f^L(x) = 1.$$

Definición 7. ([8]) Sea $h : [0,1] \to [0,1]$. Se define $h^* : [0,1] \to [0,2]$ por

$$h^*(x) = \begin{cases} 2 - h(x) & \text{si } h(x) = h^L(x) \\ h(x) & \text{en otro caso.} \end{cases}$$

Esta definición 7 nos permite establecer una caracterización del orden parcial en L

Teorema 2. ([8]) Sean $f, g \in \mathbf{L}$. $f \sqsubseteq g \Leftrightarrow f^* \leq g^*$.

3. SOBRE LA AUTOCONTRADICCIÓN

Debido a que el análisis a realizar emplea negaciones fuertes, empezamos recordando este concepto en los FSs, así como algunas de sus propiedades.

Definición 8. ([7, 15]) Una función decreciente $n : [0,1] \to [0,1]$ tal que $n(0) = 1, n(1) = 0$, se dice que es una negación. Si además satisface que $n(n(x)) = x$, para todo $x \in [0,1]$, se le dice negación fuerte o involución.

En [15] se demuestra que n es una negación fuerte en $[0,1]$, si y sólo si existe un automorfismo de orden $g : [0,1] \to [0,1]$, tal que $n(x) = g^{-1}(1 - g(x))$, para todo $x \in [0,1]$. Toda negación fuerte n es estrictamente decreciente, continua y biyectiva. Por otra parte, cada negación fuerte n tiene un único punto fijo $p \in (0,1)$, tal que $n(p) = p$, donde $p = g^{-1}(1/2)$.

Extendiendo el concepto de negación al álgebra $\mathbf{M}$, tenemos la siguiente definición de negación de tipo 2:

Definición 9. Una función decreciente $N : \mathbf{M} \to \mathbf{M}$, tal que $N(\bar{0}) = \bar{1}, N(\bar{1}) = \bar{0}$, es una negación. Si además satisface que $N(N(f)) = f$, para todo $f \in \mathbf{M}$, se le dice negación fuerte o involución.

La función N de la definición 2 satisface algunos axiomas para definir negación fuerte. Por ejemplo, $N(\bar{0}) = \bar{1}, N(\bar{1}) = \bar{0}$ y $N(N(f)) = f$, para todo $f \in \mathbf{M}$, ver [16]; sin embargo, no es decreciente en cada orden parcial, ya que en [16] se demuestra que $f \sqsubseteq g$ si y sólo si $\neg g \preceq \neg f$, o equivalentemente, $f \preceq g$ si y sólo si $\neg g \sqsubseteq \neg f$, para todo $f, g \in \mathbf{M}$. Sin embargo, si nos limitamos al subálgebra $\mathbf{L}$, dicha función N asociada a cualquier negación fuerte de tipo 1, cumple con todos los axiomas de negación fuerte; por este motivo decidimos estudiar, por el momento, la contradicción en $\mathbf{L}$, y no en todo $\mathbf{M}$.

En [16] se presenta la siguiente propiedad de la negación de tipo 2, para todo $f \in \mathbf{M}$

$$(\neg f)^L = \neg(f^R),\ (\neg f)^R = \neg(f^L),$$

y como $(N(f))(x) = f(n(x))$, entonces, $\forall x \in [0,1]$,

$$(N(f))^L(x) = f^R(n(x)),\ (N(f))^R(x) = f^L(n(x)). \quad (1)$$

Cabe destacar que todo el estudio de ahora en adelante se hará exclusivamente considerando negaciones fuertes de tipo 1 y de tipo 2, las cuales llamaremos negaciones, simplemente.

En cuanto a los conceptos de n-autocontradicción y autocontradicción en los FSs, tenemos:

Definición 10. ([14]) Sea $\mu \in [0,1]^X$ y sea $n : [0,1] \to [0,1]$ una negación. Diremos que μ es *n-autocontradictorio* si y sólo si $\mu(x) \leq n(\mu(x))$, para todo $x \in X$. Decimos que μ es *autocontradictorio*, si es n-autocontradictorio respecto a alguna negación n.

En el álgebra $\mathbf{L}$, la n-autocontradicción y autocontradicción las definimos de la siguiente manera:

Definición 11. Sea $f \in \mathbf{L}$ y sea $N : \mathbf{L} \to \mathbf{L}$ una negación. Diremos que f es *N-autocontradictorio* si y sólo si $f(x) \sqsubseteq N(f(x))$, para todo $x \in X$. Decimos que f es *autocontradictorio*, si es N-autocontradictorio respecto a alguna negación N.

En general, un conjunto borroso de tipo 2 (ver definición 1), B, cuyos grados de pertenencias f_x están en $\mathbf{L}$, es N-autocontradictorio si todos sus grados f_x lo son. Igualmente, B es autocontradictorio si todos sus grados f_x lo son.

4. CRITERIOS DE AUTOCONTRADICCIÓN EN L

En esta sección se presentan los principales resultados, correspondientes a los criterios para verificar las propiedades de N-autocontradicción y de autocontradicción en $\mathbf{L}$. En toda esta sección 4, entre otros aspectos, tendremos presente que toda negación fuerte de tipo 1, n, es estrictamente decreciente, continua y biyectiva.

Proposición 1. Sea $f \in \mathbf{L}$, y sea $N : \mathbf{L} \to \mathbf{L}$ una negación. Entonces f es N-autocontradictorio si y sólo si

$$f^* \leq (N(f))^*.$$

Demostración. Directa, aplicando la definición 11 y el teorema 2. □

Observación 1. Basta que $(N(f))^*(x) > f^*(x)$ en algún $x \in [0,1]$, para que f no sea N-autocontradictorio.

Ejemplo 1.

1) Sea $f \in \mathbf{L}$, $f(x) = 1 - x$, la cual es estrictamente decreciente, por tanto $f^*(x) = f(x) = 1 - x$ (ver definición 7). Sea la negación $N : \mathbf{L} \to \mathbf{L}$, $(N(f))(x) = f(n(x)) = f(1-x) = x$, la cual es estrictamente creciente, por tanto $(N(f))^*(x) = 2 - (N(f))(x) = 2 - x$ (ver definición 7). Evidentemente, $f^*(x) \leq (N(f))^*(x)$, para todo $x \in [0,1]$, en consecuencia, f es N-autocontradictorio.

2) Sea $f \in \mathbf{L}$, $f(x) = x^2$, la cual es estrictamente creciente, por ende $f^*(x) = 2 - f(x) = 2 - x^2$. Sea la negación $N : \mathbf{L} \to \mathbf{L}$, $(N(f))(x) = f(n(x)) = f(\sqrt{1-x^2}) = 1 - x^2$, la cual es estrictamente decreciente, por tanto $(N(f))^*(x) = (N(f))(x) = 1 - x^2$. Obviamente, $f^*(x) > (N(f))^*(x)$, para todo $x \in [0,1]$, por consiguiente, f no es N-autocontradictorio.

Proposición 2. Sean $f \in \mathbf{L}$, y sea $N : \mathbf{L} \to \mathbf{L}$ una negación, asociada a una negación de tipo 1, n. Entonces f es N-autocontradictorio si y sólo si

$$(N(f))^L \leq f^L,\ f^R \leq (N(f))^R.$$

Equivalentemente, si y sólo si

$$f^R(n(x)) \leq f^L(x),\ f^R(x) \leq f^L(n(x)), \forall x \in [0,1]$$

.

Demostración. f es N-autocontradictorio si y sólo si

$$f \sqsubseteq N(f).$$

Es decir (ver teorema 1), si y sólo si

$$(N(f))^L \leq f^L,\ f^R \leq (N(f))^R.$$

Tomando en cuenta (1), esto equivale a

$$f^R(n(x)) \leq f^L(x),\ f^R(x) \leq f^L(n(x)), \forall x \in [0,1].$$

□

Corolario 1. f es N-autocontradictorio si y sólo si

$$f^R(x) \leq f^L(n(x)), \forall x \in [0,1].$$

Equivalentemente, si y sólo si

$$f^R(n(x)) \leq f^L(x), \forall x \in [0,1].$$

Observación 2. Basta que $(N(f))^L(x) > f^L(x)$, o $f^R(x) > (N(f))^R(x)$ en algún $x \in [0,1]$, para que f no sea N-autocontradictorio. En otros términos, es suficiente que $f^R(x) > f^L(n(x))$ en algún $x \in [0,1]$, para que f no sea N-autocontradictorio. O equivalentemente, basta que $f^R(n(x)) > f^L(x)$ en algún $x \in [0,1]$, para que f no sea N-autocontradictorio.

Ejemplo 2. Sean los elementos $\bar{0}, \bar{1} \in \mathbf{L}$, en donde $N(\bar{0}) = \bar{1}$ y $N(\bar{1}) = \bar{0}$, para cualquier $N : \mathbf{L} \to \mathbf{L}$.

1) $\bar{0}$ es N-autocontradictorio respecto a cualquier negación N, ya que $(N(\bar{0}))^L(x) = \bar{1}^L(x) = \bar{1}(x) \leq \bar{0}^L(x) = 1$, y $\bar{0}^R(x) = \bar{0}(x) \leq (N(\bar{0}))^R(x) = \bar{1}^R(x) = 1$, para todo $x \in [0,1]$. De acuerdo a la definición 11, también podemos decir que $\bar{0}$ es autocontradictorio.

2) $\bar{1}$ no es N-autocontradictorio respecto a ninguna N, ya que, por ejemplo, en $x = 0$, $(N(\bar{1}))^L(0) = \bar{0}^L(0) = 1 > \bar{1}^L(0) = 0$. Según la definición 11, $\bar{1}$ no es autocontradictorio.

Corolario 2. Sea $f \in \mathbf{L}$. Si $f(0) < f(1)$ entonces f no es N-autocontradictorio respecto a ninguna negación $N : \mathbf{L} \to \mathbf{L}$, es decir, f no es autocontradictorio.

Demostración. Sea $f(0) < f(1)$, y sabiendo que

$$f^L(0) = f(0),\ f^R(0) = 1,\ f^L(1) = 1,\ f^R(1) = f(1);$$

$$\Rightarrow \text{ en } x = 1,\ f^L(0) = f^L(n(1)) = f(0) < f(1) = f^R(1);$$

para toda negación de tipo 1, n. Por tanto, decimos que f no es autocontradictorio. □

Observación 3.

1) Si $f \in \mathbf{L}$ es monótona creciente, excluyendo la función constante $f = 1$, entonces sucede que $f(0) < f(1) = 1$, por tanto, según corolario 2, f no es autocontradictorio.

2) Si f no es autocontradictorio, generalmente no implica que $f(0) < f(1)$. Por ejemplo, la función presentada en la figura 1(a) no es autocontradictoria, sin embargo $f(0) > f(1)$.

3) Si $f(0) \geq f(1)$, generalmente no implica que f es autocontradictoria. Por ejemplo, en la función de la figura 1(a), $f(0) > f(1)$, sin embargo no es autocontradictoria.

Una de las propiedades a emplear en el siguiente teorema, es: En toda $f \in \mathbf{L}$ existe un conjunto no vacío $W \subseteq [0,1]$, tal que, para todo $w \in W$

$$\lim_{x \to w^+} f(x) = 1, \text{ o } \lim_{x \to w^-} f(x) = 1, \text{ o } f(w) = 1.$$

Además, cualquier x entre dos elementos de W, presenta $f(x) = 1$, por tanto también está en W. Esto es así por que si $y < u < z$, para todo $y, z \in W$, entonces $f^L(u) = 1$ y $f^R(u) = 1$. Y debido a que f es convexa, resulta $f(u) = f^L(u) \wedge f^R(u) = 1$. En este caso, decimos que f es normalizado, ya que $f(x) = 1$ en algún $x \in [0,1]$.

El siguiente teorema proporciona una caracterización de los conjuntos N-contradictorios.

Teorema 3. Sean $f \in \mathbf{L}$, $N : \mathbf{L} \to \mathbf{L}$ una negación, asociada a una negación de tipo 1, n, con punto fijo p. Sea $\alpha = inf\{w : w \in W\}$. Entonces f es N-autocontradictorio si y sólo si

$$p \geq \alpha \text{ y } \forall x \leq \alpha,\ es\ f(n(x)) \leq f(x). \tag{2}$$

Demostración. Para la demostración utilizaremos el hecho de que f es N-autocontradictorio si y sólo si

$$f^R(n(x)) \leq f^L(x), \forall x \in [0,1].$$

"⇒" Supongamos que (2) no es cierto y veamos que entonces f no es N-autocontradictorio.

Si (2) no es cierto, se pueden dar dos casos:

I) $p < \alpha$, entonces $\exists t$, tal que $p < t < \alpha$, y $n(t) < p < \alpha$. En consecuencia,

$$f^R(n(t)) = 1,\ f^L(t) = f(t) < 1.$$

Luego,

$$f^R(n(t)) > f^L(t),$$

y no es N-autocontradictorio.

II) $p \geq \alpha$, y $\exists x \leq \alpha$ con $f(n(x)) > f(x)$. En este caso, $x \neq n(x)$. Si $x < \alpha$, $n(x) > \alpha$, entonces tenemos

$$f^R(n(x)) = f(n(x)) > f(x) = f^L(x),$$

y no es N-autocontradictorio.

Sea $x = \alpha$, $n(x) > \alpha$. Para $x = \alpha$, puede ocurrir que:

$$a)\ f^L(\alpha) = f(\alpha) < 1,\ f^R(\alpha) = 1;$$

o

$$b)\ f^R(\alpha) = f(\alpha) \leq 1,\ f^L(\alpha) = 1.$$

Como $f(n(\alpha)) > f(\alpha) < 1$, y $f^R(\alpha) = sup\{f(x) : x \geq \alpha\} \geq f(n(\alpha)) > f(\alpha)$, entonces solamente puede ocurrir el apartado *a*); luego,

$$f^R(n(\alpha)) = f(n(\alpha)) > f(\alpha) = f^L(\alpha),$$

y no es N-autocontradictorio.

"$\Leftarrow$" Comprobemos que para todo $x \in [0,1]$, $f^R(n(x)) \leq f^L(x)$. En este sentido, se pueden presentar cuatro posibilidades:

i) $x < \alpha$, entonces $n(x) > n(\alpha) \geq \alpha$, lo cual implica que

$$f^R(x) = 1 = f^L(n(x)),$$

$$f^R(n(x)) = f(n(x)) \leq f(x) = f^L(x).$$

ii) $x = \alpha$, por tanto $n(x) = n(\alpha) \geq \alpha$. Si $x = \alpha = n(x) = p \in (0,1)$, entonces no puede darse el apartado *a*), ya que aquí $\exists t$, $t < \alpha$ y $n(t) > \alpha$, tal que

$$f(t) \leq f^L(\alpha) = f(\alpha) < f(n(t)) \leq f^R(\alpha) = 1,$$

lo cual contradice la hipotesis inicial. En cuanto al apartado *b*), tenemos

$$f^R(\alpha) = f^R(n(\alpha)) = f(\alpha) \leq f^L(\alpha) = f^L(n(\alpha)) = 1.$$

Por otra parte, si $x = \alpha$ y $n(\alpha) > \alpha$, y se da el apartado *a*), entonces

$$f^R(n(\alpha)) = f(n(\alpha)) \leq f(\alpha) = f^L(\alpha),$$

$$f^L(n(\alpha)) = 1 = f^R(\alpha).$$

Y si se cumple el apartado *b*), tenemos

$$f^R(n(\alpha)) = f(n(\alpha)) \leq 1 = f^L(\alpha),$$

$$f^R(\alpha) \leq 1 = f^L(n(\alpha)).$$

iii) $\alpha < x < n(\alpha)$, entonces $\alpha < n(x) < n(\alpha)$. Aquí se cumple que

$$f^R(x) = f(x) \leq 1 = f^L(n(x)),$$

$$f^R(n(x)) = f(n(x)) \leq 1 = f^L(x).$$

iv) $x \geq n(\alpha) \geq \alpha$, en consecuencia $n(x) \leq \alpha$. Haciendo el cambio de variable $u = n(x)$, con $n(u) = n(n(x)) = x$, y $f(n(u)) = f(n(n(x)) = f(x) \leq f(n(x)) = f(u)$, entonces, este caso iv) equivale a la unión de los casos i) y ii).

Por tanto, $f^R(n(x)) \leq f^L(x)$, $\forall x \in [0,1]$; luego, f es N-autocontradictorio.

□

Corolario 3. f no es N-autocontradictorio si y sólo si

$$p < \alpha,\ \text{o}\ \exists x \leq \alpha \text{ tal que } f(n(x)) > f(x).$$

Ejemplo 3. Sea la negación $N : \mathbf{L} \to \mathbf{L}$, $(N(f))(x) = f(1-x)$, $\forall x \in [0,1]$, asociada a la negación $n(x) = 1-x$, con punto fijo $p = 0.5$. Y sean tres funciones diferentes de $[0,1]^{[0,1]}$ normales y convexas, mostradas en la siguiente figura.

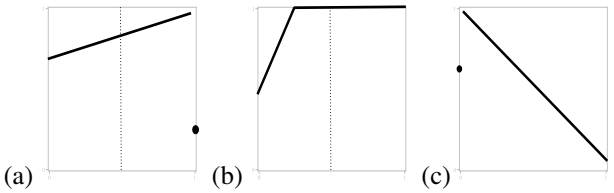


Figura 1: Ejemplos de funciones en **L**.

1) La función dibujada en la figura 1(a), no es N-autocontradictoria respecto a la negación dada, ya que $0.5 = p < \alpha = 1$. Además, esa función no es autocontradictoria porque $p < \alpha = 1, \forall p \in (0,1)$.

2) La función en la figura 1(b), no es N-autocontradictoria respecto a la negación dada, porque, por ejemplo, en $x = 0 < \alpha = 0.25$; $f(n(0)) = f(1) > f(0)$. Nótese que lo mismo sucede para cualquier negación $N : \mathbf{L} \to \mathbf{L}$, por lo que no es autocontradictoria.

3) La función presentada en la figura 1(c) es N-autocontradictoria respecto a cualquier negación $N : \mathbf{L} \to \mathbf{L}$, ya que $p > \alpha = 0$ y $\forall x \leq \alpha$, es $f(n(x)) < f(x)$.

Observación 4.

1) Si $f \in \mathbf{L}$ es monótona decreciente, incluyendo la función constante $f = 1$, entonces f es N-autocontradictorio respecto a cualquier negación N, y por tanto, f es autocontradictorio. Esto es así porque en este caso, $f(0) = 1$, entonces $x = 0 = \alpha$, con $f(n(0)) \leq 1 = f(0)$. Es decir, para toda negación $N : \mathbf{L} \to \mathbf{L}$, se cumple que $p \geq \alpha$ y $\forall x \leq \alpha$, es $f(n(x)) \leq f(x)$.

2) Que f sea N-autocontradictorio respecto a cualquier negación N, generalmente no implica que f sea monótona decreciente. Por ejemplo, la función en la gráfica 1(c) es N-autocontradictoria respecto a cualquier negación N, no obstante, la función no es monótona decreciente.

5. CONCLUSIONES

En este artículo se ha extendido los conceptos de N-autocontradicción y autocontradicción establecidos previamente en los FSs y los AIFSs, al marco de los conjuntos tipo 2 con grados de pertenencia en el conjunto de funciones **L**. A partir de dichos conceptos, han presentado algunos criterios que permiten verificar las propiedades de

N-autocontradicción y autocontradicción de un elemento en **L**, llegando a una caracterización. Desde el punto de vista práctico, tales resultados permiten analizar las posibles consecuencias contradictorias en los procesos de inferencia borrosa, en los que aparecen predicados modelizados mediante conjuntos tipo 2. Otros aspectos que merecen ser abordados en futuras investigaciones podrían ser: Estudiar la contradicción entre conjuntos en **L** y establecer los criterios para verificar dicha propiedad; finalmente, proporcionar métodos para medir la autocontradicción, N-autocontradicción, y la contradicción entre conjuntos en **L**.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la CICYT dentro del proyecto TIN2008-06890-C02-01, por la UPM-CAM, por el FONACIT, y por el Decanato de Investigación de la UNET. Los autores también agradecen a los árbitros por las valiosas sugerencias, las cuales mejoran la calidad del presente artículo.

Referencias

[1] E. Castiñeira, S. Cubillo, S. Bellido: Contradicción entre dos conjuntos. En: *Actas ESTYLF'02*, León (España), pp. 379–383, 2002.

[2] E. Castiñeira, S. Cubillo: Measures of self-contradiction in Atanassov's fuzzy sets: An Axiomatic Model. *International Journal of Intelligent Systems* 24, pp. 863–888, 2009.

[3] E. Castiñeira, S. Cubillo, C. Torres: Searching degrees of self-contradiction in Atanassov's fuzzy sets. *Mathware and Soft-Computing* 13(3), pp. 136–156, 2006.

[4] E. Castiñeira, C. Torres, S. Cubillo: Un modelo para medir la N-contradicción de conjuntos borrosos. En: *Actas ESTYLF 2008*, Mieres (España), pp. 113–119, 2008.

[5] S. Cubillo, E. Castiñeira: Contradiction in intuitionistic fuzzy sets. En: *Proceedings IPMU*, Perugia (Italia), pp. 2180–2186, 2004.

[6] S. Cubillo, C. Torres, E. Castiñeira: Self-contradiction and contradiction between two intuitionistic fuzzy sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16(3), pp. 283–299, 2008.

[7] F. Esteva, X. Domingo: Sobre funciones de negación en [0,1]. *Stochastica* 4(2), pp. 141–165, 1980.

[8] J. Harding, C. Walker, E. Walker: Lattices of convex normal functions. *Fuzzy Sets and Systems* 159, pp. 1061–1071, 2008.

[9] J. Mendel, R. Jhon: Type-2 fuzzy sets made Simple. *IEEE Transactions on Fuzzy Systems* 10(2), pp. 117–127, 2002.

[10] M. Mizumoto, K. Tanaka: Fuzzy sets of type-2 under algebraic product and algebraic sum. *System* 5, pp. 277–290, 1981

[11] M. Mizumoto, K. Tanaka: Some properties of fuzzy sets of type-2. *Inform. Control* 31, pp. 312–340, 1976.

[12] C. Torres-Blanc, S. Cubillo, E. Castiñeira: A axiomatic model for measuring contradiction and N-contradiction between two AIFSs. *Information Sciences* 180, pp. 834–845, 2010.

[13] E. Trillas, C. Alsina, J. Jacas: On contradiction in fuzzy logic. *Soft Computing* 3(4), pp. 197–199, 1999.

[14] E. Trillas, S. Cubillo: On non-contradictory input/output couples in Zadeh's CRI. En: *Proceedings NAFIPS*, New York (USA), pp.28–32, 1999.

[15] E. Trillas: Sobre funciones de negación en la teoría de conjuntos difusos. *Stochastica* 3(1), pp. 47–60, 1979.

[16] C. Walker, E. Walker: The algebra of fuzzy truth values. *Fuzzy Sets and Systems* 149, pp. 309–347, 2005.

[17] L. Zadeh: The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences* 8, pp. 199–249, 1975.

FUZZY MORPHOLOGICAL OPERATORS IN A GENERAL CONTEXT

J. Elorza[1] **R. Fuentes-González**[2] **J. Bragard**[1] **P. Burillo**[2]

[1] Dept. of Physics and Applied Mathematics. University of Navarra, {jelorza,jbragard}@unav.es
[2] Dept. of Automation and Computer Science. Public University of Navarra, {rfuentes,pburillo}@unavarra.es

Abstract

Paer we show a comprehensive summary on the relationship between fuzzy morphological operators, fuzzy consequence operators and fuzzy preorders and we motivate the study of morphological operators by using binary relations as structuring elements in a more general context, i.e., the context of fuzzy relational systems for obtaining relevant information.

Keywords: Fuzzy Preorder, Fuzzy Consequence Operator, Fuzzy Morphological Operators, Approximate Reasoning, Image Processing.

1 INTRODUCTION

Mathematical morphology was born in the sixties from the collaborative work of G. Matheron and J. Serra ([31, 36]). Some years later, the algebraic basis of mathematical morphology was introduced by H. Heijmans and C. Ronse ([29, 35]). The most important operators of morphological image processing were then analyzed and formalized (see for instance [27, 28, 38]).

Extending mathematical morphology to fuzzy sets was proposed in the early nineties, by several teams independently ([1, 8, 15, 16, 37]), and was then largely developed (see for instance [5, 9, 13, 14, 17, 30, 32]).

Generalizations of fuzzy mathematical morphology from different perspectives and interesting applications have been recently proposed (see for instance [6, 7, 13, 39, 41]).

On another topic, consequence operators were introduced in classical logic by Alfred Tarski in 1930 ([40]). Tarski axiomatized the logic consequence concept trough consequence operators. Quoting Tarski, "a logic is just a set of propositions with a consequence operator".

The concept of fuzzy consequence operator was introduced in 1979 by J. Pavelka. He extend the concept of consequence operator defined by Tarski to fuzzy sets ([33]). Fuzzy consequence operators have been largely studied in the approximate reasoning context with different fuzzy logics (see for instance [10, 11, 12, 18, 19, 22, 25, 26, 34]).

Moreover, fuzzy consequence operators are closely related to the fuzzy preorders through the concept of coherence of an operator. Coherence concept was introduced in the early nineties by J.L. Castro and E. Trillas ([12]).

In short, fuzzy closing morphological operators and fuzzy consequence operators belong to the field of fuzzy mathematical morphology and approximate reasoning, respectively. Both fields require the algebraic framework of complete lattices where the above mentioned operators can be seen as closure operators ([2, 3, 4, 42]).

In [20, 21] we establish the interesting link between fuzzy mathematical morphology and approximate reasoning through these operators and then to connect both operators with the fuzzy preorders.

In Section 2, we recall some definitions, notations and results which will be used in subsequent sections.

In Section 3, we show a summary on the relationship between fuzzy morphological operators, fuzzy consequence operators and fuzzy preorders and we motivate the study of morphological operators by using binary relations as structuring elements in a general context.

Finally, we will provide a good introductory list of references for the interested reader.

2 PRELIMINARIES AND BASIC DEFINITIONS

From now on, X will denote a nonempty (classical) universal set.

The extension of preorder concept to fuzzy $*$-preorder for any t-norm $*$ was done immediately after fuzzy set theory was proposed by L.A. Zadeh in the sixties ([43]).

Definition 1. A fuzzy (binary) relation R on X (fuzzy subset of $X \times X$) is called a $*$-fuzzy preorder if it satisfies:

(FR1) $R(x,x) = 1 \ \forall \ x \in X$ (reflexivity)

(FR2) $R(x,z) \geq R(x,y) * R(y,z) \ \forall \ x,y,z \in X$ ($*$-transitivity).

If R is also symmetric, i.e., $R(x,y) = R(y,x)$ for all $x, y \in X$, then it is called a fuzzy $*$-similarity, fuzzy $*$-indistinguishability or, equivalently, fuzzy $*$-equivalence. If R is only reflexive and symmetric, it is called a fuzzy tolerance.

L will denote a complete lattice with card $L > 1$ which will be the range of the memberships of the fuzzy subsets of X. In many applications, L is frequently taken as the chain $[0,1]$. As usual, L^X will represent the set of all functions from X into L, this is, the set of all L-fuzzy subsets of X.

Moreover, $*$ will be a left-continuous t-norm, that is, $\alpha * (\sup M) = \sup(\alpha * M)$ for all $\alpha \in L$, $M \subset L$ and $\rightsquigarrow_*$ will denote the residuated implication induced by the t-norm $*$: $\alpha \rightsquigarrow_* \beta = \sup\{w \in L / \alpha * w \leq \beta\}$. It is also called the pseudoinverse application of $*$.

In 1979, J. Pavelka introduced the concept of fuzzy consequence operator ([33]) extending the concept of consequence operator in the sense of Tarski as a natural generalization to fuzzy sets.

Definition 2. An operator $C : L^X \longrightarrow L^X$ is a fuzzy consequence operator on X if it verifies:
(FC1) $\mu \subset C(\mu)$ for all fuzzy subset $\mu \in L^X$ (inclusion)
(FC2) Si $\mu_1 \subset \mu_2 \Longrightarrow C(\mu_1) \subset C(\mu_2)$ for all $\mu_1, \mu_2 \in L^X$ (monotony)
(FC3) $C(C(\mu)) = C(\mu)$ for all $\mu \in L^X$ (idempotence)

Notice that, under the above three axioms, (FC3) may be so written equivalently as (FC3') $C(C(\mu)) \subset C(\mu) \ \forall \ \mu \in L^X$ that is a slightly weather form of (FC3).

In 1991, J.L. Castro and E. Trillas proved ([12]) the following result: If R is a fuzzy $*$-preorder then the operator C_R^* between fuzzy subsets of X given by the max-$*$ Zadeh's compositional rule

$$C_R^*(\mu) = \mu \circ^* R \tag{1}$$

is a fuzzy consequence operator (induced by R), where

$$(\mu \circ^* R)(x) = \sup_{w \in X} \{\mu(w) * R(w,x)\} \tag{2}$$

In 1999, J. Elorza and P. Burillo proved ([18]) that the converse of the previous result is also true: R is a $*$-preorder if and only if C_R^* is a fuzzy consequence operator. Therefore the only possibility to obtain fuzzy consequence operators from relations using Zadeh's max-$*$ compositional rule is to work with $*$-preorders.

For the inverse process (inducing fuzzy preorders from fuzzy consequence operators), J.L. Castro and E. Trillas added the coherence axiom to the fuzzy consequence operator concept ([12]).

Definition 3. An operator C between fuzzy sets is called $*$-coherent if

$$\mu(a) * C(\tilde{a}_1)(x) \leq C(\mu)(x) \tag{3}$$

for all $\mu \in L^X$ and for all $(a,x) \in X \times X$, where $\tilde{a}_1(t)$ is the crisp membership function of the singleton $\{a\}$, that is, $\tilde{a}_1(t) = \begin{cases} 1 & \text{if } t = a \\ 0 & \text{if } t \neq a \end{cases}$.

It is proven in [12] that if C is a $*$-coherent fuzzy consequence operator then the relation defined by

$$R_C(x,y) = C(\tilde{x}_1)(y) \tag{4}$$

is a fuzzy $*$-preorder (induced by C). It is also proven that the operator C_R^* is always $*$-coherent.

3 FUZZY MORPHOLOGICAL OPERATORS IN THE CONTEXT OF FUZZY RELATIONAL SYSTEMS

The most used morphological operators have been proposed in a new context: obtaining relevant information in fuzzy relational systems ([23, 24]). Let us give a motivation for using binary relations as structuring elements in this context.

In mathematical morphology, the effect of the basic operators dilation and erosion on an image A is produced when this image A interacts with certain family of translations of another fixed image B, called structuring element. For instance, for binary images $A \subset X$, where $X = \mathbb{R}^2$ or $X = \mathbb{Z}^2$, after assuming that $B \subset X$ is the structuring element, the dilation and the erosion of the set A by B can be expressed respectively by

$$\delta_B(A) = \{x \in X \mid B_x \cap A \neq \emptyset\} \quad (5)$$

and

$$\varepsilon_B(A) = \left\{y \in X \mid \check{B}_y \subset A\right\} \quad (6)$$

where $B_x = \{x + b \mid b \in B\}$ and $\check{B}_y = (-B)_y = \{y - b \mid b \in B\}$.

As it can be seen in the literature, morphological fuzzy dilation and erosion use a unary structuring element together with a translation operation (see for instance equations (7) in definition 2, page 2004 in [7]). Both operators dilation and erosion are related to the set of translations in X and therefore, they are also related to the binary operation $+$ in $\mathbb{R}^2$ or $\mathbb{Z}^2$.

At this point one can ask if it is possible to translate definitions set in equation (5) and (6) such that neither translations nor addition operation appear explicitly in the new definitions. This new formulation would be very useful for universal sets that do not possess a translations group and inner binary operations.

First let us propose new expressions of structuring elements and secondly we generalize them to universal sets X that are not equal to $\mathbb{R}^2$ or $\mathbb{Z}^2$.

For any structuring element $B \subset X$, we define the binary relation $R_B \subset X \times X$ as follows: $xR_By \iff y \in B_x$. From this relation, equations (5) and (6) can be translated by (7) and (8):

$$\delta_{R_B}(A) = \{x \in X \mid \exists y \in X : (y \in A) \wedge (xR_By)\} \quad (7)$$

$$\varepsilon_{R_B}(A) = \{z \in X \mid \forall y \in X : (yR_Bz) \implies (y \in A)\} \quad (8)$$

In a more general context that includes the fuzzy case, these equations can be rewritten as follows: given $R \in L^{X\times X}$ a fuzzy relation (structuring relation), dilation and erosion of $A \in L^X$ by R are defined by

$$\delta_R(A)(x) = (R \circ A)(x) = \sup_{y\in X} \{R(x, y) * A(y)\} \quad (9)$$

$$\varepsilon_R(A)(x) = (R^{op} \triangleleft A)(x) = \inf_{y\in X} \{R(y, x) \rightsquigarrow A(y)\} \quad (10)$$

for some t-norm $*$ and implication operator $\rightsquigarrow$. R^{op} denotes the inverse relation of R, $R^{op}(x, y) = R(y, x)$.

From a structuring fuzzy relation R, the most commonly used morphological operators (dilation, erosion, opening and closing) are then given respectively by $\delta_R(A) = R \circ A$, $\varepsilon_R(A) = R^{op} \triangleleft A$, $\alpha_R(A) = R \circ (R^{op} \triangleleft A)$ and $\varphi_R(A) = R \triangleleft (R^{op} \circ A)$, where the operators $\circ$, $\triangleleft$ are defined as usual:

$(R \circ \mu)(x) = \sup_{w\in X} \left\{R(x, w) * \mu(w)\right\}$, $(R \triangleleft \mu)(x) = \inf_{w\in X} \left\{R(x, w) \rightsquigarrow \mu(w)\right\}$ for some t-norm $*$ and for some implication operator $\rightsquigarrow$.

Both in the crisp and in the fuzzy cases, opening and closing operators with an isotropic structuring element are used in image processing to eliminate specific image details smaller than the structuring element. The global shape of the objects is not distorted.

For instance, closing operator connects objects that are close to each other, fills up small holes and smooths the object outline by filling up narrow gulfs. Meanings of *near*, *small* and *narrow* are related to the size and the shape of the structuring element.

Figure 1 and Figure 2 ([38]) show an example where a set is transformed by a opening and a closing operator through a disk as structuring element, respectively.

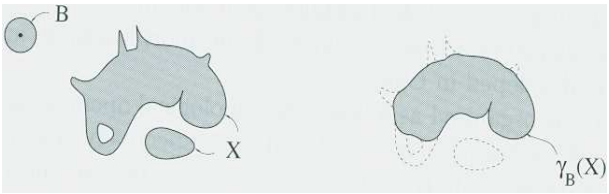


Figure 1: Effect of a opening operator on a non-smooth object

In [20] and [21] the concept of fuzzy closing operator φ^*_R induced by a fuzzy relation is given by

$$\varphi^*_R(\mu) = R^{op} \triangleleft_{\rightsquigarrow_*} (R \circ^* A) \quad (11)$$

where

$$(R \circ^* \mu)(x) = \sup_{w\in X} \left\{R(x, w) * \mu(w)\right\} \quad (12)$$

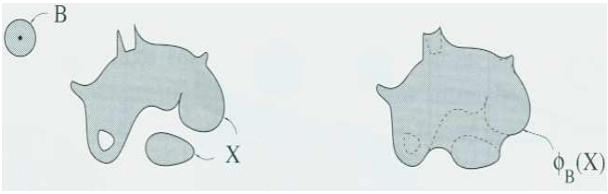


Figure 2: Effect of a closing operator on a non-smooth object

and

$$(R \triangleleft_{\rightsquigarrow_*} \mu)(x) = \inf_{w \in X} \{R(x,w) \rightsquigarrow_* \mu(w)\} \qquad (13)$$

It is well-known that φ^*_R is a closure operator and therefore a fuzzy consequence operator. In [20], we proved that given any fuzzy relation R used as the structuring element, the fuzzy closing operator $C = \varphi^*_R$ induced by such relation is always a coherent fuzzy consequence operator. Therefore such operator C induces always a preorder R_C by means of equation (4).

We have also obtained the following representation of the induced relation:

$$R_C = (R^{op} \triangleleft_{\rightsquigarrow_*} R)^{op}$$

These results have been the key for obtaining new properties.

In particular, we proved some characterizations of the reflexivity and transitivity of a fuzzy relation R. In consequence, we have characterized when a fuzzy relation is an $*$- preorder: R is an $*$-preorder if and only if $R = R^{op} \triangleleft_{\rightsquigarrow_*} R$. It should be noticed that the characterization about the transitivity does not involve the $\circ$ operator because we have characterized the transitivity of a relation R using the operator $\triangleleft$ instead of $\circ$.

In addition, we showed some properties of the preorder R_C induced by the fuzzy closing operator φ^*_R. In particular, we showed that R is a fuzzy $*$-equivalence if and only if $R_C = R$.

Finally, we showed that if R is a preorder then the fuzzy closing morphological operator induced by R by means of (11) is also induced by R as fuzzy consequence operator by means of (1). In fact, we have established that

$$\varphi^*_R = C^*_{R^{op}}$$

From this, it follows that for equivalence relations, the induced fuzzy closing morphological operator coincides with induced fuzzy consequence operator, this is,

$$\varphi^*_R = C^*_R$$

In future works we will use the connection established between both classes of operators and the class of preorders to translate well known properties from the field of approximate reasoning to the one of fuzzy relational systems. In particular we will explore its applications to image processing.

Acknowledgements

This research is partially supported by MICINN (Spain) under projects FIS2008-06335-C02-02 and FIS2011-28820-C02-02.

We thank both anonymous referees for their valuable suggestions and comments.

References

[1] H. Bandemer, W. Näther, Fuzzy Data Analysis. Theoria and Decision Library , Series B: Mathematical and Statistical Methods, Kluwer Academic Publishers, Dordrecht, 1992. 2001.

[2] R. Bělohlávek. Fuzzy closure operators. *Journal of Mathematical Analysis and Appl.* Vol 262, pp. 473-489, 2001.

[3] R. Bělohlávek. Fuzzy closure operators II. *Soft Computing* Vol 7. No 1, pp, 53-64, 2002.

[4] R. Bělohlávek. Fuzzy relational systems. Foundations and principles (IFSR International Series on Systems Science and Engineering, Vol 20, Kluwer Academic/Plenum Plubishers, New York, 2002)

[5] I. Bloch. Duality vs adjunction for fuzzy mathematical morphology and general form of fuzzy erosions and dilations. *Fuzzy sets and systems* Vol 160. pp, 1858-1867, 2009.

[6] I. Bloch. Fuzzy and Bipolar Mathematical Morphology, Applications in Spatial Reasoning. *Lecture Notes in Computer Science* Vol 5590, pp. 1-13. Springer. 2009

[7] I. Bloch. Lattices of fuzzy sets and bipolar fuzzy sets, and mathematical morphology. *Information Sciences* Vol 181, pp, 2002-2015, 2011.

[8] I. Bloch. Triangular Norms as a Tool for Constructing Fuzzy Mathematical Morphologies. *Proc. of the Internat. Workshop on Mathematical Morphology and its Applications to Signal Processing*, pp, 157-161. Barcelona, Spain. 1993.

[9] I. Bloch, H. Maître. Fuzzy Mathematical Morphologies: a comparative study. *Pattern Recognition* Vol 28. No 9, pp, 1341-1387, 1995.

[10] J.L. Castro. Fuzzy logics as families of bivaluated logics. *Fuzzy sets and systems* 64, pp. 321-332, 1994.

[11] J.L. Castro, M. Delgado, E. Trillas. Inducing Implication Relations. *International Journal of Approximate Reasoning*. Vol 10, pp. 235-250, 1994.

[12] J.L. Castro, E. Trillas. Tarski's fuzzy consequences. *Proc. of the Internat. Fuzzy Engineering Symp'91*. Vol.1, pp. 70-81, 1991.

[13] B. De Baets. Generalized idempotence in fuzzy mathematical morphology, in Fuzzy Techniques in Image Processing (E. Kerre and M. Nachtegael, eds.), Studies in Fuzziness and Soft Computing *Physica-Verlag* Vol 52, pp, 58-75, 2000.

[14] B. De Baets. Idempotent closing and opening operations in fuzzy mathematical morphology. *Proc. of the Joint Third Internat. Symposium on Uncertainty Modeling and Analysis and Annual Conference of the North American Fuzzy Information Processing Society*. (College Park, MD, USA), IEEE Computer Society Press, Los Alamitos, pp. 228-233, 1995.

[15] B. De Baets, E. Kerre, M. Gupta. The fundamentals of fuzzy mathematical morphology. Part 1: Basic concepts. *International Journal of General Systems*. Vol 23, pp. 155-171, 1994 pp. 235-250, 1994.

[16] B. De Baets, E. Kerre, M. Gupta. The fundamentals of fuzzy mathematical morphology. Part 2: Idempotence,convexity and decomposition. *International Journal of General Systems*. Vol 23, pp. 307-322, 1995 pp. 235-250, 1994.

[17] T.Q. Deng, H.J.A.M. Heijmans. Grey-scale morphology based on fuzzy logic, Journal of mathematical Imaging and Vision 16 (2002) 155-171.

[18] J. Elorza, P. Burillo. On the relation between fuzzy preorders and fuzzy consequence operators. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. Vol.7. No 3, pp. 219-234, 1999.

[19] J. Elorza, P. Burillo. Connecting fuzzy preorders, fuzzy consequence operators and fuzzy closure and co-closure systems. *Fuzzy sets and systems*. Vol.139. No 3, pp. 601-613, 2003.

[20] J. Elorza, R. Fuentes-González, J. Bragard, P. Burillo. Fuzzy Closing Operators and their Coherence as Fuzzy Consequence Operators *Proc. of the 14th Spanish Conference on Fuzzy Logic and Technologies (ESTYLF 2008)*, pp. 221-227, 2008.

[21] J. Elorza, R. Fuentes-González, J. Bragard, P. Burillo. On the relationship between Fuzzy preorders, Fuzzy Closing Morphological Operators and Fuzzy Consequence Operators *Proc. of the 15th Spanish Conference on Fuzzy Logic and Technologies (ESTYLF 2010)*, pp. 133-138, 2010.

[22] F. Esteva, P. García, L. Godo, R.O. Rodríguez. Fuzzy Approximation Relations, Modal Structures and Possibilistic Logic. *Mathware and Soft Computing*, Vol. V, n 2-3, pp. 151-166, 1998.

[23] R. Fuentes-González. Morfología matemática sobre sistemas relacionales borrosos. Una interpretación de conceptos básicos de Lógica Borrosa mediante erosiones y dilataciones morfológicas *Congreso español de Informática (CEDI 2010)*, 2010. (in Spanish)

[24] R. Fuentes-González, P. Burillo, N. Frago. Técnicas de Morfología Matemática en el tratamiento de Sistemas Relacionales Difusos *Proc. of the 11th Spanish Conference on Fuzzy Logic and Technologies (ESTYLF 2002)*, pp. 321-326, 2002. (in Spanish)

[25] L. Garmendia, J. Recasens. Approximation of Proximities by Aggregating T-indistinguishability Operators. *Mathware and Soft Computing*, Vol. 14, pp. 171-181, 2007.

[26] G. Gerla. Comparing fuzzy and crisp deduction systems, Fuzzy Sets and Systems 67, pp. 317-328. 1994

[27] J. Goutsias, H.J.A.M. Heijmans. Fundamenta morphologicae mathematicae. *Fundamenta Informaticae*, Vol. 41, pp. 1-31, 2000.

[28] H.J.A.M. Heijmans. Morphological Image Operators. *Academic Press*. 1994.

[29] H.J.A.M. Heijmans, C. Ronse. The Algebraic Basis of Mathematical Morphology. Part I: Dilations and erosions. *Computer Vision, Graphics and Image Processing* Vol 50. No 3, pp, 245-295, 1990.

[30] P. Maragos. Lattice image processing: a unification of morphological and fuzzy algebraic systems *Journal of Mathematical Imaging and Vision* Vol 22, pp, 333-353, 2005.

[31] G. Matheron. Random Sets and Integral Geometry. *Wiley*. NY. 1975

[32] M. Nachtegael, P. Sussner, T. Melange, E. Kerre. On the role of complete lattices in mathematical morphology: From tool to uncertainty model. *Information Sciences* Vol 181, pp, 1971-1988, 2011.

[33] J. Pavelka. On Fuzzy Logic I. *Zeitschr. f. Math. Logik und Grundlagen d. Math.* Bd.25, pp. 45-52, 1979.

[34] R.O. Rodríguez, F. Esteva, P. García, L. Godo. On Implicative Closure Operators in Approximate Reasoning, *Internat. Journal of Approximate Reasoning*. Vol.33, pp. 159-184, 2003.

[35] C. Ronse, H.J.A.M. Heijmans. The Algebraic Basis of Mathematical Morphology. Part II: Openings and Closings. *Computer Vision, Graphics and Image Processing* Vol 54. No 1, pp, 74-97, 1991.

[36] J. Serra. Image Analysis and Mathematical Morphology. Academic Press. Vol I (fourth printing, 1990) and Vol II (second printing, 1992).

[37] D. Sinha, E.R. Dougherty. Fuzzification of set inclusion: theory and applications. *Fuzzy sets and systems*. Vol.55, pp. 15-42, 1993.

[38] P. Soille. Morphological Image Analysis. Principles and Applications. Springer. 1999.

[39] P. Sussner, M.E. Valle. Classification of fuzzy mathematical morphologies based on concepts of inclusion measure and duality. *Journal of Mathematical Imaging and Vision* Vol 32. No 2, pp, 139-159, 2008.

[40] A. Tarski. Logic, semantics, metamathematics. Clarendon Press, Oxford, 1956.

[41] M.E. Valle, P. Sussner. A general framework for fuzzy morphological associative memories. *Fuzzy sets and systems*. Vol.159. No 7, pp. 747-768, 2008.

[42] M. Ward. The closure operators of a lattice. *Annals of Mathematics*. Vol.43. N.2, pp. 191-196, 1940.

[43] L.A. Zadeh, Fuzzy sets, *Information and Control* 8, pp. 338-353, 1965.

DECOMPOSITION OF $\mathscr{IV}$-SIMILARITIES

Ramón González-del-Campo[1], **Luis Garmendia**[2]

[1]*DSIC(UCM)*, rgonzale@estad.ucm.es
[2]*DISIA(UCM)*, lgarmend@fdi.ucm.es

Abstract

In this paper an extension of Lee's theorem for interval-valued fuzzy relations is given. An algorithm to compute an $\mathscr{IV}$-similarity decomposition is given.

Keywords: interval-valued fuzzy relations, $\mathscr{IV}$-similarity, bridge between two $\mathscr{IV}$-similarities.

1 Introduction

Interval-valued fuzzy sets ($\mathscr{IVFS}$s) were introduced in the 60s by Grattan-Guinness [12], Jahn [13], Sambuc [18] and Zadeh [21]. They are extensions of classical fuzzy sets where the membership value between 0 and 1 is replaced by an interval in $[0,1]\times[0,1]$. They easily allow to model uncertainty and vagueness: sometimes it is better to give a membership interval than a membership degree to describe a characteristic of objects on a universe. $\mathscr{IVFS}$ are a special case of type-2 fuzzy sets that simplifies the calculations while preserving their richness as well. $\mathscr{IVFS}$ have been widely applied [1, 2, 3, 5, 6, 8, 9, 11, 15, 16, 17, 19, 20].

Lee's theorem [14] was proposed for constructing the max-min equivalence closure of fuzzy similarity matrices.

This paper is organized as follows. In section 2 some basic concepts about interval-valued fuzzy sets are recalled. Lee's theorem for fuzzy relations is recalled as well. In section 3 it is showed that the bridge between two $\mathscr{IV}$-similarities is a $\mathscr{IV}$-similarity. Moreover, for any $\mathscr{IV}$-similarity there exists at least a decomposition into two $\mathscr{IV}$-subsimilarities. In subsection 3.1 an algorithm to compute the decomposition is given.

2 Preliminaries

Definition 2.1. *[6] Let $(L,\leq_L)$ be a lattice such that:*

1. $L=\{[x_1,x_2]\in[0,1]^2$ *with* $x_1\leq x_2\}$.
2. $[x_1,x_2]\leq_L[y_1,y_2]$ *if and only if* $x_1\leq y_1$ *and* $x_2\leq y_2$. *In particular:*
   $[x_1,x_2]=_L[y_1,y_2]$ *if and only if* $x_1=y_1$ *and* $x_2=y_2$.

$(L,\leq_L)$ is a complete lattice and the supremum and infimum are defined as follows.

Definition 2.2. *[7] Let $\{[v_i,w_i]\}$ be a set of intervals on L. Then*

1. $\wedge\{[v_i,w_i]\}\equiv[infimun\{v_i\},infimun\{w_i\}]$
2. $\vee\{[v_i,w_i]\}\equiv[supremun\{v_i\},supremun\{w_i\}]$

Definition 2.3. *[6] An interval-valued fuzzy set A on a universe X can be represented by the mapping:*

$$A=\{(a,[x_1,x_2])\mid a\in X,[x_1,x_2]\in L\}$$

Definition 2.4. *[6] Let A and B be two interval-valued fuzzy sets on a universe X. The equality between A and B is defined as follows:*

$$A=B \text{ if and only if } A(a)=_L B(a) \text{ for all } a\in X.$$

Definition 2.5. *[6] Let A and B be two interval-valued fuzzy sets on a universe X. The inclusion of A into B is defined as follows:*

$$A\subseteq B \text{ if and only if } A(a)\leq_L B(a) \text{ for all } a\in X.$$

Definition 2.6. *[6] Let A and B be two interval-valued fuzzy sets on a universe X. The intersection between A and B is defined as follows:*

$$A\cap B(a)=_L\wedge(A(a),B(a)) \text{ for all } a\in X.$$

Definition 2.7. *[6] A negation operator for interval-valued fuzzy sets $\mathscr{N}$ is a decreasing function $\mathscr{N}:L\to L$ that satisfies:*

1. $\mathcal{N}(0_L) =_L 1_L$
2. $\mathcal{N}(1_L) =_L 0_L$

where $0_L = [0,0]$ *and* $1_L = [1,1]$.

If $\mathcal{N}(\mathcal{N}([x_1,x_2])) =_L [x_1,x_2]$ *for all* $[x_1,x_2]$ *in L then* $\mathcal{N}$ *is called an involutive negation.*

Definition 2.8. *[6] A strong negation operator for interval-valued fuzzy sets* $\mathcal{N}$ *is a continuous, strictly decreasing and involutive negation operator.*

The t-norm operators are generalized to the lattice L in a straightforward way.

Definition 2.9. *[6] A generalized t-norm function* $\mathcal{T}$ *is a monotone increasing, symmetric and associative operator* $\mathcal{T} : L^2 \to L$ *that satisfies:* $\mathcal{T}(1_L,[x_1,x_2]) =_L [x_1,x_2]$ *for all* $[x_1,x_2]$ *in L.*

Notation. The lowest value of an interval $x = [x_1,x_2]$ is denoted $\underline{x}$ and the highest value of an interval $x = [x_1,x_2]$ is denoted $\overline{x}$, so $x = [x_1,x_2] = [\underline{x},\overline{x}]$.

Definition 2.10. *[6] A generalized t-norm operator* $\mathcal{T} : L^2 \to L$ *is t-representable if there are two t-norms* T_1 *and* T_2 *that satisfy:*

1. $\mathcal{T}([x_1,x_2],[y_1,y_2]) =_L [T_1(x_1,y_1),T_2(x_2,y_2)]$
2. $T_1(v,w) \leq T_2(v,w)$ *for all* $v,w \in [0,1]$

Definition 2.11. *[4] Let* X_1 *and* X_2 *be two universes of discourse. An interval-valued fuzzy relation* $R : X_1 \times X_2 \to L$ *is a mapping:*

$$R = \{((a,b),[x,y]) \mid a \in X_1, b \in X_2, [x,y] \in L\}$$

where $x = \underline{R(a,b)}$ and $y = \overline{R(a,b)}$.

In the rest of the paper, it is assumed that $X_1 = X_2 = X$ and that X is finite.

Definition 2.12. *[10] An interval-valued fuzzy relation* $R : X^2 \to [0,1]^2$ *is reflexive if:*

$$R(a,a) =_L 1_L \text{ for all } a \in X$$

where $1_L = [1,1]$.

Definition 2.13. *[10] An interval-valued fuzzy relation* $R : X^2 \to [0,1]^2$ *is* $[\alpha_1,\alpha_2]-$*reflexive if:*

$$R(a,a) \geq_L [\alpha_1,\alpha_2] \text{ for all } a \in X$$

where $\alpha_1,\alpha_2 \in [0,1]$ *and* $\alpha_1 \leq \alpha_2$.

Definition 2.14. *[10] An interval-valued fuzzy relation* $R : X^2 \to [0,1]^2$ *is symmetric if:*

$$R(a,b) =_L R(b,a) \text{ for all } a,b \in X$$

Definition 2.15. *[10] Let* $\mathcal{T}$ *be a generalized t-norm operator and let R be an interval-valued fuzzy relation on X. R is* $\mathcal{T}$*-transitive if:*

$$\mathcal{T}(R(a,b),R(b,c)) \leq_L R(a,c) \text{ for all } a,b,c \in X$$

Definition 2.16. *A fuzzy relation* $R : X^2 \to [0,1]$ *is a similarity if R is reflexive, symmetric and* min*-transitive.*

Definition 2.17. *Let B be a similarity on X with cardinality n. Let* $\pi = \{a_k\}$ *be a permutation of the values 1..n. Then* $P_\pi(B)$ *is the similarity where the* k^{th} *element of X is permuted by the element* a_k *of X.*

Lemma 2.1. *[14] Let B be a similarity. Let* π *be any permutation. Then* $P_\pi(B)$ *is a similarity.*

Definition 2.18. *[14] Let R and S be two similarities on universes* X_1 *and* X_2 *with cardinalities* n_1 *and* n_2 *respectively. Let t be an interval in* $[0,1]$ *such that* $t \leq \wedge(\wedge(R(a_i,a_j)),\wedge(S(a_k,a_l)))$ *for all i,j,k,l. Then, the t-bridge between R and S is a similarity defined as follows:*

$$B_{t;R_{n_1\times n_1},S_{n_2\times n_2}} = \begin{pmatrix} R & (t)_{n_1\times n_2} \\ (t)_{n_2\times n_1} & S \end{pmatrix}$$

where $(t)_{n_2\times n_1}$ *and* $(t)_{n_1\times n_2}$ *are a* $n_2 \times n_1$ *matrix and a* $n_1 \times n_2$ *respectively whose elements are t.*

Sometimes it is written $B_{t;R,S}$ instead of $B_{t;R_{n_1\times n_1},S_{n_2\times n_2}}$.

Theorem 2.1. *[14] Let C be a similarity on a universe with cardinality* $n \geq 2$*. Then, there exist two similarities (R, S) and a value t which verifies:*

$$C = P_\pi(B_{t;R_{n_1\times n_1},S_{n_2\times n_2}})$$

for some permutation π.

Note that n_1 *and* n_2 *are the cardinalities of the universes of R and S respectively with* $n = n_1 + n_2$ *and that t is the minimum value of R and S.*

3 Decomposition of $\mathcal{IV}$-similarities into $\mathcal{IV}$-subsimilarities

Let $X = \{a_1,...,a_n\}$ be a finite universe.

Definition 3.1. *Let* $\mathcal{M}([x_1,x_2],[y_1,y_2]) = [min(x_1,y_1),min(x_2,y_2)]$ *be a t-generalized t-norm. Let R be an interval-valued fuzzy relation on a universe X. R is a* $\mathcal{IV}$*-similarity if it is reflexive, symmetric and* $\mathcal{M}$*-transitive.*

Definition 3.2. *Let R and S be two* $\mathcal{IV}$*-similarities on universes* X_1 *and* X_2 *with cardinalities* n_1 *and* n_2 *respectively. Let* $[t_1,t_2]$ *be an interval in* $[0,1]^2$ *such that* $[t_1,t_2] \leq_L \wedge(\wedge(R(a_i,a_j)),\wedge(S(a_k,a_l)))$ *for all i,j,k,l. Then the* $[t_1,t_2]$*-bridge between R and S is an interval-valued fuzzy relation on* $X_1 \cup X_2$*. B is defined as follows:*

$$B_{[t_1,t_2];R,S} = \begin{pmatrix} R & ([t_1,t_2])_{n_1 \times n_2} \\ ([t_1,t_2])_{n_2 \times n_1} & S \end{pmatrix}$$

where $[t_1,t_2]$ is called the bridge interval.

Lemma 3.1. *Let R and S be two reflexive $\mathscr{IVFR}$s on universes X_1 and X_2. Then $B_{[t_1,t_2];R,S}$ is a reflexive interval-valued fuzzy relation on $X_1 \cup X_2$.*

Proof. Trivial because R and S are reflexive interval-valued fuzzy relations. □

Lemma 3.2. *Let R and S be two symmetric $\mathscr{IVFR}$s on universes X_1 and X_2. Then $B_{[t_1,t_2];R,S}$ is a symmetric interval-valued fuzzy relation on $X_1 \cup X_2$.*

Proof. Cases:

1. If $i,j \leq n_1$ then $B_{[t_1,t_2];R,S}(a_i,a_j) = R(a_i,a_j) = R(a_j,a_i) = B_{[t_1,t_2];R,S}(a_j,a_i)$ because R is symmetric.
2. If $i,j > n_1$ then $B_{[t_1,t_2];R,S}(a_i,a_j) = S(a_i,a_j) = S(a_j,a_i) = B_{[t_1,t_2];R,S}(a_j,a_i)$ because S is symmetric.
3. Otherwise, $B_{[t_1,t_2];R,S}(a_i,a_j) = [t_1,t_2] = B_{[t_1,t_2];R,S}(a_j,a_i)$ by construction.

□

Lemma 3.3. *Let R and S be two $(min-min)$-transitive interval-valued fuzzy relations on universes X_1 and X_2 with cardinalities n_1 and n_2 respectively. Let $B_{[t_1,t_2];R,S}$ be the $[t_1,t_2]$-bridge between R and S. Then, $B_{[t_1,t_2];R,S}$ is (min,min)-transitive interval-valued fuzzy relation.*

Proof. It will be proved that $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) \leq_L B_{[t_1,t_2];R,S}(a_i,a_j)$ for all i,j,k.
Cases:

1. For $i,j \leq n_1$:
   (a) For $k \leq n_1$:
   $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) =_L$
   $=_L \mathscr{M}(R(a_i,a_k), R(a_k,a_j))$
   $\leq_L R(a_i,a_j) =_L B_{[t_1,t_2];R,S}(a_i,a_j)$
   so $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) \leq_L B_{[t_1,t_2];R,S}(a_i,a_j)$ is verified due to the fact that R is a $\mathscr{IV}$-similarity.
   (b) $k > n_1$:
   $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) =_L$
   $\mathscr{M}([t_1,t_2],[t_1,t_2]) =_L [t_1,t_2] \leq_L B_{[t_1,t_2];R,S}(a_i,a_j)$.
2. For $i \leq n_1$ and $j > n_1$:
   (a) For $k \leq n_1$:
   $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) =_L$
   $\mathscr{M}(R(a_i,a_k),[t_1,t_2]) =_L [t_1,t_2] \leq_L$
   $B_{[t_1,t_2];R,S}(a_i,a_j)$
   (b) For $k > n_1$:
   $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) =_L$
   $\mathscr{M}([t_1,t_2],S(a_k,a_j)) =_L [t_1,t_2] \leq_L$
   $B_{[t_1,t_2];R,S}(a_i,a_j)$
   so $\mathscr{M}(B_{[t_1,t_2];R,S}(a_i,a_k), B_{[t_1,t_2];R,S}(a_k,a_j)) =_L$
   $[t_1,t_2] \leq_L B_{[t_1,t_2];R,S}(a_i,a_j)$ for all k.
3. For $i,j > n_1$: similar to case 1 using S instead of R.
4. For $i > n_1$ and $j \leq n_1$: similar to case 2 using R instead of S.

□

Theorem 3.1. *Let R and S be two $\mathscr{IV}$-similarities on universes X_1 and X_2. Let $B_{[t_1,t_2];R,S}$ be the $[t_1,t_2]$-bridge between R and S. Then, $B_{[t_1,t_2];R,S}$ is a $\mathscr{IV}$-similarity on $X_1 \cup X_2$.*

Proof. Trivial according to the lemmas 3.1,3.2,3.3. □

Lemma 3.4. *Let R be an interval-valued fuzzy relation on a finite universe X. If R is a $\mathscr{IV}$-similarity then there exists an interval $R(a_p,a_q)$ such that $R(a_p,a_q) \leq_L R(a_i,a_j)$ for all i, j.*

Proof. It will be proved using mathematical induction:

1. Base case: Let X be a universe with cardinality 3. Let R be a $\mathscr{IV}$-similarity on X. Then R is as follows:

   $$\begin{pmatrix} 1_L & [y_1,y_2] & [x_1,x_2] \\ [y_1,y_2] & 1_L & [x_1,x_2] \\ [x_1,x_2] & [x_1,x_2] & 1_L \end{pmatrix}$$

   where $[x_1,x_2] \leq_L [y_1,y_2]$, so there exists an interval $[y_1,y_2]$ which is minimum.
2. Inductive step: Let X be a universe with cardinality n. Let R be a $\mathscr{IV}$-similarity on X which has a minimum interval. Let's add an element a_{n+1} to the universe. Let S be a $\mathscr{IV}$-similarity on $X \cup \{a_{n+1}\}$ such that $S(a_i,a_j) = R(a_i,a_j)$ for all a_i,a_j in X. It is denoted $S(a_i,a_{n+1})$ as c_i for all $i : 1..n$. For any a_i,a_j in X, it is verified:

   $$min(R(a_i,a_j),c_j) \leq_L c_i$$
   $$min(c_i,c_j) \leq_L R(a_i,a_j)$$
   $$min(R(a_i,a_j),c_i) \leq_L c_j$$

   so $c_i \leq_L c_j$ or $c_j \leq_L c_i$. In similar way it is possible to prove that $R(a_i,a_j) \leq_L c_j$ or $c_j \leq_L R(a_i,a_j)$ and $c_i \leq_L R(a_i,a_j)$ or $R(a_i,a_j) \leq_L c_i$.

   Therefore, there exists a minimum interval in S.

□

Definition 3.3. *Let R be a $\mathscr{IV}$-similarity on X. Let $I \subseteq X$ be a subset of X. A subsimilarity of R restricted to T $(R)_I$ is defined as $(R)_I : I \times I \to [0,1]^2$ such that $(R)_I(a_r,a_s) = R(a_r,a_s)$ for all a_r and a_s are in I. $(R)_I$ is denoted I-subsimilarity of R.*

Lemma 3.5. *Let R be a $\mathscr{IV}$-similarity on X. Any I-subsimilarity of R on $I \subseteq X$ is a $\mathscr{IV}$-similarity.*

Proof. Trivial. □

Lemma 3.6. *Let C be a $\mathscr{IV}$-similarity and let $C(a_p,a_q)$ be the minimum value of C. Let $I = \{a_i \mid a_i \in X \text{ and } C(a_i,a_q) =_L C(a_p,a_q)\}$ and $I' = X \setminus I$. Then, $C(a_i,a_j) >_L C(a_p,a_q)$ for all a_i and a_j in I'.*

Proof. Trivial by construction of I'. □

Lemma 3.7. *Let C be a $\mathscr{IV}$-similarity and let $C(a_p,a_q)$ be the minimum value of C. Let $I = \{a_i \mid a_i \in X \text{ and } C(a_i,a_q) =_L C(a_p,a_q)\}$ and $I' = X \setminus I$. Then, $C(a_i,a_j) =_L C(a_p,a_q)$ for all a_i in I and for all a_j in I'.*

Proof. Let's suppose there exists a_r in I and a_s in I' such that $C(a_r,a_s) >_L C(a_p,a_q)$. Let $C(a_r,a'_s)$ be such that r is in I and s' in I'. By lemma 3.6 $C(a_s,a_{s'}) >_L C(a_p,a_q)$ for a_s and a'_s in I'. Then, it is verified $C(a_p,a_q) <_L min(C(a_r,a_s),C(a_s,a_{s'})) \leq_L C(a_r,a_{s'})$ so a_r and a'_s are in I' which contradicts the assumption.

□

Theorem 3.2. *Let C be a $\mathscr{IV}$-similarity on a universe X with cardinality n. Then, there exist two similarities (R, S) and an interval $[t_1,t_2]$ which verifies:*

$$C = P_\pi(B_{[t_1,t_2];R_{n_1\times n_1},S_{n_2\times n_2}})$$

for some permutation π.

Note that n_1 and n_2 are the cardinalities of the universes of R and S respectively and $[t_1,t_2]$ is the minimum value of R and S.

Proof. Let C be a $\mathscr{IV}$-similarity and let $C(a_p,a_q)$ be the minimum value of C according to lemma 3.4.

Let $I = \{a_i \mid a_i \in X \text{ and } C(a_i,a_q) = C(a_p,a_q)\}$ and $I' = X \setminus I$.

$(C)_I$ and $(C)_{I'}$ are $\mathscr{IV}$-similarities according to lemma 3.5

For all a_i in I and for all a_j in I' it is verified $C(a_i,a_j) = C(a_p,a_q)$ according to lemma 3.7.

Let $\pi = \{a_{i_1},...,a_{i_{n_1}},a_{i'_1},...,a_{i'_{n_2}}\}$ be a permutation where $a_{i_1},...,a_{i_{n_1}}$ belong to I and $a_{i'_1},...,a_{i'_{n_2}}$ belong to I'.

Then: $C = P_\pi(B_{[t_1,t_2];R_{n_1\times n_1},S_{n_2\times n_2}})$ where $[t_1,t_2] =_L C(a_p,a_q)$, $n_1 = card(I)$ and $n_2 = card(X \setminus I)$ □

3.1 Algorithm to decompose an $\mathscr{IV}$-similarity into $\mathscr{IV}$-subsimilarities

- *Input*: A $\mathscr{IV}$-similarity C on X with cardinality n such that $n \geq 2$.
- *Output*: Two $\mathscr{IV}$-subsimilarities R,S on X_1, X_2 such that $X = X_1 \cup X_2$ and a brdge interval t such that $C = P_\pi(B_{[t_1,t_2];R,S})$
- The algorithm runs as follows:
  - *Step 1*: Search for the minimum interval of C (it always exists by lema 3.4): $t = [a_p,a_q]$.
  - *Step 2*: Choose X_1 such that $X_1 \subset X$. Compute I such that $I = \{a_i \mid a_i \in X_1 \text{ and } C(a_i,a_q) =_L C(a_p,a_q)\}$
  - *Step 3*: Compute: $R \leftarrow C_I$ and $S \leftarrow C_{X\setminus I}$

Example 3.1. *Let C be a $\mathscr{IV}$-similarity:*

$$C = \begin{pmatrix} [1,1] & [0.1,0.2] & [0.8,1] \\ [0.1,0.2] & [1,1] & [0.1,0.2] \\ [0.8,1] & [0.1,0.2] & [1,1] \end{pmatrix}$$

Then:

$$R = \begin{pmatrix} [1,1] & [0.8,1] \\ [0.8,1] & [1,1] \end{pmatrix}$$

$$S = \begin{pmatrix} [1,1] \end{pmatrix}$$

and

$$t = [0.1,0.2]$$

$$B_{t;R,S} = \begin{pmatrix} [1,1] & [0.8,1] & [0.1,0.2] \\ [0.8,1] & [1,1] & [0.1,0.2] \\ [0.1,0.2] & [0.1,0.2] & [1,1] \end{pmatrix}$$

4 Conclusions

The decomposition of similarities into subsimilarities is a usefull result that can be applied, for example, in optimal computation of min-transitive closures [14]. In this paper those results are extended to decompose and $\mathscr{IV}$-similarity into $\mathscr{IV}$-subsimilarities, including an algorithm to do it.

Acknowledgment

This research is partially supported by the Spanish Ministry of Science and Technology, grant number TIN2009-07901, the Research Group CAM GR35/10-A at Complutense University of Madrid.

References

[1] T. Bilgic. Interval-valued preference structures. *European Journal of Operational Research*, 105(1):162–183, February 1998.

[2] H. Bustince, E. Barrenechea, M. Pagola, and R Orduna. Image thresholding computation using Atanassov's intuitionistic fuzzy sets. *JACIII*, pages 187–194, 2007.

[3] H. Bustince and P. Burillo. Mathematical analysis of interval-valued fuzzy relations: Application to approximate reasoning. *Fuzzy Sets Syst*, 113:205–219, 2000.

[4] H. Bustince and P. Burillo. Mathematical analysis of interval-valued fuzzy relations: application to approximate reasoning. *Fuzzy Sets Syst.*, 113(2):205–219, 2000.

[5] H. Bustince, V. Mohedano, E. Barrenechea, and M. Pagola. An algorithm for calculating the threshold of an image representing uncertainty through a-ifss. *Proceedings of 11th Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 2383–2390, 2006.

[6] C. Cornelis, G. Deschrijver, and E. Kerre. Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application. *Int. J. Approx. Reasoning*, 35(1):55–95, 2004.

[7] C. Cornelis, G. Deschrijver, and E. Kerre. Advances and challenges in interval-valued fuzzy logic. *Fuzzy Sets and Systems*, 157(5):622–627, 2006.

[8] G. Deschrijver and E. E. Kerre. On the composition of intuitionistic fuzzy relations. *Fuzzy Sets Syst.*, 136(3):333–361, 2003.

[9] A. Dziech and M. B. Gorzalczany. Decision making in signal transmission problems with interval-valued fuzzy sets. *Fuzzy Sets Syst.*, 23:191–203, August 1987.

[10] R. González-del Campo, L. Garmendia, and J. Recasens. Transitive closure of interval-valued relations. *Proceedings IFSA-EUSFLAT'09*, pages 837–842, 2009.

[11] M. B. Gorzalczany. Interval-valued fuzzy controller based on verbal model of object. *Fuzzy Sets Syst.*, 28:45–53, October 1988.

[12] I. Grattan-Guiness. Fuzzy membership mapped onto interval and many-valued quantities. *Math. Logik. Grundladen Math*, 22:149–160, 1975.

[13] K.U. Jahn. Intervall-wertige Mengen. *Math. Nach.*, 68:115–132, 1975.

[14] H. Lee. An optimal algorithm for computing the max-min transitive closure of a fuzzy similarity matrix. *Fuzzy Sets and Systems*, 123(1):129–136, 2001.

[15] A. Pankowska and M. Wygralak. On hesitation degrees in if-set theory. In *ICAISC*, pages 338–343, 2004.

[16] A. Pankowska and M. Wygralak. General if-sets with triangular norms and their applications to group decision making. *Inf. Sci.*, 176(18):2713–2754, 2006.

[17] M. K. Roy and R. Biswas. I-v fuzzy relations and sanchez's approach for medical diagnosis. *Fuzzy Sets Syst.*, 47:35–38, April 1992.

[18] E. Sanchez and R. Sambuc. Fuzzy relationships. phi -fuzzy functions. application to diagnostic aid in thyroid pathology. *Proceedings of an International Symposium on Medical Data Processing*, pages 513–524, 1976.

[19] E. Szmidt and J. Kacprzyk. Group decision making under intuitionistic fuzzy preference relations. In *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 172–178. IPMU, 1998.

[20] H.R. Tizhoosh. Image thresholding using type-2 fuzzy sets. *Pattern Recognit*, 38:2363–2372, 2005.

[21] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I. *Information Sciences*, 8:199–249, 1975.

Isomorfismos entre Indistinguibilidades, Conjuntos de Extensionales y Aproximaciones Superiores e Inferiores

Gabriel Mattioli[1], **Jordi Recasens**[1]

[1]*Escola Tècnica Superior d'Arquitectura del Vallès*
Universitat Politècnica de Catalunya
C/Pere Serra 1-15 Sant Cugat del Vallès, Barcelona
{gabriel.mattioli,j.recasens}@upc.edu

Resumen

Dada una relación de T-indistinguibilidad E se puede derivar su correspondiente conjunto de extensionales H_E y sus operadores de aproximación por conjuntos extensionales superior ϕ_E e inferior ψ_E por conjuntos extensionales. Recíprocamente, dado cualquiera de éstos la indistinguibilidad inicial puede ser hallada. Este trabajo prueba que las conocidas biyecciones entre estos conceptos son en realidad isomorfismos de retículos. Se estudia también como son afectados el orden y las operaciones de los retículos bajo el efecto de dichos isomorfismos.

Palabras Clave: Indistinguibilidad, Conjunto de Extensionales, Aproximación Superior, Aproximación Inferior.

1. INTRODUCCIÓN

El conjunto H_E de conjuntos extensionales asociados a una indistinguibilidad E ha sido estudiado desde múltiples ángulos en la literatura del tema. En [4] se estudia desde una perspectiva algebraica y en [6] se analiza su estructura topológica. H_E también ha sido caracterizado como el conjunto de generadores de E [5], como un conjunto de estados lógicos y como el conjunto de puntos fijos de dos operadores conocidos como ϕ_E y ψ_E [6]. Dado un subconjunto difuso μ, estos operadores proporcionan su mejor aproximación superior e inferior respectivamente por conjuntos extensionales.

En [9] se desarrolla un conjunto de propiedades que caracteriza completamente a los operadores de aproximación superior e inferior por extensionales y en [4] se dan propiedades muy similares que caracterizan a los conjuntos de difusos que son el conjunto de extensionales de una cierta indistinguibilidad E. En consecuencia se tiene que el conjunto $\mathscr{E}$ de operadores de indistinguibilidad sobre un universo X, el conjunto $\mathscr{H}$ de conjuntos de extensionales asociados a una indistinguibilidad E sobre X y los conjuntos $\mathscr{U}$ y $\mathscr{L}$ de operadores de aproximación por extensionales superior e inferior respectivamente asociados a una indistinguibilidad E sobre X están en biyección. Sin embargo, dado que estos conceptos están tan fuertemente interrelacionados y que son caracterizados por propiedades bastante similares parece natural preguntarse si la relación entre ellos es más fuerte que una simple biyección conjuntista y alcanza algún nivel estructural.

Este trabajo se ha concentrado en resolver esta pregunta y el resultado obtenido es que estas biyecciones son en realidad isomorfismos ya que el orden y las operaciones elementales son respetadas (en algún caso de forma dual). Más aún, dado que puede darse a estos conjuntos la estructura de retículos, la conclusión del trabajo es que estos retículos son isomorfos.

Ésto es importante tanto desde un punto de vista teórico como práctico. Tiene relevancia teórica porque proporciona una luz unificadora a los muchos estudios realizados alrededor de estos cuatro conceptos de forma independiente y que llegaban a conclusiones similares tras mucho esfuerzo demostrativo. A la vista de estos isomorfismos se hace evidente que la única diferencia entre estos retículos es su aproximación semántica y la notación y por tanto muchos resultados y teoremas pueden ser traducidos de un ámbito a otro de forma directa a través de estos isomorfismos.

Además, en contextos prácticos se da muchas veces la situación de que en un cierto problema múltiples criterios o atributos deben ser considerados y por ello deben ser agregados de alguna manera. En [7], por ejemplo, se estudia la agregación de relaciones de indistinguibilidad. Hasta ahora, en los estudios sobre agregación realizados es necesario que las magnitudes o conceptos a agregar estén codificados en la misma estructura formal. Sin embargo, en muchos conceptos prácticos ésto requiere un importante esfuerzo ya que la misma naturaleza de la medida de los criterios es esencialmente diferente (datos estadísticos, relacionales, cualitativos...) y sería mucho más sencillo y menos arbitrario poder codificar cada atributo en la estructura formal

que mejor se le adecúe. Los isomorfismos desarrollados en este trabajo permiten que cada atributo se trate como mejor convenga y después todos puedan ser traducidos a la misma estructura para realizar la agregación.

Este trabajo está estructurado de la manera siguiente:

En la sección de Preliminares se darán las definiciones y resultados básicos sobre los que se desarrolla el estudio posterior.

En la tercera sección se prueban los principales resultados del trabajos. La sección esta dividida en tres subsecciones que desarrollan separadamente los isomorfismos entre los retículos $\mathscr{E}$, $\mathscr{H}$, $\mathscr{U}$ y $\mathscr{L}$ respectivamente. Los isomorfismos son explícitamente dados y se estudia su acción sobre el orden y las operaciones de los retículos

Por último, las conclusiones del trabajo se dan en la última sección.

2. PRELIMINARES

En este apartado se dan las definiciones y resultados básicos de los conceptos tratados en este trabajo. En primer lugar se da la definición de operador de indistinguibilidad y de algunos conceptos relacionados. Posteriormente se da la definición del conjunto de extensionales H_E y los operadores de aproximación por conjuntos extensionales superior ϕ_E e inferior ψ_E asociados a una T-indistinguibilidad E.

Definición 2.1. *[12] Sea T una t-norma. Una relación difusa E sobre un conjunto X es un operador de T-indistinguibilidad si y solo si $\forall x,y,z \in X$*

a) $E(x,x) = 1$ *(Reflexividad)*

b) $E(x,y) = E(y,x)$ *(Simetría)*

c) $T(E(x,y),E(y,z)) \leq E(x,z)$ *(T-transitividad).*

Es inmediato comprobar que dadas dos indistinguibilidades E y F, su intersección $E \cap F$ también es una indistinguibilidad. Por contra, la unión $E \cup F$ en general no preserva la T-transitividad.

Definición 2.2. *Sea T una t-norma y R una relación reflexiva y simétrica sobre X. La clausura transitiva $\overline{R}$ de R es la T-indistinguishabilidad definida como*

$$\overline{R} = \bigcap_{E \in A} E$$

donde A es el conjunto de T-indistinguibilidades sobre X mayores o iguales a R.

Ahora podemos otorgar una estructura reticular al conjunto de T-indistinguibilidades $\mathscr{E}$ sobre X, con el orden natural $\leq$, y dos operaciones: la intersección $E \cap F$ y la unión cerrada $\overline{E \cup F}$.

Se pueden generar indistinguibilidades de múltiples maneras. La más sencilla es generar la indistinguibilidad asociada a un subconjunto difuso.

Definición 2.3. *Sea X un conjunto, T una t-norma, $\overleftrightarrow{T}$ su operador birresidual y μ un subconjunto difuso de X. La relación difusa E en X definida por*

$$E_\mu(x,y) = \overleftrightarrow{T}(\mu(x),\mu(y))$$

$\forall x,y \in X$ es una T-indistinguibilidad.

Uno de los principales resultados de la teoría de indistinguibilidades es el Teorema de Representación. Éste prueba que cualquier familia de difusos en un conjunto X genera una T-indistinguibilidad y, recíprocamente, que cualquier T-indistinguibilidad sobre X puede ser generada por una familia de difusos. El enunciado es el siguiente:

Teorema 2.4. [13] *Sea R una relación difusa en un conjunto X y T una t-norma continua. Entonces R es una T-indistinguibilidad si y sólo si existe una familia $(\mu_i)_{i\in I}$ de difusos en X tal que $\forall x,y \in X$*

$$R(x,y) = \inf_{i\in I} E_{\mu_i}(x,y).$$

El Teorema de Representación caracteriza las T-indistinguibilidades por su conjunto de generadores. Es bien sabido que éstos coinciden con los conjuntos extensionales que se definen a continuación.

Definición 2.5. *Sea X un conjunto y E una T-indistinguibilidad sobre X. Un subconjunto difuso μ de X es extensional respecto a E si y sólo si:*

$$\forall x,y \in X \;\; T(E(x,y),\mu(y)) \leq \mu(x).$$

Denotaremos H_E al conjunto de todos los extensionales respecto a E en X.

Proposición 2.6. *Sea X un conjunto, E una T-indistinguibilidad en X y μ un subconjunto difuso de X. Entonces:*

$$\mu \in H_E \Leftrightarrow E_\mu \geq E$$

En [4] se prueba que los conjuntos de extensionales quedan caracterizados por las propiedades siguientes.

Proposición 2.7. *Sea E una T-indistingubilidad y H_E su conjunto de extensionales. Entonces, $\forall \mu \in H_E, (\mu_i)_{i\in I}$ familia de extensionales y $\forall \alpha \in [0,1]$ se verifica:*

1. $\bigvee_{i\in I} \mu_i \in H_E$
2. $\bigwedge_{i\in I} \mu_i \in H_E$
3. $T(\alpha,\mu) \in H_E$
4. $\overrightarrow{T}(\mu|\alpha) \in H_E$
5. $\overrightarrow{T}(\alpha|\mu) \in H_E$

Dado un difuso $\mu \in [0,1]^X$, es interesante encontrar el menor extensional que contiene a μ. Ésto es logrado por el operador ϕ_E, definido a continuación. ϕ_E también puede ser visto como un operador de clausura en $[0,1]^X$ [6].

Definición 2.8. *Sea X un conjunto y E una T-indistinguibilidad sobre X. El operador ϕ_E: $[0,1]^X \to [0,1]^X$ se define $\forall x \in X$ como:*

$$\phi_E(\mu)(x) = \sup_{y \in X} T(E(x,y), \mu(y)).$$

ϕ_E está caracterizdo por las siguientes propiedades.

Proposición 2.9. *$\forall \mu, \mu' \in [0,1]^X$ se verifica:*

1. $\mu \leq \mu' \Rightarrow \phi_E(\mu) \leq \phi_E(\mu')$
2. $\mu \leq \phi_E(\mu)$
3. $\phi_E(\mu \vee \mu') = \phi_E(\mu) \vee \phi_E(\mu')$
4. $\phi_E(\phi_E(\mu)) = \phi_E(\mu)$
5. $\phi_E(\{x\})(y) = \phi_E(\{y\})(x)$
6. $\phi_E(T(\alpha, \mu)) = T(\alpha, \phi_E(\mu))$

Análogamente, dada una T-indistinguibilidad E, se define el operador ψ_E como la función que asigna a cada difuso μ el mayor extensional contenido en μ respecto a E. El operador ψ_E es el dual de ϕ_E y puede verse como un operador topológico de interior.

Definición 2.10. *Sea E una T-indistinguibilidad en X. El operador ψ_E: $[0,1]^X \to [0,1]^X$ se define $\forall x \in X$ como:*

$$\psi_E(\mu)(x) = \inf_{y \in X} \overrightarrow{T}(E(x,y) | \mu(y)).$$

ψ_E queda completamente caracterizado por las propiedades siguientes:

Proposición 2.11. *$\forall \mu, \mu' \in [0,1]^X$ se satisface:*

1. $\mu \leq \mu' \Rightarrow \psi_E(\mu) \leq \psi_E(\mu')$
2. $\psi_E(\mu) \leq \mu$
3. $\psi_E(\mu \wedge \mu') = \psi_E(\mu) \wedge \psi_E(\mu')$
4. $\psi_E(\psi_E(\mu)) = \psi_E(\mu)$
5. $\psi_E(\overrightarrow{T}(x|\alpha)(y)) = \psi_E(\overrightarrow{T}(y|\alpha))(x)$
6. $\psi_E(\overrightarrow{T}(\alpha|\mu)) = \overrightarrow{T}(\alpha$
   $psi_E(\mu))$

Un resultado interesante es que $Im(\phi_E) = Im(\psi_E) = H_E$, y por tanto el conjunto de extensionales H_E se puede caracterizar como el conjunto de puntos fijos de ϕ_E y ψ_E. Este hecho es consecuencia directa de las definiciones y trivial desde la perspectiva topológica.

Además de la interpretación topológica que se ha dado a estos operadores de aproximación, cabe remarcar que estos conceptos también aparecen de forma natural en el marco de los fuzzy rough sets y de las lógicas modales fuzzy. En este último campo, ϕ_E se puede ver como un operador de posibilidad y ψ_E como una de necesidad.[10]

Por último, recordamos el concepto de Conexión de Galois entre conjuntos ordenados ya que aparecerá más adelante al analizar la relación entre los conceptos introducidos:

Definición 2.12. *Sean $(A, \leq)$, $(B, \leq)$ dos conjuntos parcialmente ordenados. Una Conexión de Galois entre estos conjuntos es un par (F,G) de funciones $F : A \to B$ and $G : B \to A$ tales que:*

$F(a) \leq b \quad \Leftrightarrow \quad a \leq G(b)$ *(Conexión de Galois monótona)*
$b \leq F(a) \quad \Leftrightarrow \quad a \leq G(b)$ *(Conexión de Galois antítona)*

3. Isomorfismos y Dualidades entre Indistinguibilidades, Conjuntos de Extensionales y Aproximaciones Superiores e Inferiores

Salta a la vista que los conjuntos de extensionales y las aproximaciones superior e inferior asociadas a una indistinguibilidad E quedan caracterizados por unas propiedades muy similares. En consecuencia parece natural preguntarse hasta qué punto se hallan relacionados estos conceptos.

En este apartado se darán explícitamenmte las conocidas biyecciones entre estos conceptos ($\mathscr{E}$, $\mathscr{H}$, $\mathscr{U}$, $\mathscr{L}$) y se demostrará que estas biyecciones mantienen la estructura dada por las operaciones y el orden. En conclusión, se tiene que más allá de biyecciones estas aplicaciones son isomorfismos de retículos.

Este apartado está dividido en tres subapartados. En el primero se dará el isomorfismo entre indistinguibilidades y conjuntos de extensionales. En el segundo se tratarán indistinguibilidades y aproximaciones superiores y en el tercero se estudiará el caso análogo de las aproximaciones inferiores.

3.1. Indistinguibilidades y Conjuntos de Extensionales

Esta sección estudiará el isomorfismo entre el retículo de indistinguibilidades $\mathscr{E}$ sobre un conjunto X y el retículo $\mathscr{H}$ de conjuntos de extensionales.

De forma totalmente análoga a lo que sucede con las indistinguibilidades, dados dos conjuntos H_E y H_F de extensionales, su intersección $H_E \cap H_F$ verifica las propiedades y por ende es el conjunto de extensionales de alguna indistin-

guibilidad, pero su unión $H_E \cup H_F$ no. Para cerrar la operación unión en el conjunto $\mathscr{H}$ de conjuntos de extensionales y tener una estructura reticular hay que introducir el concepto de clausura extensional de un difuso $\mu \in [0,1]^X$.

Definición 3.1. *Sea $J \subseteq [0,1]^X$. Definimos su clausura extensional $\overline{J}$ como:*

$$\overline{J} := \bigcap_{H \in [0,1]^X, J \subseteq H} H$$

donde H es un conjunto de difusos de X que satisface las propiedades de la proposición 2.7.

El retículo $\mathscr{E}$ de indistinguibilidades está naturalmente ordenado por $\leq$, mientras que el retículo de conjuntos de extensionales $\mathscr{H}$ está ordenado por la inclusión $\subseteq$. La proposición siguiente muestra que el paso de un retículo a otro invierte el orden.

Proposición 3.2. *Sean E, F dos indistinguibilidades sobre X con conjuntos de extensionales asociados H_E y H_F respectivamente. Entonces:*

$$E \leq F \Leftrightarrow H_F \subseteq H_E$$

Demostración. $E \leq F \Leftrightarrow \forall x,y \in X \;\; E(x,y) \leq F(x,y)$, que es equivalente a $\forall x,y \; \forall \mu \in H_F \;\; T(E(x,y),\mu(y)) \leq T(F(x,y),\mu(y)) \leq \mu(x) \Leftrightarrow H_F \subseteq H_E$. □

El teorema siguiente prueba que $\mathscr{E}$ y $\mathscr{H}$ están en biyección.

Teorema 3.3. [4] *Sea H un subconjunto de $[0,1]^X$ que satisface las propiedades de 2.7. Entonces existeuna única T-indistinguibilidad E tal que $H = H_E$.*

Corolario 3.4. *$(\Gamma : E \to H_E, \Sigma : H_E \to E)$ definen una conexión de Galois antítona entre $\mathscr{E}$ y $\mathscr{H}$.*

Se ha mostrado que existe una biyección entre $\mathscr{E}$ y $\mathscr{H}$ y se ha estudiado cómo ésta al orden de los retículos. Ahora se demostrará que la estructura también se preserva (de forma inversa) por esta biyección y en consecuencia se tendrá un isomorfismo de retículos. De hecho quedará probado que $\mathscr{H}$ es el espacio dual de $\mathscr{E}$.

Teorema 3.5. *Sean E, F dos indistinguibilidades sobre un conjunto X con conjuntos de extensionales asociados H_E y H_F respectivamente. Entonces:*

1. $H_{\overline{E \cup F}} = H_E \cap H_F$
2. $H_{E \cap F} = \overline{H_E \cup H_F}$

Demostración. 1. $\mu \in H_E \cap H_F \Leftrightarrow \mu \in H_E$ y $\mu \in H_F$. Equivalentemente $E_\mu \geq E$ y $E_\mu \geq F \Rightarrow E_\mu \geq E \cup F$ y por definición de clausura extensional $E_\mu \geq \overline{E \cup F} \Leftrightarrow \mu \in H_{\overline{E \cup F}}$

2. $\supseteq$)

$E \cap F \leq E \Rightarrow H_E \subseteq H_{E \cap F}$.
$E \cap F \leq F \Rightarrow H_F \subseteq H_{E \cap F}$.
Por tanto $H_E \cup H_F \subseteq H_{E \cap F} \Rightarrow \overline{H_E \cup H_F} \subseteq H_{E \cap F}$.

$\subseteq$)

Sea $\mu \in H_{E \cap F}$. Supongamos $\mu \notin \overline{H_E \cup H_F}$. Entonces existe un conjunto de extensionales H_G de una indistinguibilidad G tal que $H_G \supseteq H_E \cup H_F$ y $\mu \notin H_G \Rightarrow \exists (x,y)$ tal que $E_\mu(x,y) < G(x,y)$. Sin pérdida de generalidad podemos asumir que $E(x,y) \leq F(x,y)$. Como $H_G \supseteq H_E \cup H_F$, $H_E \subseteq H_G$ y $H_F \subseteq H_G \Rightarrow G \leq E$ and $G \leq F$. Por tanto

$E_\mu(x,y) < G(x,y) \leq E(x,y) \leq F(x,y)$. Luego

$E_\mu(x,y) < (E \cap F)(x,y)$.

En consecuencia, $\mu \notin H_{E \cap F}$ lo que contradice la hipótesis.
Así $\mu \in \overline{H_E \cup H_F}$.

□

3.2. Indistinguibilidades y Aproximaciones Superiores

Este subapartado se centrará en demostrar que el retículo de aproximaciones superiores $\mathscr{U}$ es isomorfo al de indistinguibilidades $\mathscr{E}$ sobre un conjunto X. El hilo argumental que llevará a esta conclusión es totalmente análogo al del subapartado anterior.

Tal como se hizo con $\mathscr{E}$ y $\mathscr{H}$, para darle al conjunto $\mathscr{U}$ de aproximaciones superiores ϕ_E ($E \in \mathscr{E}$) una estructura reticular, hay que definir una ϕ-clausura.

Definición 3.6. *Sea f definida en $[0,1]^X \to [0,1]^X$. Definimos su ϕ-clausura $\overline{f}$ de f como:*

$$\overline{f} = \bigwedge_{f \leq \phi, \phi \in U} \phi$$

donde U es el conjunto de aplicaciones $\phi : [0,1]^X \to [0,1]^X$ satisfaciendo las propiedades de 2.9.

Así, la estructura de retículo es obtenida con el orden $\leq$, y las operaciones $\wedge$ y $\overline{\vee}$.

Contrariamente a lo que sucedía con los conjuntos de extensionales, en el caso de aproximaciones superiores el orden es preservado directamente (y no de forma inversa).

Proposición 3.7. *Sean E, F dos T-indistinguibilidades sobre X con aproximaciones superiores asociadas ϕ_E y ϕ_F respectivamente. Entonces:*

$$E \leq F \Leftrightarrow \phi_E \leq \phi_F$$

Dado un operador $\phi \in \mathscr{U}$ se puede obtener una indistinguibilidad de la manera siguiente.

Proposición 3.8. *Sea $\phi : [0,1]^X \to [0,1]^X$ satisfaciendo las propiedades de 2.9. Entonces la relación difusa E_ϕ definida como*

$$E_\phi(x,y) = \phi(\{x\})(y)$$

es una T-indistinguibilidad sobre X.

Demostración. Inmediato a partir de la definición. □

El teorema siguiente demuestra que existe una biyección entre $\mathscr{E}$ y $\mathscr{U}$.

Teorema 3.9. *Existe una biyección entre el conjunto de T-indistinguibilidades sobre X y las aplicaciones $\phi : [0,1]^X \to [0,1]^X$ que satisfacen 2.9.*

Demostración. Hay que probar que $\phi_{E_\phi} = \phi$ y $E_{\phi_E} = E$

- Sea ϕ una aplicación que satisface 2.9. $\forall x,y \in X$ y cualquier difuso $\mu \in [0,1]^X$ se tiene:

$$\begin{aligned} \phi_{E_\phi}(\mu)(x) &= \sup_{y\in X} T(E_\phi(x,y),\mu(y)) \\ &= \sup_{y\in X} T(\phi(\{y\})(x),\mu(y)) \\ &= \phi(\mu)(x). \end{aligned}$$

- Sea E una T-indistinguibilidad. $\forall x,y \in X$ se tiene:

$$\begin{aligned} E_{\phi_E}(x,y) &= \phi_E(\{x\})(y) \\ &= \sup_{z\in X} T(E(y,z),\{x\}(z)) \\ &= E(x,y). \end{aligned}$$

□

Corolario 3.10. *$(\Gamma : E \to \phi_E, \Sigma : \phi_E \to E)$ definen una conexión de Galois monótona entre $\mathscr{E}$ y $\mathscr{U}$.*

Hemos construido una biyección entre $\mathscr{E}$ y $\mathscr{U}$ y visto que el orden es respetado por ésta. De nuevo, veamos ahora que las operaciones también son respetadas y por tanto la biyección dada es en realidad un isomorfismo.

Teorema 3.11. *Sean E, F dos T-indistinguibilidades con aproximaciones superiores asociadas ϕ_E and ϕ_F respectivamente. Entonces:*

1. $\phi_{\overline{E\cup F}} = \overline{\phi_E \vee \phi_F}$
2. $\phi_{E\cap F} = \phi_E \wedge \phi_F$

Demostración. 1. $\geq$)

Sea $\mu \in [0,1]^X$.

$\phi_{\overline{E\cup F}}(\mu)(x) = \sup_{y\in X} T(\overline{E\cup F}(x,y),\mu(y))$.

Como $\overline{E\cup F} \geq E,F \Rightarrow \phi_{\overline{E\cup F}}(\mu)(x) \geq \phi_E(\mu)(x)$ y $\phi_{\overline{E\cup F}}(\mu)(x) \geq \phi_F(\mu)(x)$. Luego:

$\phi_{\overline{E\cup F}}(\mu)(x) \geq (\phi_E \vee \phi_F)(\mu(x)) \Rightarrow \phi_{\overline{E\cup F}} \geq \overline{\phi_E \vee \phi_F}$.

$\leq$)

Por la definición 3.6 y la proposición 2.9, existe una única T-indistinguibilidad G sobre X tal que $\overline{\phi_E \vee \phi_F} = \phi_G$. Tenemos que demostrar que $\overline{E\cup F} \leq G$.

En efecto, $\phi_E \leq \overline{\phi_E \vee \phi_F} = \phi_G$ y por el lema 3.7 $E \leq G$. Análogamente tenemos que $F \leq G$, y por tanto $E \cup F \leq G \Rightarrow \overline{E\cup F} \leq G$, lo que prueba la desigualdad.

2. $\leq$

$E \cap F \leq E \Rightarrow \phi_{E\cap F} \leq \phi_E$
$E \cap F \leq F \Rightarrow \phi_{E\cap F} \leq \phi_F$
Por tanto $\phi_{E\cap F} \leq \phi_E \wedge \phi_F$

$\geq$

$\phi_{E\cap F}(\mu)(x) = \sup_{y\in X} T(\text{mín}\{E(x,y),F(x,y)\},\mu(y))$
$= \sup_{y\in X} \text{mín}\{T(E(x,y),\mu(y)),T(F(x,y),\mu(y))\} \geq \phi_E(\mu)(x) \wedge \phi_F(\mu)(x)$.

□

3.3. Indistinguibilidades y Aproximaciones Inferiores

En este subapartado se construirá y estudiará de forma análoga a los anteriores el isomorfismo entre el retículo de indistinguibilidades y el de aproximaciones inferiores por extensionales. El operador de aproximación inferior ψ_E asociado a una indistinguibilidad E se comporta de forma dual al de aproximación superior ϕ_E y como se verá los resultados son duales a los de la subsección anterior.

En primer lugar, para darle al conjunto de aproximaciones inferiores $\mathscr{L}$ una estructura de retículo hay que definir el concepto de ψ-clausura de una aplicación.

Definición 3.12. *Sea f definida en $[0,1]^X \to [0,1]^X$. Definimos la ψ-clausura $\underline{f}$ de f como:*

$$\underline{f} = \bigvee_{\psi\in L, f\leq\psi} \psi$$

donde L es el conjunto de aplicaciones $\psi : [0,1]^X \to [0,1]^X$ que satisfacen 2.11.

Veamos ahora c´0mo el orden es preservado en el paso de un retículo a otro.

Proposición 3.13. *Sean E, F dos indistinguibilidades sobre X con aproximaciones inferiores asociadas ψ_E and ψ_F respectivamente. Entonces:*

$$E \leq F \Leftrightarrow \psi_E \leq \psi_F.$$

De esta forma, la estructura de retículo en $\mathscr{L}$ es dada por el orden $\leq$ y las operaciones $\wedge$ y $\veebar$

Proposición 3.14. *Sea $\psi : [0,1]^X \to [0,1]^X$ una aplicación que satisface 2.11. Entonces la relación difusa E_ψ definida como*

$$E_\psi(x,y) = \inf_{\alpha\in[0,1]} \overrightarrow{T}(\psi(\overrightarrow{T}(\{x\}|\alpha)(y)|\alpha)).$$

es una T-indistinguibilidad sobre X.

La proposición siguiente prueba que $\mathscr{E}$ y $\mathscr{L}$ están en biyección.

Teorema 3.15. [9] *Hay una biyección entre el conjunto de T-indistinguibilidades y el de aplicaciones ψ que satisfacen 2.11.*

Corolario 3.16. $(\Gamma : E \to \psi_E, \Sigma : \psi_E \to E)$ *definen una conexión de Galois monótona entre $\mathscr{E}$ and $\mathscr{L}$.*

Demostración. Trivial por las propiedades de la residuación. □

Sólo queda por ver que la biyección dada también respeta la estructura de las operaciones $\wedge$ y $\widetilde{\vee}$. Ésto se demuestra en el siguiente teorema.

Teorema 3.17. *Sean E, F dos T-indistinguibilidades con aproximaciones inferiores asociadas ψ_E y ψ_F. Entonces:*

1. $\psi_{\overline{E\cup F}} = \underline{\psi_E \wedge \psi_F}$
2. $\psi_{E\cap F} = \psi_E \vee \psi_F$

La demostración del teorema es dual a la dada en el teorema análogo de la sección anterior.

4. Conclusion

En este trabajo la relación entre los conjuntos de T-indistinguibilidades $\mathscr{E}$, conjuntos de extensionales $\mathscr{H}$, y aproximaciones superiores e inferiores por extensionales $\mathscr{U}$ y $\mathscr{L}$ sobre un conjunto X ha sido totalmente estudiada. Se ha demostrado que estos conjuntos son en realidad retículos isomorfos y estos isomorfismos han sido dados explícitamente. Estos resultados se pueden resumir en el siguiente diagrama:

$$\begin{array}{ccccccc}
\mathscr{L} & \approx & \mathscr{U} & \approx & \mathscr{E} & \approx & \mathscr{H} \\
\psi_E \vee \psi_F & \leftrightarrow & \phi_E \wedge \phi_F & \leftrightarrow & E\cap F & \leftrightarrow & \overline{H_E \cup H_F} \\
\underline{\psi_E \wedge \psi_F} & \leftrightarrow & \overline{\phi_E \vee \phi_F} & \leftrightarrow & \overline{E\cup F} & \leftrightarrow & H_E \cap H_F
\end{array}$$

Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto número TIN2009-07235.

El primer autor está financiado por la Universitat Politècnica de Catalunya (UPC).

Referencias

[1] Boixader, D., Jacas, J., Recasens, J.: Upper and lower approximation of fuzzy sets. Int. J. of General Systems 29, 555-568 (2000)

[2] Boixader, D., Jacas, J., Recasens, J.: Fuzzy equivalence relations: advanced material. In: Dubois, D., Prade, H. (eds.) Fundamentals of Fuzzy Sets, pp. 261-290. Kluwer Academic Publishers, New York (2000)

[3] Bělohlavek, R.: Fuzzy Relational Systems: Foundations and Principles. Kluwer Academic Publishers, New York (2002)

[4] Castro, J.L., Klawonn, F.: Similarity in Fuzzy Reasoning. Mathware & Soft Computing 2, 197-228 (1996)

[5] Jacas, J., Recasens, J.: Fixed points and generators of fuzzy relations. J. Math. Anal. Appl 186, 21-29 (1994)

[6] Jacas, J., Recasens, J.: Fuzzy T-transitive relations: eigenvectors and generators. Fuzzy Sets and Systems 72, 147-154 (1995)

[7] Jacas, J., Recasens, J.: Aggregation of T-Transitive Relations. Int J. of Intelligent Systems 18, 1193-1214 (2003)

[8] Klement, E.P., Mesiar, R., Pap E.: Triangular norms. Kluwer Academic Publishers, Dordrecht (2000)

[9] Morsi, N.N., Yakout, M.M.: Axiomatics for fuzzy rough sets. Fuzzy Sets and Systems 100, 327-342 (1998)

[10] Radzikowska, A.M., Kerre, E.E: Characterisation of main classes of fuzzy relations using fuzzy modal operators. Fuzzy Sets and Systems 152, 223-247 2005

[11] Recasens, J. Indistinguishability Operators. Modelling Fuzzy Equalities and Fuzzy Equivalence Relations. Series: Studies in Fuzziness and Soft Computing, Vol. 260 (2011)

[12] Trillas, E.: Assaig sobre les relacions d'indistingibilitat. Proc. Primer Congrés Català de Lògica Matemàtica, Barcelona, 51-59 (1982) (En Catalán)

[13] Valverde, L.: On the Structure of F-indistinguishability Operators. Fuzzy Sets and Systems 17, 313-328 (1985)

[14] Zadeh. L.A. Similarity relations and fuzzy orderings. Inform. Sci. 3, 177-200 (1971)

A RECURSIVE ALGORITHM TO COMPUTE A BASIS OF A SIMILARITY

Luis Garmendia[1], Jordi Recasens[2], Adela Salvador[3]

Complutense University of Madrid, lgarmend@fdi.ucm.es
Technical University of Catalunya, j.recasens@upc.edu
Technical University of Madrid, ma09@caminos.upm.es

Abstract

This paper studies some theory and methods to build a representation theorem basis of a similarity from the basis of its subsimilarities, providing an alternative recursive method to compute the basis of a similarity.

Keywords: T-indistinguishabilities representation theorem, similarity, basis of a similarity, dimension of a similarity, similarity representation theorem.

1 INTRODUCTION

Similarity fuzzy relations were introduced by Zadeh [1971] [10] to represent a degree of equality or closeness between the elements of a universe. They generalises the equivalence relations; in fact, the similarities are the only T-indistinguishabilities that satisfy that all their alpha cuts are crisp equivalence relations. Similarities are not only a powerful tool to represent equality information, but they are also useful to classify the universe into clusters with uncertainty.

The Valverde´s representation theorem of T-indistinguisabilities is one of the strongest theorems in fuzzy logic theory. It opened an interesting area of investigation on T-indistinguishabilities, the computation of its basis, specially for both the minimum t-norm and Archimedean t-norms, and the study of special T-indistinguishabilities such as the one dimensional ones.

On the other hand, the computational investigation on the computation of T-transitive closures, or the search of the structure of a similarity, specially its decomposition into subsimilarities, were not related in previous studies with the computation of the basis of a similarity.

Joan Jacas [4] studied two algorithms to compute basis of a similarity in 1990, but did not consider the decomposition of similarities to do it. The aim of this paper is to introduce a new approach to compute them, providing useful theory and methods for both better understanding the concept of structure of a similarity and the Valverde's representation theorem for similarities and the computation of their basis using a new decomposition approach.

2 PRELIMINARIES

Let $X = \{x_1, ..., x_n\}$ be a finite universe.

Let T be a t-norm. A T-indistinguishability operator E on X is a fuzzy relation $E: X \times X \to [0, 1]$, satisfying for all x, y, z in X:

1. $E(x, x) = 1$ (Reflexivity)
2. $E(x, y) = E(y, x)$ (Symmetry)
3. $T(E(x, y), E(y, z)) \leq E(x, z)$ (T-transitivity)

Definition 2.2. A **similarity** [Zadeh, 1971] [10] is a reflexive, symmetric and min-transitive fuzzy relation, it is, a similarity is a Min-indistinguishability operator.

Notation 2.1.

We can denote $x_{ij} = E(x_i, x_j)$.

Lemma 2.1. Let π be a permutation on X. If E is a similarity on X, then the fuzzy relation $P_\pi(E)$ is also a fuzzy similarity.

Proof. It is obvious. $P_\pi(E)$ is reflexive and symmetric. If $x_{ij} \geq \text{Min}\{x_{ik}, x_{kj}\}$ for all i, j, k then $x_{rs} = x_{\pi(i)\ \pi(j)} \geq \text{Min}\{x_{\pi(i)\ k}, x_{k\pi(j)}\} = T\{x_{rk}, x_{ks}\}$ for all $1 \leq r, s, k \leq n$.

2.1. CONSTRUCTION OF A FUZZY SIMILARITY FROM SUBSIMILARITIES

Let C and D be two similarities on two disjoint sets X_1 and X_2. A similarity relation E(F; C, D) on $X_1 \cup X_2$ can be built with the following shape:

$$E(F; C, D) = \begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$$

A method for giving the bridging values f_{ij} in F, (when $j \leq card(X_1) < i$) is the assignation of a unique value f, in all the $card(X_1) \times card(X_2)$ values in F. This value must be chosen in an interval [0, a] where $a = \min\{\min(C), \min(D)\}$. The values in F^T are the symmetric values f of the computed F.

So the computed values in F are equal and satisfy that $f \leq \min\{\min(C), \min(D)\}$.

Lemma 2.1.1. [5] Lee [2001]

If C and D are fuzzy similarities, then E(f; C, D) is also a fuzzy similarity, $\forall f \in [0, \min(\min(C), \min(D))]$.

Lemma 2.1.2. [5] Lee [2001]

If $S_{n\times n}$ is a fuzzy similarity, then there exists a decomposition such that $S_{n\times n} = P_{\pi}(E(f; C_{n1\times n1}, D_{n2\times n2}))$.

2.2. THE REPRESENTATION THEOREM OF T-INDISTINGUISHABILITY OPERATORS

The representation theorem allow us to generate a *T*-indistinguishability operator on a set X from a family of subsets on X, and reciprocally states that every *T*-indistinguishability can be obtained in this form.

Definition 2.2.1 The residuation $\vec{T}$ or quasi inverse of a t-norm *T* is the map $\vec{T}$: $[0,1] \times [0,1] \to [0,1]$ defined for all x, y in [0, 1] by

$$\vec{T}\,(x \mid y) = \sup\{\propto \in [0,1] \mid T(x,\alpha) \leq y\}\,.$$

The residuation $\vec{T}$ of a t-norm *T* is a *T*-preorder (reflexive and *T*-transitive) on [0, 1], and then it is a useful operator to generate implication relations from fuzzy sets on X.

Note that $\overrightarrow{Min}\,(x \mid y) = \begin{cases} 1 & if\ x \leq y \\ y & if\ x > y \end{cases}$.

Definition 2.2.2 The biresiduation $\overleftrightarrow{T}$ (or also E_T) of a t-norm *T* is the map $\overleftrightarrow{T}$: $[0,1] \times [0,1] \to [0,1]$ defined for all x, y in [0, 1] by

$$\overleftrightarrow{T}\,(x,\ y) = E_T(x,y) = Min\,(\vec{T}\,(x \mid y),\ \vec{T}\,(y \mid x)).$$

Note that $\overleftrightarrow{Min}\,(x,\ y) = \overleftrightarrow{Min}(\mathrm{Max}(x,\ y),\ \mathrm{Min}(x,y))$

$$= \begin{cases} 1 & if\ x = y \\ Min(x,y) & if\ x \neq y \end{cases}.$$

The biresiduation $\overleftrightarrow{T}$ of a t-norm *T* is a *T*-indistinguishability operator on [0, 1], and then it is a useful operator to generate *T*-indistinguishability relations on X from two fuzzy sets on X.

Lemma 2.2.1.

Let μ be a fuzzy set on X, and *T* a continuous t-norm. The fuzzy relation E_μ on X defined for all x, y in X by $E_\mu(x,y) = E_T(\mu(x),\mu(y))$ is a T-indistinguishability operator.

Note that E_μ is a one-dimensional *T*-indistinguishability generated by a basis of one fuzzy set μ.

Lemma 2.2.2.

Let $(E_i)_{i\in I}$ be a family of *T*-indistinguishability operators on a set X. The relation E on X defined for all x, y in X by

$$E(x,y) = \inf_{i\in I} E_{\mu i}(x,y)$$

is a T-indistinguishability operator.

The next theorem is crucial to understand the structure of a T-indistinguishability operator. It allows us to generate *T*-indistinguishabilities from a family of fuzzy sets, and reciprocally that any *T*-indistinguishability can be generated from a family of fuzzy sets.

Representation theorem 2.2.1. [9]

Let R be a fuzzy relation on X and T a continuous t-norm. R is a T-indistinguishability operator if and only if there exists a family $(\mu_i)_{i\in I}$ of fuzzy sets on X such that for all x, y in X

$$R(x,y) = \inf_{i\in I} E_{\mu i}(x,y).$$

Definition 2.2.1. Dimension and basis of a T-indistinguishability operator

The dimension of a *T*- indistinguishability operator E is the minimal of the cardinalities *d* of the generating families of E, in the sense of the representation theorem.

That minimal basis of generating fuzzy sets is called a basis of the *T*-indistinguishability operator.

3 DECOMPOSITION OF SIMILARITIES AND REPRESENTATION THEOREM

This chapter gives some theoretical results toward a method to compute a basis of a similarity from the bases of its subsimilarities.

Lemma 3.1. Let S be the following similarity on $X_1 \cup X_2$,

S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$

Then dim(S) ≥ dim(C).

Lemma 3.2.

Let $(\mu_i)_{i\in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1. Let $(\gamma_i)_{i\in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$ where F = $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix.

Suppose that r > s and f < min(C) and $f \leq$ min(D) then dim(S) ≤ dim(C) and a generator set of S is $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix}, \ldots, \begin{pmatrix}(\mu_s)\\(\gamma_s)\end{pmatrix}, \begin{pmatrix}(\mu_{s+1})\\F'\end{pmatrix}, \ldots, \begin{pmatrix}(\mu_r)\\F'\end{pmatrix}\right\}$ where F' = $\begin{pmatrix} f \\ \vdots \\ f \end{pmatrix}$ is a $card(X_2) \times 1$ matrix.

Proposition 3.1.

Let $(\mu_i)_{i\in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1. Let $(\gamma_i)_{i\in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$ where F = $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix.

Suppose that r > s and $f <$ min(C) and $f \leq$ min(D) then

1) dim(S) = dim(C) = r
2) a basis of S is $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix}, \dots, \begin{pmatrix}(\mu_s)\\(\gamma_s)\end{pmatrix}, \begin{pmatrix}(\mu_{s+1})\\F'\end{pmatrix}, \dots, \begin{pmatrix}(\mu_r)\\F'\end{pmatrix}\right\}$

where F' = $\begin{pmatrix} f \\ . \\ . \\ f \end{pmatrix}$ is a $card(X_2) \times 1$ matrix.

Proof:

By Lemma 3.1 dim(S) ≥ dim(C).

By Lemma 3.2 dim(S) ≤ dim(C). Also dim(S) = dim(C) = r and

$\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix}, \dots, \begin{pmatrix}(\mu_s)\\(\gamma_s)\end{pmatrix}, \begin{pmatrix}(\mu_{s+1})\\F'\end{pmatrix}, \dots, \begin{pmatrix}(\mu_r)\\F'\end{pmatrix}\right\}$ is a basis of S(F; C, D). □

Corollary 3.1.

Let $(\mu_i)_{i\in I}$ be a representations theorem basis of a similarity C on a finite set X = $\{x_1, \dots, x_n\}$.

Let S be the similarity on X ∪ $\{x_{n+1}\}$ such that S(F; C, 1) = $\begin{pmatrix} C & F'^T \\ F' & 1 \end{pmatrix}$ where F' = (f ... f) is a $1 \times card(X)$ matrix.

If $f <$ min(C), then a basis of S is

$\{\mu_1, \dots, \mu_{\dim(C)}\} \times \{f\} = \left\{\begin{pmatrix}\mu_1(x_1)\\\vdots\\\mu_1(x_n)\\f\end{pmatrix}, \dots, \begin{pmatrix}\mu_{\dim(C)}(x_1)\\\vdots\\\mu_{\dim(C)}(x_n)\\f\end{pmatrix}\right\}$

Example 3.1.

Let S be the similarity $\begin{pmatrix} 1 & a & b & c \\ a & 1 & b & c \\ b & b & 1 & c \\ c & c & c & 1 \end{pmatrix}$, already ordered with c < b < a.

S= $\begin{pmatrix}\begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix} & \begin{matrix} c \\ c \\ c \end{matrix} \\ \begin{matrix} c & c & c \end{matrix} & 1 \end{pmatrix}$,

and a base of $\begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix}$ *is* $\left\{\begin{pmatrix}1\\a\\b\end{pmatrix}\right\}$, then by the theorem 3.1 a basis of S is

$\left\{\begin{pmatrix}1\\a\\b\end{pmatrix}\right\} \times \{c\} = \left\{\begin{pmatrix}1\\a\\b\\c\end{pmatrix}\right\}$.

Note that S is one dimensional.

Example 3.2.

Let S be the similarity $\begin{pmatrix} 1 & a & a & b \\ a & 1 & a & b \\ a & a & 1 & b \\ b & b & b & 1 \end{pmatrix}$, already ordered with b < a.

S= $\begin{pmatrix}\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix} & \begin{matrix} b \\ b \\ b \end{matrix} \\ \begin{matrix} b & b & b \end{matrix} & 1 \end{pmatrix}$,

A basis of $\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$ is $\left\{\begin{pmatrix}1\\a\\a\end{pmatrix}, \begin{pmatrix}1\\1\\a\end{pmatrix}\right\}$

Then, as b < min$\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$, by theorem 3.1, a basis of S is $\left\{\begin{pmatrix}1\\a\\a\end{pmatrix}, \begin{pmatrix}1\\1\\a\end{pmatrix}\right\} \times \{b\} = \left\{\begin{pmatrix}\begin{pmatrix}1\\a\\a\end{pmatrix}\\b\end{pmatrix}, \begin{pmatrix}\begin{pmatrix}1\\1\\a\end{pmatrix}\\b\end{pmatrix}\right\}$

Lemma 3.2.

Let $(\mu_i)_{i\in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1. Let $(\gamma_i)_{i\in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$, where F= $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix

Suppose that r > s, f = min(C) and $f \leq$ min(D), then:

1) dim(S) > dim(C) = r.

Lemma 3.3.

Let $(\mu_i)_{i\in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1 Let $(\gamma_i)_{i\in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$,

where F = $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix

Suppose that r > s, f = min(C) and f ≤ min(D), then:

1) dim(S) ≤ dim(C) +1 = r + 1
2) A generator set of S is A basis of S is $\left\{ \begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_s) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (\mu_{s+1}) \\ F' \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_r) \\ F' \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix} \right\}$,

where $(1) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ is a $\mathrm{card}(X_1) \times 1$ matrix and F' = $\begin{pmatrix} f \\ \vdots \\ f \end{pmatrix}$ is a $\mathrm{card}(X_2) \times 1$ matrix.

Proof:

Now $\left\{ \begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_s) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (\mu_{s+1}) \\ F' \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_r) \\ F' \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix} \right\}$ is a generator set of S, because

$inf\left\{ \inf_{1 \le i \le s} E_{\binom{(\mu_i)}{(\gamma_i)}}, \inf_{r+1 \le j \le r} E_{\binom{(\mu_j)}{F}} \right\} = \begin{pmatrix} C & F* \\ F* & D \end{pmatrix}$ = S(F; C, D).

Also $\begin{pmatrix} (1) \\ F' \end{pmatrix} o_{\overline{Min}} ((1), F') = \begin{pmatrix} (1) & F^T \\ F & (1) \end{pmatrix}$, so the infimum of all the similarities is $\begin{pmatrix} C & F^T \\ F & D \end{pmatrix}$ = S(F; C, D).□

Proposition 3.2.

Let $(\mu_i)_{i \in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1. Let $(\gamma_i)_{i \in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$, where F= $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix

Suppose that r > s, f = min(C) and $f \le$ min(D), then:

1) dim(S) = dim(C) +1 = r + 1
2) A basis of S is: $\left\{ \begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_s) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (\mu_{s+1}) \\ F' \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_r) \\ F' \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix} \right\}$

where $(1) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ is a $\mathrm{card}(X_1) \times 1$ matrix and F' = $\begin{pmatrix} f \\ \vdots \\ f \end{pmatrix}$ is a $\mathrm{card}(X_2) \times 1$ matrix.

Proof:

By lemma 3.2 dim(S) > dim(C) = r, and by lemma 3.3 dim(S) ≤ dim(C) +1 = r + 1, then dim(S) = dim(C) +1 = r + 1, and a basis of S is $\left\{ \begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_s) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (\mu_{s+1}) \\ F' \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_r) \\ F' \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix} \right\}$,

Example 3.5

Let S be the similarity $\begin{pmatrix} 1 & a & b & b \\ a & 1 & b & b \\ b & b & 1 & b \\ b & b & b & 1 \end{pmatrix}$, already ordered with b < a.

S = $\left(\begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix} \begin{matrix} b \\ b \\ b \end{matrix} \right)$, and a basis of $\begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix}$, is $\left\{ \begin{pmatrix} 1 \\ a \\ b \end{pmatrix} \right\}$, then by the propositions 3.1 and 3.2 a generator set of S is = $\left\{ \left(\begin{pmatrix} 1 \\ a \\ b \\ b \end{pmatrix} \right), \begin{pmatrix} 1 \\ 1 \\ 1 \\ b \end{pmatrix} \right\}$.

Proposition 3.3.

Let $(\mu_i)_{i \in I}$ be a representation theorem basis of a similarity C of dimension r on a finite set X_1. Let $(\gamma_i)_{i \in J}$ be a representation theorem basis of a similarity D of dimension s on a finite set X_2.

Let S be the following similarity on $X_1 \cup X_2$, S(F; C, D) = $\begin{pmatrix} \boxed{C} & \boxed{F^T} \\ \boxed{F} & \boxed{D} \end{pmatrix}$,

where F = $\begin{pmatrix} f & \cdots & f \\ \vdots & \ddots & \vdots \\ f & \cdots & f \end{pmatrix}$ is a $card(X_1) \times card(X_2)$ matrix.

Suppose that $r = s$ (dim(C) = dim(D) and f < min(C, D), then a generator set of S is:

$$\left\{ \begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \ldots, \begin{pmatrix} (\mu_r) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix} \right\}$$

and dim(S) ≤ r + 1.

Proof:

Let $\{\mu_1, \ldots, \mu_r\}$ be a representation theorem basis of a similarity C, so

C = $\inf_{i \in I} E_{\mu i}$ where I = {1, ..., r}.

In the same way, let $\{\gamma_1, ..., \gamma_s\}$ be a representation theorem basis of a similarity D, so

D = $\inf_{j\in J} E_{\gamma j}$ where J = {1, ..., r}

Now, for all $1 \leq i \leq r$,

$\begin{pmatrix}(\mu_i)\\(\gamma_i)\end{pmatrix}$ $o_{\overrightarrow{Min}}((\mu_i) \quad (\gamma_i)) = \begin{pmatrix}(E_{\mu i}) & (x_{nm})\\(x_{mn}) & (E_{\gamma i})\end{pmatrix}$

Also $\begin{pmatrix}(1)\\F'\end{pmatrix}$ $o_{\overrightarrow{Min}}$ $((1) \quad F') = \begin{pmatrix}(1) & F^T\\F & (1)^T\end{pmatrix}$,

So the infimum of all the similarities is $\begin{pmatrix}C & F^T\\F & D\end{pmatrix} = S$(F; C, D), and then $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix}, ..., \begin{pmatrix}(\mu_r)\\(\gamma_s)\end{pmatrix}, \begin{pmatrix}(1)\\F'\end{pmatrix}\right\}$ is a generator set of S. □

Example 3.4.

Let S be the similarity $\begin{pmatrix}1 & a & b & b\\a & 1 & b & b\\b & b & 1 & a\\b & b & a & 1\end{pmatrix}$, already ordered, with b < a.

S = $\begin{pmatrix}\begin{pmatrix}1 & a\\a & 1\end{pmatrix} & \begin{matrix}b & b\\b & b\end{matrix}\\ \begin{matrix}b & b\\b & b\end{matrix} & \begin{pmatrix}1 & a\\a & 1\end{pmatrix}\end{pmatrix}$

A basis of $\begin{pmatrix}1 & a\\a & 1\end{pmatrix}$, is $\left\{\begin{pmatrix}1\\a\end{pmatrix}\right\}$, then by the proposition 3.3 a generator set (and a basis in this example) of S is $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix}, ..., \begin{pmatrix}(\mu_r)\\(\gamma_s)\end{pmatrix}, \begin{pmatrix}(1)\\F'\end{pmatrix}\right\} = \left\{\begin{pmatrix}1\\a\\1\\a\end{pmatrix}, \begin{pmatrix}1\\1\\b\\b\end{pmatrix}\right\}$.

Note that there exists other bases of S, for example $\left\{\begin{pmatrix}1\\a\\b\\b\end{pmatrix}, \begin{pmatrix}b\\b\\1\\a\end{pmatrix}\right\}$.

Let S be the similarity $\begin{pmatrix}1 & a & b & b\\a & 1 & b & b\\b & b & 1 & a\\b & b & a & 1\end{pmatrix}$, already ordered, with b < a.

S= $\begin{pmatrix}\begin{pmatrix}1 & a\\a & 1\end{pmatrix} & \begin{matrix}b & b\\b & b\end{matrix}\\ \begin{matrix}b & b\\b & b\end{matrix} & \begin{pmatrix}1 & a\\a & 1\end{pmatrix}\end{pmatrix}$

A base of $\begin{pmatrix}1 & a\\a & 1\end{pmatrix}$ *is* $\left\{\begin{pmatrix}1\\a\end{pmatrix}\right\}$,

Then by proposition 3.3 a basis of S is

$= \left\{\begin{pmatrix}1\\a\\b\\b\end{pmatrix} \quad \begin{pmatrix}b\\b\\1\\a\end{pmatrix}\right\}$.

4 GENERATIING A DECOMPOSITION OF A GIVEN SIMILARITY

Input: A similarity S on X.

Output:

- Two similarities S_1, and S_2 on X_1 and X_2 such that X = $X_1 \cup X_2$
- A bridging value f such that f < min(S_1) and f ≤ min(S_2)

Algorithm 5.1:

1. Sort the universe X having descending columns under the diagonal using algorithm 4.1
2. Compute the frequency k of the lowest value v in the first column
3. If k ≥ card(X) -1 then STOP. S_1 = S.
4. Let X_1 be $\{x_1, ..., x_{k+1}\}$ and let X_2 be $\{x_{k+2}, ..., x_{card(X)}\}$
5. Let S_1 = S| X_1, S_2 = S| X, and f = v.

5 A RECURSIVE ALGORITHM TO COMPUTE A GENERATING SET OF A SIMILARITY

Input: a similarity S on a universe X.

Output: A generating set of S, and so, an upper bound of the dimension of S.

A recursive algoritm:

1) If card(X) ≤ 2 then STOP. dim(S) = 1 and a basis of S is the first column.
2) Descompose S into S_1, and S_2 using algorithm 5.1. Note that f < min(S_1)
3) IF card(X_2) = 1 then use corollary 3.1:
   a. Recursively compute a generating set of S_1, $(\mu_i)_{i\in I}$
   b. A generating set of S is $\{(\mu_i)_{i\in I}\} \times \{f\}$
   c. STOP
4) Recursively compute a generating set of S_1, $(\mu_i)_{i\in I}$ and S_2, $(\gamma_i)_{i\in J}$
5) IF dim(S_1) > dim(S_2) -or vice versa- then use proposition 3.1
   a. A generating set of S is $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix} ... \begin{pmatrix}(\mu_s)\\(\gamma_s)\end{pmatrix} \begin{pmatrix}(\mu_{s+1})\\F\end{pmatrix} ... \begin{pmatrix}(\mu_r)\\F\end{pmatrix}\right\}$
   b. STOP
6) IF dim(S_1) = dim(S_2) then use proposition 3.3
   a. A generating set of S is $\left\{\begin{pmatrix}(\mu_1)\\(\gamma_1)\end{pmatrix} ... \begin{pmatrix}(\mu_r)\\(\gamma_s)\end{pmatrix} \begin{pmatrix}(1)\\F\end{pmatrix}\right\}$
   b. STOP

Example 6.1.

Let S be the sorted similarity with descending columns under the diagonal given by the following matrix:

$$S = \begin{pmatrix} 1 & 0,9 & 0,4 & 0,4 & 0,3 \\ 0,9 & 1 & 0,4 & 0,4 & 0,3 \\ 0,4 & 0,4 & 1 & 0,7 & 0,3 \\ 0,4 & 0,4 & 0,7 & 1 & 0,3 \\ 0,3 & 0,3 & 0,3 & 0,3 & 1 \end{pmatrix}$$

It will be computed the algorithm 6.1 step by step.

A decomposition of S is given by algorithm 5.1 is

$$S = \begin{pmatrix} \begin{pmatrix} 1 & 0,9 & 0,4 & 0,4 \\ 0,9 & 1 & 0,4 & 0,4 \\ 0,4 & 0,4 & 1 & 0,7 \\ 0,4 & 0,4 & 0,7 & 1 \end{pmatrix} & \begin{matrix} 0,3 \\ 0,3 \\ 0,3 \\ 0,3 \end{matrix} \\ \begin{matrix} 0,3 & 0,3 & 0,3 & 0,3 \end{matrix} & (1) \end{pmatrix}, S_1 = \begin{pmatrix} 1 & 0,9 & 0,4 & 0,4 \\ 0,9 & 1 & 0,4 & 0,4 \\ 0,4 & 0,4 & 1 & 0,7 \\ 0,4 & 0,4 & 0,7 & 1 \end{pmatrix},$$

Now that S is ordered and decomposed, we can build its basis as follows:

A basis of $\begin{pmatrix} 1 & 0,9 \\ 0,9 & 1 \end{pmatrix}$ is $\left\{\begin{pmatrix} 1 \\ 0,9 \end{pmatrix}\right\}$. A basis of $\begin{pmatrix} 1 & 0,7 \\ 0,7 & 1 \end{pmatrix}$ is $\left\{\begin{pmatrix} 1 \\ 0,7 \end{pmatrix}\right\}$, and dimC = 2 = dimD, f = 0,4 < min(C, D). Then by proposition 3.3, a generator set of $S_1 = \begin{pmatrix} \begin{pmatrix} 1 & 0,9 \\ 0,9 & 1 \end{pmatrix} & \begin{matrix} 0,4 & 0,4 \\ 0,4 & 0,4 \end{matrix} \\ \begin{matrix} 0,4 & 0,4 \\ 0,4 & 0,4 \end{matrix} & \begin{pmatrix} 1 & 0,7 \\ 0,7 & 1 \end{pmatrix} \end{pmatrix}$ is $\left\{\begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \dots, \begin{pmatrix} (\mu_r) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (1) \\ F' \end{pmatrix}\right\} = \left\{\begin{pmatrix} 1 \\ 0.9 \\ 1 \\ 0.7 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0,4 \\ 0,4 \end{pmatrix}\right\}$

And then, by proposition 3.1, a generator set of $S = \begin{pmatrix} S_1 & 0,3 \\ 0,3 & (1) \end{pmatrix}$ is $\left\{\begin{pmatrix} (\mu_1) \\ (\gamma_1) \end{pmatrix}, \dots, \begin{pmatrix} (\mu_s) \\ (\gamma_s) \end{pmatrix}, \begin{pmatrix} (\mu_{s+1}) \\ F' \end{pmatrix}, \dots, \begin{pmatrix} (\mu_r) \\ F' \end{pmatrix}\right\} =$

$$\left\{\begin{pmatrix} 1 \\ 0,9 \\ 1 \\ 0,7 \\ 0,3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0,4 \\ 0,4 \\ 0,3 \end{pmatrix}\right\}$$

6 CONCLUDING REMARKS

This paper´s main contribution is a method to compute a representation theorem basis of a similarity from the bases of its subsimilarities.

These results can be used to propose an alternative algorithm to build a basis of similarities.

The bases of all structures of similarities with dimension four are computed in the examples using the new construction method.

7 ACKNOWLEDGMENT

We thank the support of the research projects TIN2009-07901, CAM GR35/10-A and TIN2009-07235.

8 REFERENCES

[1]. J. L. Castro, F. Klawonn, F.: Similarity in Fuzzy Reasoning. *Mathware & Soft Computing* 2, 197-228, 1996

[2]. J. C. Fodor, M. Roubens: Structure of transitive valued binary relations. *Math. Soc. Sci. 30*,71-94, 1995.

[3]. J. Jacas: Similarity relations. The calculation of minimal generating families. *Fuzzy Sets and Systems* 35, 151-162, 1990.

[4]. J. Jacas, R. Recasens: J. One dimensional indistinguishability operators. *Fuzzy Sets and Systems* 109, 447-451, 2000

[5]. H. S. Lee: An optimal algorithm to compute the max-mintransitive closure of a fuzzy similarity. *Fuzzy Sets and Systems* 123, 129-136, 2001.

[6]. J. Recasens: *Indistinguishability operators: Modelling Fuzy Equallities and Fuzzy Equivalence relations.* Studies in Fuzziness and Soft Computing. Springer. 2010.

[7]. E. Trillas: Assaig sobre les relacions d'indistingibilitat. *Proc. Primer Congrés Català de Lògica Matemàtica*, Barcelona, 51-59, 1982

[8]. E. Trillas, L. Valverde: An inquiry into indistinguishability operators. *Aspects of vagueness,* 231-256, Theory Decis. Lib., 39, Reidel, Dordrecht, 1984.

[9]. L. Valverde: On the structure of F-indistinguishability Operators. *Fuzzy Sets and Systems* 17, 313-328, 1985.

[10]. L. A. Zadeh, Similarity relations and fuzzy orderings, *Inform. Sci.* 3, 177–200, 1971.

"ROUGH SETS" Y "SOFT COMPUTING"

Rafael Bello [1], José Luis Verdegay [2]

[1] *Departamento de Ciencias de la Computación, Universidad Central de Las Villas, Cuba, rbellop@uclv.edu.cu*

[2] *Departamento de Ciencias de la Computación e I.A. -**CITIC**- Universidad de Granada, España, verdegay@ugr.es*

Resumen

La solución de problemas en presencia de incertidumbre es un tema relevante para lograr sistemas computacionales eficaces. Se han desarrollado diversas técnicas para considerar la incertidumbre, su desarrollo ha llevado a la consolidación de la Soft Computing como un área de alto interés. Por otra parte, la Teoría de los "Rough Sets" se ha convertido en una vía eficaz para tratar la incertidumbre, especialmente cuando esta se presenta por inconsistencia en los datos. En este artículo se presenta un análisis de la relación existente entre los conjuntos aproximados y algunas componentes de la Soft Computing, y como esta hibridación permite mejorar el desempeño de los sistemas.

Palabras Clave: Rough Sets, Soft Computing.

1 INTRODUCCIÓN

La información imperfecta es omnipresente. Casi toda la información disponible es incierta, incompleta o imprecisa. A nivel empírico, la incertidumbre es una compañía inseparable de casi cualquier medida, resultante de una combinación de inevitables errores de medición y límites en la resolución de los instrumentos. A nivel cognitivo emerge de la vaguedad y ambigüedad inherente al lenguaje natural. Por eso los sistemas de información tienen que ser capaces de considerar la incertidumbre. Esto es particularmente importante en el caso de la Inteligencia Artificial (IA), donde es necesario disponer de mecanismos de inferencia que permitan extraer conclusiones sugeridas pero no aseguradas por sus premisas; en estos casos el sistema puede tener información que es relevante a una conclusión y que hace razonable presumir que la conclusión es verdadera. De ahí que la modelación de la incertidumbre sea de especial interés en el campo de la IA.

La naturaleza multidimensional de la incertidumbre se oculta cuando la incertidumbre se concibe sólo en términos de la teoría de las probabilidades, en la cual se manifiesta sólo una de sus dimensiones. Al reconocerse su carácter diverso varios autores han establecido diferentes tipos de incertidumbre. Bonnisone y Tong [Bon85] clasifican la imperfección del conocimiento en incertidumbre, imprecisión e información incompleta. Este último término se utiliza para indicar la ausencia de un valor, el término imprecisión denota la existencia de un valor el cual no puede ser medido con la precisión adecuada, e incertidumbre indica el hecho de que un agente ha formulado una opinión subjetiva sobre la veracidad de un hecho que no conoce con certeza. Bosc y Prade [Bos93] añaden dos nuevos términos, vaguedad e inconsistencia.

La vaguedad es una propiedad de los conjuntos (conceptos) y puede ser atribuida a los límites del conjunto, mientras que la inconsistencia describe una situación en la cual hay dos o más valores en conflicto para ser asignados a una variable.

La necesidad de considerar la incertidumbre en el proceso de solución de problemas, teniendo en cuenta que esta puede estar presente tanto en los datos de la instancia del problema que se resuelve, como en el conocimiento que se utiliza para resolver esa clase de problemas, ha dado lugar al desarrollo de diversos métodos de solución de problemas que pretenden ser robustos. El proceso de solución de problemas en estas condiciones se conoce como razonamiento bajo incertidumbre, razonamiento aproximado, o razonamiento inexacto. En este contexto es en el que se desarrolla la Soft Computing (SC), punto más avanzado de ese camino en la solución de problemas en un ambiente de incertidumbre.

Zadeh propuso la primera definición de SC en 1994 [Zad94], aunque la idea de establecer el área de SC apareció en el 1990 [Zad01]. Hasta ese momento, los conceptos incluidos en SC se trataban de forma aislada. La definición propuesta por Zadeh fue la siguiente:

Básicamente, SC no es un cuerpo homogéneo de conceptos y técnicas. Mas bien es un grupo de métodos que de alguna manera tienen el propósito de explotar la tolerancia a la imprecisión y la incertidumbre para alcanzar tratabilidad, robustez y soluciones de bajo costo. Las principales componentes de la SC son la lógica borrosa, la neurocomputación y el razonamiento probabilístico, con la posterior integración de los Algoritmos Genéticos (AG),

las redes de creencias, los sistemas caóticos y partes de la teoría del aprendizaje. De esta forma, Zadeh definió la SC por extensión, por medio de diferentes técnicas y conceptos desarrollados para enfrentar problemas que tengan elementos de incertidumbre en su dominio de definición.

Mas recientemente, Verdegay et al. [Ver08] dan una definición de SC considerándola como una antítesis de lo que denominaría "hard computing": SC puede por eso ser vista como una serie de técnicas y métodos con las que manejar las situaciones prácticas reales en la misma manera en que los humanos tratan con ellas, es decir, en base a inteligencia, sentido común, consideración de analogías, aproximaciones, etc. En este sentido, Soft Computing es una familia de métodos de solución de problemas encabezados por el Razonamiento Aproximado y los Métodos de Aproximación Funcional y de Optimización incluyendo los métodos de búsqueda. Por eso, la SC está en la base teórica del área de los sistemas inteligentes.

Se puede apreciar una relación entre los tipos de incertidumbre antes mencionados (incertidumbre, imprecisión, información incompleta, vaguedad, inconsistencia) y las diferentes componentes de la SC. Los conjuntos borrosos y la lógica borrosa para modelar la vaguedad, los modelos probabilísticos (Teoría de Dempster–Shafer y Redes Bayesianas) para la incertidumbre e imprecisión. Como resulta natural que en un mismo problema aparezcan diferentes tipos de incertidumbre, entonces procede la hibridación de estas técnicas (por ejemplo, lógica borrosa con redes neuronales y algoritmos evolutivos). De modo que lo más importante en la SC es que las metodologías que la integran deben ser vistas como el resultado de la cooperación, asociación, complementariedad o hibridación de las componentes antes señaladas [Ver08].

Este articulo pretende ubicar los "Rough Sets" (RS) en este contexto. Para ello se realiza un análisis desde dos perspectivas: como una vía para modelar la incertidumbre cuando esta está dada por vaguedad o inconsistencia en la información, y analizando diversas formas de hibridación de los RS con dos de las componentes antes mencionadas de la SC. Por limitaciones de espacio el análisis es mas descriptivo que crítico, pero suficientemente informativo para los investigadores interesados en el tema. En [Bel12] se puede encontrar una versión mucho mas extensa de este trabajo.

2 ANÁLISIS DE LA INCONSISTENCIA DE LA INFORMACIÓN MEDIANTE RS

En 1982 [Paw82] Z. Pawlak presentó los RS inferior y superiormente, abriendo una nueva dirección en el desarrollo de teorías sobre la información incompleta. Según L. A. Zadeh, [Orl98], la Teoría de los RS (RST) ha evolucionado hasta convertirse en una metodología para enfrentar una amplia variedad de problemáticas, entre ellas la incertidumbre presente en la información.

De este modo se agrega al conjunto de teorías que están orientadas a modelar el problema de la incertidumbre. La RST es la mejor herramienta para modelar la incertidumbre cuando esta se manifiesta como inconsistencia. A la vez se relaciona con otros miembros de ese conglomerado como la Teoría de los conjuntos borrosos y la Teoría de evidencia [Paw95a], y [Paw95b]. En [Won01] se analiza el empleo de los RS para razonar en condiciones de incertidumbre, combinándolos con un enfoque probabilístico. También, la integración de los RS con las redes neuronales artificiales ha tenido una creciente atención como vía para construir sistemas más eficientes e inteligentes, como se muestra, entre otros, en [Jel94], [Lin96], [Ngu97] y [Swi98].

La filosofía de los RS se basa en la suposición que alguna información es asociada con cada objeto del universo de discurso [Paw99]. En esta sección se definen los conceptos básicos de la RST: sistema de información y aproximación de conjuntos. El punto de partida del análisis de datos basado en RS es el sistema de información [Paw02]. Cuando la información se representa mediante una tabla donde cada fila representa un objeto y cada columna representa un rasgo esta tabla es llamada Sistema de Información; más formalmente, es un par S = (U, A), donde U es un conjunto finito no vacío de objetos llamado Universo y A es un conjunto finito no vacío de atributos. Un Sistema de Decisión es cualquier sistema de información de la forma DS = (U, A$\cup${d}), donde $d \notin A$ es el atributo de decisión. Empleando este sistema de Información (o sistema de Decisión) y una relación de inseparabilidad (indiscernibility relation) se pueden establecer los conceptos básicos de RST, la aproximación inferior y superior de un concepto. La definición clásica de aproximación inferior y superior fue originalmente introducida con referencia a una relación de inseparabilidad la cual es una relación de equivalencia.

Sea DS = (U, A$\cup${d}) un sistema de decisión y B$\subseteq$A y X$\subseteq$U. B define una relación de equivalencia y el subconjunto X es un concepto en el universo U. X se puede aproximar usando solo la información contenida en B mediante la construcción de las aproximaciones B-inferior y B-superior denotadas por B_*X y B^*X respectivamente. A partir de estas aproximaciones la RST ofrece una serie de medidas que permiten medir la incertidumbre presente en la información, entre ellas, precisión de la aproximación y calidad de la aproximación, que dan una medida de la vaguedad del concepto que se aproxima.

Usando el enfoque de los RS se puede analizar uno de los tipos de incertidumbre de la información, en este caso la consistencia del sistema de decisión. Cuando dos objetos inseparables pertenecen a clases de decisión diferentes se dice que el sistema de decisión es *inconsistente*, en otro caso es *consistente*. También se conocen como sistemas *no deterministas* o *deterministas*, respectivamente. Si la calidad de la clasificación es igual a 1 el sistema de decisión es consistente, lo que significa que no existe un par

de objetos en *U* que son inseparables y pertenecen a clases distintas; si es menor que 1 es inconsistente y ese valor indica el grado de consistencia.

Esta medida de consistencia del sistema de decisión ha sido empleada como valor heurístico en procesos de selección de rasgos basados en metaheurísticas, como se analiza en [Zho01], [Jen03], [Wan06], [Wan07], [Cab07], [Bel08] y [Gom08]. Otro interesante estudio sobre el empleo de esta medida en el campo de la complejidad de datos se muestra en [Cab08] y [Cab10], donde se establece una relación entre grado de consistencia de un sistema de decisión y la calidad del resultado de un proceso de descubrimiento de conocimiento, en particular en la construcción de clasificadores, que lo utiliza como conjunto de aprendizaje. Lo que muestra la utilidad de los RS en los procesos de aprendizaje.

3 COMBINACIÓN DE LOS RS CON LA SOFT COMPUTING.

Como se ha planteado antes, una característica relevante de la SC es la combinación de sus componentes para producir modelos computacionales híbridos más eficientes o eficaces. En este empeño de crear sistemas híbridos con un mejor desempeño los RS han sido combinados con diversas componentes de la SC. Esto lo sostiene S. Pal en [Pal04], cuando plantea que la RST y la computación granular han probado ser otra herramienta de la SC que, en varias alternativas de sinergia con la lógica borrosa, las Redes Neuronales Artificiales (RNA) y los AG, ofrecen un marco más fuerte para alcanzar tratabilidad, robustez, soluciones de bajo costo, acercándose muy bien a la forma de decidir de los humanos.

Li et al [Li07] plantean que la lógica borrosa, las RNA y el razonamiento probabilístico fueron las componentes iniciales de la SC, a las que después se les añadieron otras como los AG, los sistemas caóticos, y los RS. Analizan la necesidad de la complementariedad de estas componentes y por tanto del desarrollo de sistemas híbridos basados en ella de la forma siguiente. Cada metodología de la SC tiene sus propias fortalezas y ofrece diferentes ventajas. Así, los conjuntos borrosos son apropiados en la modelación del razonamiento humano y ofrecen un mecanismo natural para tratar la incertidumbre.

Las RNA son robustas al ruido y tienen una alta capacidad para modelar relaciones altamente no lineales. Los AG son particularmente útiles para búsqueda óptima. Los RS son muy eficientes para reducción de atributos y extracción de reglas. Por otra parte, las técnicas de SC tienen restricciones que no permiten su aplicación individual en algunos casos. Los conjuntos borrosos son dependientes del conocimiento experto. El tiempo de entrenamiento de las RNA es excesivo y tedioso cuando hay muchas entradas y la mayoría de los modelos de redes fallan en las facilidades de explicación. Los AG pueden tener problemas de convergencia. Los RS tienen problemas NP en la reducción de atributos y descubrimiento de reglas óptimas. Para superar estas limitaciones la respuesta es la hibridación. En [Li07] también se presenta una revisión de los sistemas híbridos que integran RS y técnicas de SC en el campo de la minería de datos, analizando el papel de los RS en los sistemas híbridos basados en SC en tres líneas: preprocesamiento de los datos, cálculo de las medidas de incertidumbre y descubrimiento de conocimiento.

En esta sección se presentan varios ejemplos de estas combinaciones que ilustran como los RS se integran de forma natural con otras componentes de la SC con el propósito de generar modelos computacionales más flexibles y robustos. Primero se muestra como los conjuntos borrosos y los RS se han integrado, y luego las diferentes maneras en que los RS se han introducido en la computación neuro-borrosa.

3.1. RS Y CONJUNTOS BORROSOS

La similaridad entre los términos conjuntos borrosos y RS pueden crear confusión. Ambos tienen que ver con la aproximación de conceptos; mientras que la teoría de los conjuntos borrosos caracteriza un concepto vago por una función de pertenencia con un rango entre cero y uno, la RST ofrece una aproximación inferior y otra superior dependiendo de una relación. La RST y la Teoría de los conjuntos borrosos son generalizaciones de la teoría clásica de conjuntos para modelar la vaguedad y la incertidumbre [Yao98c]. Ambas teorías modelan diferentes tipos de incertidumbre, mientras la RST considera la inseparabilidad entre los objetos típicamente caracterizada por una relación de equivalencia, la FST supone una mala definición de los limites de una clase.

Formalmente un conjunto borroso A definido sobre un universo X es un par de la forma $(x, \varphi A(x))$. Donde $\varphi A(x)$ se llama función de pertenencia del conjunto A y asigna a cada elemento de X un grado de pertenencia en el intervalo [0,1]. X se llama universo de discurso y puede ser un espacio discreto o continuo. Más específicamente, un conjunto borroso es una clase con límites no rigurosamente definidos mientras que un RS es un conjunto duro, con límites bien establecidos. Ambas teorías han evolucionado en direcciones diferentes y actualmente se complementan, en lugar de competir (L.A. Zadeh en [Orl98]).

Existen diversas formas de integrar ambos modelos. Los RS permiten aproximar conceptos definidos mediante conjuntos duros. Pero cuando el concepto no es representable de esta forma, se llega a una situación donde los RS y los conjuntos borrosos pueden ser integrados. Los métodos de los RS ofrecen descripciones aproximadas de conceptos y pueden ser usados para construir descripciones aproximadas de conceptos borrosos. Combinar los RS y los conjuntos borrosos permite obtener aproximaciones de conjuntos borrosos, así como aproximaciones de con-

juntos mediante relaciones de similaridad borrosas, lo que es particularmente importante en el caso de que el dominio de los rasgos es continuo (real), en cuyo caso el enfoque clásico de los RS basados en relaciones de equivalencias resulta inadecuado.

Las definiciones previas de RS se basan en una relación *B, B⊆A,* que define como inseparables a pares de objetos que tienen igual valor para un subconjunto de rasgos, es decir, los objetos (x,y) son inseparables según B si $a_i(x) = a_i(y)$, para todo rasgo i en B, donde $a_i(x)$ denota el valor del rasgo i en el objeto x. Definida de esta forma, R es una relación de equivalencia. Pero cuando el dominio de los rasgos en B no es discreto, una relación de inseparabilidad definida de esta forma no es aplicable. La relación de equivalencia es muy estricta en caso de dominios continuos pues ligeras diferencias entre los valores de los objetos para un rasgo pueden no ser significativas al analizar su inseparabilidad. Esto es muy importante en el caso de rasgos numéricos en los cuales pequeños errores en la medición de los mismos pueden generar estas diferencias. En estos casos es mas adecuado usar una relación de semejanza en lugar de una de equivalencia.

Es necesario usar otros tipos de relaciones de inseparabilidad entre los objetos del universo U. Esto se logra extendiendo el concepto de inseparabilidad entre objetos de modo que se agrupen en la misma clase los objetos similares, no idénticos, según una relación de semejanza R. Las relaciones de semejanza no inducen una partición del universo U, sino generan clases de semejanza para cualquier objeto $x \in U$. Remplazando la relación de equivalencia por una relación binaria más débil, se obtiene una extensión del enfoque clásico de la RST. Algunas extensiones se estudian en [Yao98a, Slo00, Sko05, Paw07].

En [Dub90] se analizaron los conjuntos borrosos y los RS concluyendo que no son teorías rivales, sino dos diferentes herramientas matemáticas con propósitos diferentes, y se logró combinar de una forma natural y útil ambos tipos de conjuntos definiendo los conjuntos "rough-fuzzy sets" y "fuzzy-rough sets". En [Dub90] y [Dub92] y [Bha05], se propuso el concepto de "fuzzy-rough set" remplazando las relaciones binarias duras por relaciones borrosas sobre el universo.

En [Wyg89] los RS se expresan mediante FP de la forma $\varphi \rightarrow \{0, 0.5, 1\}$, para representar las regiones negativa, límite y positiva respectivamente. Todos los objetos del universo que pertenecen a la aproximación inferior tienen valor de pertenencia 1, los que pertenecen al conjunto $BN_B(X)$ tienen valor 0.5, y el resto de los objetos 0.

Un interesante estudio sobre la cantidad de "vaguedad" (fuzziness) presente en un RS se presenta en [Cha00]. Si X es un RS, entonces existe alguna cantidad de "grado de aproximación" (roughness) debido a la frontera aproximada no vacía. Esto se puede ver en el sentido de que todos los miembros del universo son de alguna forma miembros del conjunto X. En [Paw03] se analizan las propiedades de esta medida y como la misma es una generalización de la función de pertenencia borrosa.

Otra idea que relaciona los conjuntos borrosos y los RS se muestra en [Qua00]. Se presenta una generalización de la RST llamada α-RST. Todos los conceptos de la RST son extendidos usando un enfoque borroso y está dirigida al caso de datos reales imprecisos y con ruidos. En lugar de la relación de inseparabilidad clásica se introduce una relación parametrizada, notada como $IND(R,\alpha)$ o de forma abreviada R_α. Usando esta relación, dos elementos son inseparables si y solo si tienen los mismos valores para todos los rasgos y su grado de posibilidad es mayor que un umbral de similaridad α.

En [Yao04], se formula un marco semántico para conjuntos borrosos dentro de la RST, en el que las funciones de pertenencia aproximadas se ven como un tipo especial de función de pertenencia borrosa, también interpretable usando probabilidades condicionales; las aproximaciones de los RS se relacionan con los conceptos de núcleo y soporte de un conjunto borroso. De modo que la RST ofrece una vía especifica y concreta de construir las funciones de pertenencia de los conjuntos borrosos.

Una medida sobre el grado de precisión de los conjuntos borrosos se introduce en [Ban96] y [Zha04]. También en [Won92] los autores demostraron que una noción generalizada de RS se puede describir convenientemente por el concepto de conjunto borroso empleando operaciones apropiadas sobre conjuntos borrosos. La combinación de los conjuntos borrosos con el modelo VPRS (Variable Precisión Rough Set) fue estudiada en [Hon05] y [Tsa05], el modelo VPRS es una versión generalizada del enfoque convencional de la RST, en el que se permiten ciertas imprecisiones en la información [Zia93]. Otras alternativas de hibridación de los RS y los borrosos se pueden encontrar en [Bea99, Jen04, She04, Bha05, Qin05, Tha05]

3.2. RS Y REDES NEURONALES ARTIFICIALES

La combinación de los RS y las Redes neuronales artificiales ha tenido diferentes propósitos, entre ellos crear nuevos modelos de neuronas, preprocesar los datos de entradas a la red neuronal, ayudar al diseño de la red neuronal etc. Como ejemplo del último caso se pueden presentar los trabajos descritos en [Hun97] y [Szc98]. En los trabajos de [Yas95], [Has98] y [Ahn00] los RS combinados con las redes neuronales artificiales se han usado para resolver problemas de predicción en economía y finanzas, particularmente los RS se usaron en el preprocesamiento del conjunto de entrenamiento de la red neuronal, pues permitieron reducir el conjunto de atributos, y eliminar los objetos inconsistentes lo que mejoró el desempeño de la red y redujo el tiempo de entrenamiento. Los resultados alcanzados fueron superiores a los obtenidos usando análisis discriminante y modelos de redes neuronales artificiales [Tay02].

Un ejemplo de RNA aproximada [Lin96] consiste en combinar en una red neuronas aproximadas y neuronas convencionales. Las neuronas aproximadas se definen como un par, que representan los valores inferior y superior, estos valores representan cotas permitiendo un rango de tolerancia. Una neurona aproximada (rough neuron) r se puede ver como un par de neuronas, r* y r*. Su entrada, al igual que en las neuronas convencionales se calcula a partir de las entradas que recibe desde todas las neuronas que contribuyen a ella. Las salidas de las neuronas convencionales y aproximadas se diferencian, y quedan establecidas por las expresiones siguientes. Para el caso de las neuronas convencionales la salida de esta se calcula a partir de su entrada total y el modelo de neurona que se utilice (lineal, lineal con umbral, sigmoide, etc.). Para las neuronas aproximadas se definen dos salidas, una superior y otra inferior. La topología de la red incluye las conexiones entre neuronas convencionales o entre estas y las aproximadas se establecen como normalmente ocurre. En el caso de las neuronas aproximadas estas se pueden conectar entre si en tres variantes: completamente, excitante o inhibitoriamente.

Otra alternativa para combinar ambos modelos es usar técnicas basadas en los RS para procesar la entrada a la RNA, lo cual puede influir sustancialmente en el costo computacional de implementar la red neuronal. Se han desarrollado diversos modelos computacionales basados en RS para realizar el preprocesamiento de conjuntos de entrenamientos, en particular mediante la selección de rasgos y la edición de los conjuntos.

En [Szc98] se presenta un amplio estudio de los efectos de aplicar este procedimiento sobre nueve conjuntos de entrenamiento de dominios diferentes. En todos los casos se aprecian decrecimientos significativos tanto en el tamaño de la entrada a la red como en el tamaño de la red (cantidad de neuronas), lo que provoca que la red sea más manejable computacionalmente, sin provocar pérdida de eficacia de la red.

Por su parte en [Mau96] se presenta un estudio comparativo sobre la efectividad de aplicar la técnica estadística conocida como Análisis de Componentes Principales y la RST para realizar el preprocesamiento en el entrenamiento de una RNA; en el cual los experimentos realizados demuestran que ambos son similarmente competentes para esta tarea, y que efectivamente los RS pueden ser una alternativa interesante al análisis por componentes principales para reducir el tiempo de aprendizaje y mejorar el aprendizaje de una red neuronal.

En [Jel95] se plantea que la reducción de los datos basada en RS coopera efectivamente con el entrenamiento de redes neuronales mediante el algoritmo de retropropagación de los errores, los experimentos computacionales realizados muestran que es posible reducir significativamente (incluso varias veces) el tiempo de aprendizaje, manteniendo aproximadamente estable el error que se comete. También Pawlak [Paw95] encontró que la reducción de los datos usando RS mejoró 4.72 veces el tiempo de aprendizaje de una RNA.

Otro aspecto que se ha estudiado es como usar métodos basados en RS para diseñar redes neuronales [Yas95], [Qin00], [Pal03] y [Tia06]. La determinación de la topología y otros parámetros de una RNA es una tarea difícil en la mayoría de los modelos de estas. Se ha investigado el empleo de los conjuntos borrosos para ayudar en esta tarea, incluyendo la determinación del numero de capas ocultas, la cantidad de nodos en las capas ocultas y el conjunto inicial de pesos de la red; para ello lo más usual es descubrir reglas usando métodos basados en RS a partir de la misma información que será usada para entrenar la red. También se ha analizado el ajuste de la arquitectura de una RNA usando los RS [Yon99] y [Has02] . El efecto de este enfoque híbrido es acortar el tiempo de entrenamiento y lograr una mayor capacidad de generalización [Li04].

Un ejemplo de lo antes descrito son los estudios presentados en [Ban98] y [Pal03], que muestran la integración de varias componentes de la SC para obtener un modelo de RNA más eficiente y efectivo en problemas complejos, en el cual se combinan una red neuronal artificial tipo Perceptrón Multicapa (MLP), los conjuntos borrosos, los RS y los AG. La RNA es el soporte o estructura del modelo, los conjuntos borrosos ayudan a modelar las entradas y salidas de la red, los RS sirven para extraer el conocimiento del dominio de aplicación y a partir de este conocimiento diseñar una red inicial cercana a la red óptima, lo cual reduce el espacio de búsqueda. Los AG se emplean para ajustar los elementos de la red, entre ellos los parámetros de las funciones de pertenencias de los términos lingüísticos de cada rasgo; así como los pesos asociados a cada conexión en la red.

El trabajo desarrollado por [Yah00] se enmarca en el campo de los sistemas híbridos inteligentes en los que diferentes técnicas inteligentes como la lógica borrosa, las redes neuronales, los AG y los sistemas expertos se combinan para superar las limitaciones de cada técnica individualmente. En su trabajo proponen un sistema híbrido que combina las RNA y los RS para complementar la tecnología de los sistemas expertos. La RNA es utilizada en el proceso de adquisición del conocimiento, apoyando la construcción de la base de conocimientos del sistema experto, los RS sirven para el preprocesamiento de la información usada en el entrenamiento de la RNA y para refinar el proceso de inferencia del sistema experto.

Uno de los más recientes trabajos en que se muestra la combinación de RS y RNA se presenta en [Chu10], para implementar una aplicación en el campo del diseño de productos. En este caso se utiliza como modelo de red un SOM (Self-organizing maps). Los RS se usan en la generación de reglas de configuración de productos que ayudan a los diseñadores de estos a partir de sistemas de información conteniendo configuraciones incompletas; la RST y su aplicación en el descubrimiento de conocimien-

to en el caso de sistemas de información incompletos tiene sus propias particularidades como se muestra en [Kry98], [Gre01], [Hon02] y [Grz04]; el SOM es empleado para analizar las variables de diseño y optimizar una función objetivo. Ambas componentes se insertan en un sistema experto de diseño de configuraciones.

Describimos a continuación otros trabajos que abordan la combinación de RS y RNA. Un modelo similar al anterior fue propuesto por Mitra et al. [Mit01], en el cual se integran cuatro técnicas de la SC (RNA, conjuntos borrosos, AG y RS) para diseñar una red basada en el conocimiento para la generación de reglas [Tha09]. Li y Wang [Li04] presentaron un enfoque híbrido donde integran los RS y las RNA para descubrir reglas de clasificación a partir de grandes bases de datos, obteniendo reglas más precisas y concisas. Wang et al. [Wan05] usaron los RS en la etapa de preprocesamiento de los datos para entrenar RNA. En [Che10] se muestra la combinación de RS y AG en la formulación de un modelo híbrido para el pronóstico de precios; basado en la RST se realiza la selección de indicadores y la extracción de reglas las cuales son refinadas usando AG para incrementar la precisión del pronóstico, en particular para lograr la predicción de eventos inesperados, lo cual se logra mediante las mutaciones que se realizan por el AG sobre las reglas descubiertas usando RS; en este trabajo se destaca el hecho de la ventaja de usar RS pues la mayoría de los métodos estadísticos establecen condiciones sobre las variables usadas en el análisis, lo que hace que no puedan ser aplicados a bases de datos que no sigan esas distribuciones estadísticas.

Conclusiones

En este trabajo se ha realizado una revisión del lugar de los RS en la SC, considerando que esta se puede ver como un conjunto de métodos computacionales orientados a ser más eficiente y eficaz la solución de problemas en la presencia de incertidumbre. El mismo comprende dos aspectos principales. Se analizan los RS como un modelo que permite tratar la incertidumbre cuando la misma se genera por inconsistencias en la información, así como diferentes formas de hibridar los RS con algunas componentes de la SC.

El estudio muestra como los RS permiten determinar la inconsistencia en la información a partir del empleo de relaciones de inseparabilidad y medir el grado de inconsistencia presente en un conjunto de datos. Por otra parte, se integra perfectamente con otras componentes de la SC por la complementariedad de las mismas, para lograr sistemas híbridos que son más flexibles y robustos. El empleo de los RS en la solución de problemas de aprendizaje ha sido especialmente relevante, entre ellos la reducción de información y el descubrimiento de conocimiento. Por todo lo anterior se puede concluir que los RS se pueden situar sin dudas en el contexto de la SC.

Agradecimientos

Este trabajo ha sido realizado con la ayuda de los proyectos TIC-02970 (Junta de Andalucía), TIN2008-06872-C04-04/TIN y TIN2011-27696-C02-01 (ASCETAS) del Ministerio de Ciencia e Innovación y el proyecto ANIMO-CHEVERE del Programa Erasmus Mundus (Comisión Europea).

Referencias

[Ahn00] Ahn, B.S. et al.. ESWA 18, 65-74. 2000.

[Ban96] Banerjee, M. y Pal, S.K.. Inf. Comp. Sci, 235–246. 1996.

[Ban98] Banerjee, M. et al. IEEE Trans. on NN, 9, 6, 1203-1216. 1998.

[Bea99] Beaubouef, T. et al.. En Pal, S.K. and Skowron, A. (Eds). Springer-Verlag, 1999.

[Bel08] Bello, R. et al. LNCS, 348-355. 2008.

[Bel12] Bello, R y J.L. Verdegay Inf. Sci (en prensa)

[Bha05] Bhatt, R. y Gopal, M.. PRL 26, 1632-1640. 2005.

[Bon85] Bonnissone, P.P. y Tong, R.M. Int. J. Man Machine Studies, 22, 241-250, 1985.

[Bos93] Bosc, P. y Prade, H. Proc II Work. UMIS 1993

[Cab07] Caballero, Y. et al. Proc of ISDA07, 407-411.

[Cab08] Caballero, Y., et al.. Stud. in Comp Intel 224

[Cab10] Caballero, Y., et al. Stud in Comp Intel 262.

[Cha00] Chakrabarty, K. et al.. FSS 110, 247-251. 2000.

[Che10] Cheng, C.H., et al.. Inf Sci 180, 1610–1629 2010

[Chu10] Chu, X.Z. et al. ESWA 37, 7364–7372. 2010.

[Dub90] Dubois, D. y Prade, H.. IJGS 17 191-208. 1990.

[Dub92] Dubois, D. y Prade, H.. Intelligent Decision Support, Slowinski, R. (Ed.). Kluwer Academic, 1992.

[Gom08] Gómez, Y. et al. JUCS 14, 15, 2582-2596, 2008

[Gre01] Greco, S. et al.. EJOR 129, 1-47, 2001.

[Grz04] Grzymala-Busse, J.W. y Siddhaye, S. Proc IPMU 2004, 2, 923-930, 2004.

[Has98] Hashemi, R. et al.. EJOR 109 (2), 390-402. 1998.

[Has02] Hassan, Y., y Tazaki, E.: Int. J. of Neural System 12, 435–446. 2002.

[Hon02] Hong, T.P. et al. ESWA 22, 285-293. 2002.

[Hon05] Hong, T-P. et al. ESWA, Elsevier Ltd., 2005.

[Hun97] Hung, S.N. et al.. LNAI 1263, 359-366. 1997.

[Jel94] Jelonek, J. et al.. Ziarko, W. (Ed.) Springer 1994.

[Jel95] Jelonek, J. et al. Comp. Intel. 11(2), 339-347 1995

[Jen03] Jensen, R. y Shen. Q. Procs UK Workshop on Comp. Intel.. 15-22. 2003.

[Jen04] Jensen, R. y Shen. Q. FSS 141, 469-485. 2004.

[Jen05] Jensen, J. y Shen, Q.. FSS 149(1): 5-20 (2005)

[Kry98] Kryszkiewicz, M.. Inf. Sci, 112, 39-49, 1998.

[Li04] Li, R. y Wang, Z-O.. EJOR.439–448. 2004.

[Li07] Li, R. et al.. LNAI. 4632, pp. 35 – 44. 2007.

[Lin96] Lingras, P. In Proc. of the VI IPMU'96).

[Mau96] Maudal, O. Report, Dept. of Artificial Intelligence, University of Edinburgh. 1996.

[Mit01] S. Mitra, S. et al. Neurocomp. 36, 45–66. 2001.

[Ngu97] Nguyen, H.S. et al.. LNAI 1263 (1997).

[Orl98] Orlowska, E. (Ed.).. Rough sets analysis. Physica-Verlag. 1998.

[Pal03] Pal, S.K. et al. IEEE Trans KDE 15, 14-25. 2003.

[Pal04] Pal, S.K. Information Sciences 163, 5–12. 2004.

[Paw82] Pawlak, Z.. Int'l J. Inf Comp Sci 11, 341-356, 1982

[Paw95a] Pawlak, Z. et al.. Comm. of ACM, 38(11), 89-95, 1995.

[Paw95b] Pawlak, Z.. Comp. Intel. 11, no. 2, 1995.

[Paw99] Pawlak, Z. In Pal, S.K. and Skowron, A. (Eds). Springer Verlag, 1999.

[Paw02] Pawlak, Z.. EJOR 136, 181-189. 2002.

[Paw03] Pawlak, Z.. In B. Bouchon-Meunier et al. (Eds), Elsevier Science. 2003.

[Paw07] Pawlak, Z. y Skowron, Inf. Sci 177, 28-40. 2007.

[Qin00] Qin, Z., Mao, Z.: Proc 3rd World Congress on Intel. Control and Aut 795–799. 2000.

[Qin05] Qin, K. y Pei, Z.. FSS 151, 601-613. 2005.

[Qua00] Quafafou, M.. Inf. Sci 124, 301-316. 2000.

[She04] Shen. Q. and Jensen, R.. Pattern Rec 37, 1351-1363. 2004.

[Sko05] Skowron, A. et al. LNCS 3400, 175-189. 2005.

[Slo00] Slowinski, R. and Vanderpooten, D. IEEE Trans on DKE 12 (2) 331-336. 2000.

[Swi98] Swiniarski, R.. LNAI 1424 (1998)

[Szc98] Szczuka, M.S.. In Polkowski and Skowron, A. (Eds.), Physica-Verlag, 449-470. 1998.

[Tay02] Tay, F.E. y Shen, L.. EJOR 141, 641-659. 2002.

[Tha05] Thangavel, K. y Pethalakshmi, A.., Int'l J. AIML 6 (1), 11–17. 2005.

[Tha09] Thangavel, K. y Pethalakshmi, A. Applied Soft Computing 9 , 1–12. 2009.

[Tia06] Tian, K., Wang, J., et al.: Chinese Control Conference, 834–838. IEEE Computer Society Press, 2006.

[Tsa05] Tsai, Y-C. et al. ESWA. 1–8. 2005.

[Ver08] J. L. Verdegay et al. FSS, 2008.

[Wan05] Wang, Y. et al. Int'l J. Comp. Cog. 3 (1). 2005.

[Wan06] Wang, X. et al. Comp. Methods and Programs in Biomedicine 83, 147-156. 2006.

[Wan07] Wang, X. et al.. PRL 28, 459-471. 2007.

[Won01] Wong, S.K.M. and Butz, C.J.. LNCS 2005, 511-519, 2001.

[Won92] Wong, SKM., Ziarko, W. FSS 21, 357-362 1992

[Wyg89] Wygralak, M.. FSS 29, 241-243, 1989.

[Yah00] Yahia, M.E. et al.. ESWA 18, 87–99. 2000.

[Yao98a] Yao, Y.Y.. Inf. Sci., 111, 239-259, 1998.

[Yao98c] Yao, Y.Y.. J. Inf. Sci 109, 227-242. 1998.

[Yao04] Yao, Y.Y. LNCS Trans. on RS II, 310-331. 2004

[Yao08] Yao, Y.Y.. IJAR 49, 255–271. 2008.

[Yas95] Yasdi, R.. Neurocomputing 7, 61-84. 1995.

[Yon99] Yon, J.H., et al. FUZZ IEEE 1666–1670, 1999.

[Zad65] Zadeh, L.A.. Inf. & Control 8, 338-353. 1965.

[Zad94] Zadeh, L.A.. Proc. IIZUKA'94, 1–2. 1994.

[Zad01] Zadeh, L.A.. Applied Soft Comp 1(1), 1–2. 2001.

[Zia93] Ziarko, W.. J. Comp and Sys Sci 46, 39-59. 1993.

[Zha04] Zhang, H. et al.. Inf. Sci 166,147–165. 2004.

[Zho01] Zhong, N. et al. J. Intel Inf Sys 16 199-214, 2001

UNA TIPOLOGÍA CONJUNTA DE LA BIPOLARIDAD Y LA INCERTIDUMBRE BORROSA

J. Tinguaro Rodríguez[1], Camilo A. Franco[1], Javier Montero[1]

[1] *Departamento de Estadística e Investigación Operativa, Facultad de Matemáticas, Universidad Complutense de Madrid*
jtrodrig@mat.ucm.es

Resumen

El objetivo de este trabajo es presentar una metodología que permita clarificar las relaciones existentes entre las nociones de incertidumbre borrosa (*fuzziness*) y bipolaridad. Esta metodología está basada en el uso de extensiones continuas 4-valoradas sobre pares de evidencia positiva y negativa (o bipolares). Estas extensiones permiten definir rigurosamente la semántica, en términos cognitivos o epistémicos, de algunos formalismos habituales, diferentes a la luz de este método y que, sin embargo, suelen considerarse equivalentes .

Palabras Clave: Bipolaridad, incertidumbre borrosa o lingüística, intuicionismo

1 INTRODUCCIÓN

El desarrollo de la inteligencia artificial durante la segunda mitad del siglo pasado ha conllevado un notable esfuerzo científico y tecnológico orientado a replicar y plasmar en procesos automáticos (i.e., informáticos) algunas de las capacidades y habilidades características de la mente humana.

En tanto que la habilidad para el manejo de lenguajes complejos es quizás la faceta esencial de la naturaleza humana, el tratamiento del lenguaje natural ha sido (y es) uno de los retos fundamentales a enfrentar dentro de este contexto.

La aparición de las nociones de conjunto y lógica *borrosos* (*fuzzy set* y *fuzzy logic*, [19]) es sin duda uno de los hitos claves en relación con el tratamiento informático del lenguaje natural, en tanto que permiten una modelización matemática rigurosa de la imprecisión y ambigüedad inherentes al lenguaje humano. Recordemos que muchos conceptos (i.e. palabras o predicados) P, como *bueno* o *joven*, no tienen una definición totalmente precisa, de manera que es habitual la aparición de incertidumbre y/o dudas en relación a si determinados objetos x del universo de discurso X considerado (e.g. *alternativas* de un problema de decisión o *edades* de posibles clientes) verifican o no esos predicados. En este marco, por medio de la asignación de grados continuos de pertenencia o verificación $\mu_P(x) \in [0,1]$, la lógica borrosa permite flexibilizar y salvar las dificultades de la lógica clásica binaria para afirmar la ocurrencia de $P(x)$ o $\neg P(x)$.

Paralelamente, la psicología cognitiva ha ido señalando [4,12] que el razonamiento humano tiende a analizar y organizar la realidad (e.g. una decisión a tomar) en términos de las relaciones positivas (de distensión, placer, beneficio, etc.) y/o negativas (de tensión, dolor, pérdida, etc.) que, a través de los aprendizajes pertinentes, guardan los objetos y la información disponible con respecto a la emotividad o racionalidad del individuo (e.g., una alternativa puede ser *buena* respecto a ciertos criterios y *mala* respecto de otros). De este modo, la realidad mental es construida en base al carácter y la relevancia afectivos de los objetos en consideración, o en otras palabras, en términos de pares de *polos de referencia P/Q*, como *bueno/malo* o *verdadero/falso*, que organizan y dan relevancia a la información disponible. Así pues, lo relevante no es saber si un objeto x es *bueno* ($P(x)$) o *no-bueno* ($\neg P(x)$), sino si es *bueno* ($P(x)$) o *malo* ($Q(x)$).

Esta coexistencia de pares de polos opuestos de referencia semántica suele ser referida en la literatura cognitiva como *bipolaridad*, aunque es preciso señalar que algunos autores [1] han usado también el término *intuicionismo* en relación a estas ideas.

Además, diferentes *tipos* de incertidumbre lingüística o borrosa (*fuzziness*) [20] y de bipolaridad [8] han sido definidos y estudiados con el objetivo de generalizar y ampliar el marco de aplicación de estas nociones seminales.

Así, mientras que el tratamiento habitual (o de tipo 1, F1) de la incertidumbre borrosa permite evaluar la imprecisión lingüística de una manera precisa (asignando valores de verdad graduables pero precisos $\mu_P(x) \in [0,1]$ a la proposición $P(x)$), la incertidumbre borrosa de tipo 2 (F2) hace referencia a la posibilidad de medir esa imprecisión de forma *imprecisa* (asignando conjuntos borrosos de tipo 1 como grados de verdad de $P(x)$).

De manera similar, en tanto que la bipolaridad de tipo 1 (B1) asume que la información negativa viene dada por la negación o complementación de la positiva, la bipolaridad de tipo 2 (B2) permite que la relación entre los polos no sea tan simple (e.g. *malo* ≠ *no-bueno*), lo que suele conllevar la necesidad de evaluar el par $(\mu_P(x), \mu_Q(x))$ para capturar toda la información relevante.

Nótese que, en tanto que tratan diferentes tipos de incertidumbre, las nociones de incertidumbre borrosa y bipolaridad parecen a primera vista independientes y complementarias, de modo que, en principio, un formalismo B2 puede ser asociado a una incertidumbre borrosa de tipo F1 o F2 (o incluso ser nítido), e igualmente un marco F2 puede estar asociado a un formalismo B1 o B2.

Sin embargo, el descubrimiento de la equivalencia formal [5,6] entre los conjuntos borrosos *intuicionistas* de Atanassov (AFS [1]), originalmente ideados como objetos de tipo F1 y B2 y los conjuntos borrosos intervalo-valorados (IVFS [11]), en principio concebidos como objetos F2 y B1, ha conducido a identificar esos marcos en cierta medida, lo cual ha acaecido no sin algunas disputas y enfrentamientos [3,7] entre algunos autores y escuelas.

El objetivo de este trabajo es arrojar algo de luz sobre las relaciones entre bipolaridad e incertidumbre borrosa desde una perspectiva diferente. Con este fin, se utilizará la noción de extensión 4-valorada (que recuerda claramente a la de estructura de preferencia [9, 10]) para definir y comparar rigurosamente la semántica y la estructura lógica subyacente de las posibles combinaciones de tipos de bipolaridad e incertidumbre borrosa. Esto permitirá separar y distinguir los AFS de los IVFS de una manera práctica, mostrando a su vez la independencia de esas nociones.

2 TIPOS DE INCERTIDUMBRE BORROSA

La teoría de conjuntos borrosos [19] permite que los objetos de un universo de discurso X verifiquen de forma parcial un predicado impreciso P. Para ello, la función característica del conjunto $P = \{x \in X \mid P(x)\}$ es generalizada mediante una *función de pertenencia* $\mu_P : X \to [0,1]$, de modo que $\mu_P(x) \in [0,1]$ indica el grado en que el objeto $x \in X$ verifica el predicado P. Así, la semántica (o el uso) de P sobre X es modelada a través de un conjunto borroso de tipo 1 (T1FS) $P = \{(x, \mu_P(x)) \mid x \in X\}$.

En respuesta a algunas críticas que señalaban la posible complejidad de asignar grados de pertenencia totalmente precisos, L.A. Zadeh introdujo en [20] la noción de incertidumbre borrosa de tipo n, permitiendo evaluar la veracidad de la proposición $P(x)$ por medio de un conjunto borroso de tipo n-1. Particularmente, si $F([0,1])$ denota el conjunto de todos los conjuntos borrosos de tipo 1 sobre [0,1], un conjunto borroso P de tipo 2 (T2FS) se asocia a una función de pertenencia $\psi_P : X \to F([0,1])$, de manera que $\psi_P(x) : [0,1] \to [0,1]$ expresa la *plausibilidad* de la proposición "x verifica P con grado de verdad μ" para cada grado de verdad $\mu \in [0,1]$. Así se introduce la imprecisión en la evaluación de los grados de verdad de tipo 1 $\mu_P(x)$ y, en general, los tipos de incertidumbre borrosa más elevados (i.e. tal que n>1) permiten introducir una mayor imprecisión en la evaluación de los grados de verdad.

Los conjuntos borrosos intervalo-valorados (IVFS, [11]) son quizás la clase más simple y común de conjuntos borrosos de tipo 2, en tanto que asignan a cada objeto x y predicado P un intervalo $[\mu_L(x), \mu_U(x)]_P$ como valores (igualmente y totalmente) plausibles del grado de verdad $\mu_P(x)$. Por tanto, el espacio de valuación de los IVFS viene dado por el conjunto $L^I = \{[a,b] \subseteq [0,1] \mid a \le b\} \subset F([0,1])$. Además, a mayor amplitud del intervalo $[\mu_L(x), \mu_U(x)]$, mayor es la incertidumbre asociada a él, de modo que su longitud $u(x) = \mu_U(x) - \mu_L(x)$ suele verse como el grado de *ignorancia* asociada a esa evaluación. Así, $u(x) = 0$ es cierto si, y solo si (sii), no hay incertidumbre sobre el grado $\mu_P(x)$ en que x verifica P, esto es, sii $\mu_L(x) = \mu_P(x) = \mu_U(x)$.

Los operadores lógicos clásicos, como la negación, la disyunción o la conjunción pueden ser generalizados al caso difuso mediante diversos operadores borrosos. Así, la negación habitual para conjuntos borrosos de tipo 1 está dada por $n(\mu) = 1 - \mu$, de modo que $\mu_{\neg P}(x) = n(\mu_P(x))$, con $\neg P$ =*no-P*. La negación habitual sobre los IVFS vienen dada por

$$n_{L^I}([\mu_L(x), \mu_U(x)]_P) = [1 - \mu_U(x), 1 - \mu_L(x)]_{\neg P}.$$

Similarmente, t-normas y t-conormas [17] suelen tomarse como operadores para la conjunción y la disyunción, respectivamente. Ambos tipos de conectores están relacionados a través de la negación, de manera que si T es una t-norma, entonces $S(\mu_P, \mu_Q) = n[T(n(\mu_P), n(\mu_Q))]$ es una t-conorma. Para T1FS, ejemplos habituales de operadores así relacionados son la t-norma y conorma de Lukasiewicz, respectivamente dados, para $a,b \in [0,1]$, por $a \odot b = \max\{a+b-1, 0\}$ y $a \oplus b = \min\{a+b, 1\}$, o también la t-norma $a \wedge b = \min\{a,b\}$ y su respectiva t-conorma $a \vee b = \max\{a,b\}$. Para IVFS, estos operadores son extendidos de la siguiente forma:

$$[a,b] \odot_I [c,d] = [a \odot c, b \odot d], [a,b] \oplus_I [c,d] = [a \oplus c, b \oplus d],$$

$[a,b] \wedge_I [c,d] = [a \wedge c, b \wedge d], [a,b] \vee_I [c,d] = [a \vee c, b \vee d]$.

Por otro lado, nótese que la lógica borrosa parece infravalorar la noción de información negativa, en tanto que asume que la falsedad de $P(x)$ ha de ser igual a la verdad de $\neg P(x)$, i.e., $\mu_{\neg P}(x)$, aunque este grado de falsedad puede ser de naturaleza diferente a la información negativa. En cualquier caso, es claro que la lógica borrosa no considera una evaluación independiente junto a los grados (precisos o no) $\mu_P(x)$ o $\psi_P(x)$.

3 TIPOS DE BIPOLARIDAD

Básicamente, los modelos de representación bipolares asumen la existencia de un par de polos de referencia semántica *P*/*Q*, tales como *verdadero*/*falso* o *bueno*/*malo*, los cuales proporcionan marcadores absolutos que confieren a la información su carácter intrínsicamente positivo o negativo. La información sin carácter positivo ni negativo es entonces considerada irrelevante o *neutral* en términos de esas referencias, del mismo modo que la información con carácter simultáneamente positivo y negativo es considerada como *conflictiva* o contradictoria. Así, los polos *P*/*Q* organizan y dan relevancia a la información disponible, de manera que la información positiva para uno de ellos es considerada como negativa para el otro.

Si la relación entre los polos está dada por la negación, i.e. si $Q = no\text{-}P = \neg P$ (e.g. *malo=no-bueno*), entonces una única evaluación $\mu_P(x) \in [0,1]$ (o $\mu_P(x) \in L^I$) es suficiente para capturar toda la información relevante en términos de la polaridad *P*/*Q*, ya que entonces $\mu_Q(x) = \mu_{\neg P}(x) = 1 - \mu_P(x)$ y por lo tanto la información negativa es solo la negación de la positiva. El término bipolaridad de tipo 1 (B1, [8]) es habitualmente usado para referirse a esta situación, y nótese que los conjuntos borrosos (tanto F1 como F2) y la lógica clásica pertenecen a esta categoría.

Sin embargo, si la relación entre *P* y *Q* es más compleja (e.g. $malo \neq no\text{-}bueno$), entonces es normalmente necesario evaluar también $\mu_Q(x)$ para capturar toda la información relevante sobre $P(x)$. En particular, es posible usar dos escalas L_P^+ y L_P^- para evaluar, respectivamente, $\mu_P(x)$ y $\mu_Q(x)$, de manera que el par $(\mu^+(x), \mu^-(x))_P = (\mu_P(x), \mu_Q(x)) \in L_P^+ \times L_P^-$ mide de manera conjunta el grado de información positiva y negativa sobre $P(x)$. Por otro lado, si se asume que la relación entre los polos es simétrica (aunque es posible eliminar esta asunción, ver [16]), entonces

$(\mu^+(x), \mu^-(x))_Q = (\mu_Q(x), \mu_P(x)) \in L_Q^+ \times L_Q^- = L_P^- \times L_P^+$.

Estos pares de evidencia son típicos de la bipolaridad de tipo 2 (B2, [8]), y nótese que, en tanto que *P* y *Q* no son necesariamente disjuntos, los modelos B2 admiten en principio 4 casos posibles respecto de la relación entre un objeto *x* y una polaridad *P*/*Q*:

- *Veracidad positiva t*: *x* verifica *P* pero no *Q*, i.e., la información es positiva y no negativa. El par $(\mu^+, \mu^-) = (1,0)$ es el representante extremo (o nítido) de esta situación.
- *Veracidad negativa* (o *falsedad positiva) f*: *x* cumple *Q* pero no verifica *P*, luego la información es negativa y no positiva. El par extremo es $(\mu^+, \mu^-) = (0,1)$.
- *Irrelevancia o neutralidad i: x* no verifica *P* ni *Q*, por lo que la información es no positiva y no negativa y en el caso extremo se tiene $(\mu^+, \mu^-) = (0,0)$;
- *Conflicto k: x* verifica *P* y *Q* simultáneamente, luego la información es positiva y negativa, como en el caso extremo $(\mu^+, \mu^-) = (1,1)$.

Es importante señalar que los casos tercero y cuarto no son posibles en un marco B1, en el que *P* y *Q* están estrechamente ligados a través de la negación. Si en un marco B2 la relación entre los polos es restringida, el tercer o cuarto caso pueden dejar de ser posibles, pero no simultáneamente, ya que entonces ese marco sería B1.

Es interesante examinar con más detalle el caso de los AFS [1]: estos asignan a $P(x)$ un grado de pertenencia $\mu_P(x)$ así como un grado de *no-pertenencia* $\mu_{\neg P}(x)$, cumpliendo que $\mu_P(x), \mu_{\neg P}(x) \in [0,1]$ y $\mu_P(x) + \mu_{\neg P}(x) \leq 1$. Por tanto, el espacio de valuación de los AFS es el conjunto $L^* = \{(x,y) \in [0,1]^2 \mid x + y \leq 1\} \subset [0,1]^2$, y nótese que la misma escala $L^+ = L^- = [0,1]$ es usada para ambas evaluaciones.

Aunque la presencia de dos evaluaciones hace pensar que los AFS pertenecen a un marco B2, lo cierto es que, implícitamente, Atanassov trata un marco B1 (en el que la polaridad relevante es $P / \neg P$) como un marco B2, asumiendo que la verdad de $\neg P(x)$ no puede ser directamente obtenida a partir de la de $P(x)$, por tanto requiriendo una evaluación independiente $\mu_{\neg P}(x) = \nu_P(x)$. En tanto que $\mu_P(x) + \nu_P(x) \leq 1$, se obtiene que la disyunción de P y su negación (en términos de la t-conorma de Lukasiewicz $\oplus$) puede no ser una tautología, violando por ello la ley del tercio excluso ($P \vee \neg P$), en consonancia con la intención de Atanassov de formular un modelo *intuicionista* [1].

Sin embargo, otras estructuras lógicas borrosas también incumplen esa ley sin considerar una negación independiente (e.g. tomando la t-conorma máximo $\vee$ junto a $\mu_{\neg P}(x) = 1 - \mu_P(x)$). Por ello, el significado

concreto de los AFS es relativamente oscuro, pues muestran la semántica de un marco B1 mientras que formalmente se asemejan a un marco B2. Además, en tanto que el modelo de Atanassov nunca proporciona una definición clara de su estructuras lógica y semántica subyacente (ver [12]), no hubo razones sólidas para separar los AFS de los IVFS cuando se demostró la equivalencia [5,6] de ambos formalismos (mediante el isomorfismo $\Phi: L^* \to L^I$ tal que

$$\Phi(\langle \mu_P(x), \mu_{\neg P}(x) \rangle) = [\mu_P(x), 1 - \mu_{\neg P}(x)]).$$

Esta equivalencia produjo una fuerte controversia (ver [3,7]) entre Atanassov y sus partidarios, por un lado, que defendieron la validez de los AFS como objetos B2 (o al menos su validez como modelo *intuicionista*) diferentes de los IVFS, y por otro lado, una importante parte de la comunidad *fuzzy*, que consideraba que el significado intuicionista de los AFS no estaba en absoluto claro, y, además, en tanto que los IVFS fueron formulados previamente a los AFS, la relevancia de estos como objetos bipolares debía ser reducida debido a tal equivalencia (a pesar de que los IVFS fueron originalmente concebidos como objetos B1).

En nuestra opinión, la semántica original de los AFS (sea cual sea) no es de tipo B2. Sin embargo, si ν es interpretada como la función de pertenencia de un polo Q tal que $Q \neq \neg P$, i.e. $\nu_P(x) = \mu_Q(x)$, y se mantiene la restricción $\mu_P(x) + \nu_P(x) \leq 1$, entonces se obtiene fácilmente una semántica B2 (con una relación particular entre los polos). Nuestra opinión es que estos *nuevos* objetos, (que denominaremos AFS *bipolares*, BAFS) no son semánticamente equivalentes a los IVFS. Para sostener esta afirmación, en la próxima sección definiremos formalmente la semántica de los objetos B2, mostrando que la semántica de los BAFS (y en general la de los formalismos F1 y B2) difiere de la de los IVFS (y en general de la semántica de los formalismos F2 y B1). Esta diferenciación semántica será introducida mediante la noción de extensión 4-valorada.

4 EXTENSIONES 4-VALORADAS

Como se señaló en la sección anterior, en un marco B2 la información positiva y negativa proporcionada por los pares $(\mu^+, \mu^-) \in \{0,1\}^2$ conduce de forma natural a cuatro casos que reflejan el significado o semántica de esos pares en términos cognitivos. Esto es, la descripción de la información contenida en los pares se traduce en 4 diferentes *estados epistémicos*, que reflejan la situación de conocimiento, en términos de una polaridad *P*/*Q*, a la que aboca la evidencia disponible sobre un objeto *x*.

Cuando la evidencia es evaluada de forma borrosa, esto es, cuando $(\mu^+, \mu^-) \in [0,1]^2$ ($o \in (L^I)^2$), la traducción de los pares en estados epistémicos es más compleja, en tanto que un par puede ser parcialmente compatible con varios de esos 4 estados. En otras palabras, en un contexto borroso esos cuatro estados son también graduales, y es entonces preciso proporcionar una extensión continua 4-valorada de esos pares que especifique la semántica de los pares no extremos. Aunque es posible definir concretamente estas extensiones de diversas maneras (dando lugar a diferentes semánticas), en este trabajo se opta por una extensión que generaliza las nociones de verdad y falsedad habituales de la lógica borrosa.

Así pues, suponiendo que $L^+ = L^- = L$, $L = [0,1]$ ó L^I, con cada par $(\mu^+, \mu^-) \in L^+ \times L^-$ se asocia una *matriz de evidencia*

$$EM = \begin{bmatrix} t(\mu^+, \mu^-) & i(\mu^+, \mu^-) \\ k(\mu^+, \mu^-) & f(\mu^+, \mu^-) \end{bmatrix},$$

obtenida mediante las fórmulas

$$t(\mu^+, \mu^-) = \mu^+ \wedge \neg\mu^-,\ f(\mu^+, \mu^-) = \neg\mu^+ \wedge \mu^-, \qquad (1)$$

$$i(\mu^+, \mu^-) = \neg\mu^+ \odot \neg\mu^-,\ k(\mu^+, \mu^-) = \mu^+ \odot \mu^-, \qquad (2)$$

donde $\neg\mu$ denota la negación (en *L*) de la evaluación μ. Nótese que $t, f, i, k \in L$, y por tanto *EM* evalúa, en la misma escala en que se valoran μ^+ y μ^-, el grado en que se verifica cada uno de los cuatro estados epistémicos.

La extensión dada en las fórmulas (1)-(2) fue propuesta para el caso *L*=[0,1] en [14] (y ha seguido siendo analizada en [18]), donde se prueba que es la *única* extensión continua basada en t-normas verificando simultáneamente que los estados *i* y *k* son mutuamente exclusivos (ya que si $i>0$ entonces $k=0$ y viceversa) y que $t + f + i + k = 1$ para todo $(\mu^+, \mu^-) \in L^+ \times L^-$. Nótese que intercambiando las posiciones de la t-normas mínimo $\wedge$ y de Lukasiewicz $\odot$ se obtiene otra extensión (utilizada en modelización de preferencias [9]), que también cumple $t + f + i + k = 1$ pero en la que ahora *t* y *f* son exclusivos. Sin embargo, como veremos, esta última extensión no generaliza las nociones clásicas de verdad y falsedad en el caso en que $Q = \neg P$, mientras que la dada por (1)-(2) sí lo hace.

Los resultados en [14] relativos a la extensión dada por (1)-(2) pueden ser generalizados al caso $L = L^I$ mediante el siguiente teorema, cuya demostración puede encontrarse en [15]:

Teorema: Sean $\mu^+ = [\mu_L^+, \mu_U^+]$ y $\mu^- = [\mu_L^-, \mu_U^-]$ ($0 \leq \mu_L^- \leq \mu_U^- \leq 1$, $0 \leq \mu_L^+ \leq \mu_U^+ \leq 1$) las evaluaciones intervalo-valoradas positiva y negativa de la proposición "$x \in P$". Sean también $t = [t_L, t_U]$, $f = [f_L, f_U]$, $i = [i_L, i_U]$ y $k = [k_L, k_U]$ la extensión 4-valorada asociada al par (μ^+, μ^-), obtenida a través de las fórmulas (1)-(2).

Entonces k_L y i_U así como i_L y k_U son mutuamente exclusivos y se cumple que

$$t_L + f_L + i_U + k_U = t_U + f_U + i_L + k_L = 1 + m,$$

donde $m = \min\{u^+, u^-, u_t, u_f, u_u, u_k\}$ y $u^+ = \mu_U^+ - \mu_L^+$, $u^- = \mu_U^- - \mu_L^-$, $u_t = t_U - t_L$, $u_f = f_U - f_L$, $u_i = i_U - i_L$, $u_k = k_U - k_L$ son los grados de ignorancia asociados con cada intervalo.

A continuación, la extensión (1)-(2) será aplicada a diferentes combinaciones de tipos de bipolaridad y de incertidumbre borrosa para ilustrar la semántica particular de cada caso:

Incertidumbre borrosa + bipolaridad de tipo 1 **(F1,B1)**

En este caso es $L=[0,1]$ y $Q = \neg P$, por lo que $\mu^+ = \mu_P$ y $\mu^- = \mu_{\neg P} = n(\mu_P) = 1 - \mu_P$. Esto implica que $t = \mu^+$, $f = \mu^-$ y $k = i = 0$. En particular, $m = 0$ en tanto que no se admite ignorancia respecto de los valores de μ^+, μ^-. Como ejemplo, supóngase un $x \in X$ tal que $\mu_P(x) = 0.6$, esto es, se tiene el par $(\mu^+, \mu^-) = (0.6, 0.4)$ y la matriz

$$EM_1 = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.4 \end{bmatrix}.$$

Por tanto, la extensión (1)-(2) generaliza las nociones de verdad y falsedad de la lógica borrosa (identificadas con μ_P y $\mu_{\neg P}$), y refleja la semántica B1 de este marco, que excluye los estados *k* e *i*.

Incertidumbre borrosa tipo 2+bipolaridad tipo 1 **(F2,B1)**

Supóngase ahora que, en el ejemplo anterior, el valor de $\mu_P(x)$ no se conoce con precisión, sino que se sabe que cae en el intervalo $[\mu_L(x), \mu_U(x)]_P$. Por tanto, se asume aún que $Q = \neg P$, pero ahora $L = L^I$ y entonces $\mu^+ = [\mu_L, \mu_U]_P$ y $\mu^- = n_{L^I}([\mu_L, \mu_U]_P) = [1 - \mu_U, 1 - \mu_L]_{\neg P}$. Así, se obtiene de nuevo que $t = \mu^+$ y $f = \mu^-$ pero ahora $k = i = [0, u]$ y $m=u$, donde $u = \mu_U - \mu_L$ es la ignorancia asociada a μ_P. Por ejemplo, si $\mu^+ = [0.5, 0.7]$ (luego $\mu^- = [0.3, 0.5]$), se obtiene la matriz

$$EM_2 = \begin{bmatrix} [0.5, 0.7] & [0, 0.2] \\ [0, 0.2] & [0.3, 0.5] \end{bmatrix},$$

mostrando que ahora las evaluaciones de la veracidad y falsedad positivas son también afectadas por ignorancia, permitiendo que *i* y *k* sean no nulas. Sin embargo, si la ignorancia se reduce, i.e. al tender *u* a 0, *k* e *i* tienden a [0,0], aproximándose al caso anterior (F1,B1).

Incertidumbre borrosa tipo 1+bipolaridad tipo 2 **(F1,B2)**

Si se permite que la información negativa sobre "$x \in P$" sea independiente de la negación *n* del primer caso (F1,B1), se obtiene un marco B2 en el que $Q \neq \neg P$, $\mu^+ = \mu_P$, $\mu^- = \mu_Q$ y $L^+ = L^- = L = [0,1]$. Luego $m=u=0$, pero *t* (*f*) ya no puede ser identificada con μ^+ (μ^-), y uno de los estados *i*,*k* puede cumplirse. Por ejemplo, si $(\mu^+, \mu^-) = (0.6, 0.7)$, se obtiene la matriz

$$EM_3 = \begin{bmatrix} 0.3 & 0 \\ 0.3 & 0.4 \end{bmatrix},$$

que muestra que $t \neq \mu^+$, $f \neq \mu^-$ e informa de que la información es parcialmente conflictiva (k=0.3), implicando que $t < \mu^+$ y $f < \mu^-$.

Nótese que $(\mu^+, \mu^-) = (0.5, 0.3) \in L^*$ es el par asociado al intervalo $[0.5, 0.7] \in L^I$ del caso previo a través del isomorfismo $\Phi : L^* \to L^I$, y que produce la matriz

$$EM_4 = \begin{bmatrix} 0.5 & 0.2 \\ 0 & 0.3 \end{bmatrix}.$$

Se tiene que $EM_2 \neq EM_4$, y no solo porque EM_2 introduce ignorancia mientras que EM_4 no lo hace. Si la ignorancia *u* se reduce, entonces podría tenerse $\mu^+_{EM_2} = [0.5, 0.5]$, de manera que $t_{EM_2} = [0.5, 0.5] = 0.5 = t_{EM_4}$, pero entonces sería $f_{EM_2} = [0.5, 0.5] \neq 0.3 = f_{EM_4}$ e $i_{EM_2} = [0,0] \neq 0.2 = i_{EM_4}$ ya que $\mu^-_{EM_2} = n_{L^I}(\mu^+_{EM_2}) = [0.5, 0.5]$ por el marco B1 de los IVFS. Sin embargo, los BAFS permiten que μ^- sea independiente de μ^+ al exhibir una bipolaridad de tipo 2, aunque la condición $\mu^+ + \mu^- \leq 1$ implique que k=0. Entonces, es claro que la semántica en términos de estados epistémicos de los BAFS es diferente de la de los IVFS, y su equivalencia debería ser entendida como una coincidencia formal entre dos escalas que pertenecen a universos diferentes, pues $[0,1] \subset L^I \subset F([0,1])$ mientras que $L^* \subset [0,1]^2 \subset L^I \times L^I \subset F([0,1]) \times F([0,1])$.

Incertidumbre borrosa tipo 2+bipolaridad tipo 2 **(F2,B2)**

Finalmente, supóngase que, en un marco B2, se permite ignorancia sobre las evaluaciones μ^+, μ^-, de modo que se obtiene un par $(\mu^+, \mu^-) \in L^I \times L^I$. Nótese que este caso generaliza todos los anteriores y coincide con las hipótesis del teorema anterior. Entonces, si $\mu^+ = [0.5, 0.7]$ y $\mu^- = [0.6, 0.9]$, se obtiene la matriz

$$EM_5 = \begin{bmatrix} [0.1, 0.4] & [0,0] \\ [0.1, 0.6] & [0.3, 0.5] \end{bmatrix},$$

y obsérvese que ahora cada uno de sus elementos tiene asociada un grado de ignorancia diferente (y m=0), y que estos resultados extienden los de EM_1, EM_2 y EM_3. Nótese también que u^+ o u^- pueden ser reducidas sin afectar a la otra, y en el caso límite en que $u^+ = u^- = 0$, se alcanza un marco F1 y B2. Un ejemplo de formalismo B2 y F2 son los IVFS de Atanassov (AIVFS, ver [2]), aunque es válido el mismo argumento de la Sección 3 para afirmar que estos no pueden ser vistos como objetos B2 a menos que se asuma la existencia de un polo negativo $Q \neq \neg P$ (dando lugar a AIVFS *bipolares* o BAIVFS).

5 CONCLUSIONES

Las nociones de incertidumbre borrosa y de bipolaridad han sido revisadas y analizadas. En tanto fueron desarrolladas para afrontar la modelización de características distintas del razonamiento humano, ambas nociones son de hecho complementarias e independientes. Un formalismo F2 puede ser B2 o no serlo, y recíprocamente un marco B2 puede permitir ignorancia (siendo entonces también una marco F2) o no. En particular, esta diferencia se ha hecho evidente en este trabajo, clarificando algunos aspectos de las relaciones entre los formalismos B1,F2 y los B2,F1. De manera implícita, este trabajo muestra que existe cierta diferencia entre las nociones de *intuicionismo* (en el sentido de usar una negación sub-aditiva $n(\mu) \leq 1-\mu$) y de bipolaridad. Sin embargo, si se asume la presencia de un polo negativo $Q \neq \neg P$, entonces los BAFS y BAIVFS constituyen formalismos relevantes dentro de un marco B2. La siguiente Tabla 1 presenta una tipología de los casos y formalismos estudiados en este trabajo, los cuales muestran semánticas diferentes en términos de los estados cognitivos o epistémicos a los que conducen.

Tabla 1. Combinaciones de tipos 1 y 2 de incertidumbre borrosa y bipolaridad.

Fuzziness / **Bipolaridad**	Tipo 1 (F1)	Tipo 2 (F2)
Tipo 1 (B1)	T1FS	T2FS, IVFS
Tipo 2 (B2)	BAFS, BFS	BT2FS, BAIVFS

En el futuro, este trabajo se extenderá en diversas direcciones: se introducirán retículos (*lattices*) para obtener un marco lógico más general y riguroso. También, se explorarán con más detenimiento las relaciones entre bipolaridad, intuicionismo, el uso de negaciones sub-aditivas y los operadores de disimilaridad [16]. Además, se estudiará el posible uso de los formalismos F2 para combinar la incertidumbre borrosa con otras teorías de la incertidumbre (probabilística, posibilística, etc.).

Agradecimientos

Este trabajo ha sido realizado en el marco del proyecto TIN2009-07901.

Referencias

[1] Atanassov KT (1986) Intuitionistic Fuzzy-Sets, F Set Syst, 20 (1) 87-96

[2] Atanassov KT, Gargov G (1989) Interval Valued Intuitionistic Fuzzy Sets, F Set Syst, 31(3) 343-349

[3] Atanassov KT (2005) Answer to D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk and H. Prade's paper "Terminological difficulties in fuzzy set theory - the case of "Intuitionistic fuzzy sets", Fuzzy Set Syst, 156 (3) 496-499

[4] Cacioppo JT, GardnerWL, Berntson GG (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. Pers Soc Psychol Rev 1:3–25.

[5] Cornelis C, Atanassov K, Kerre E, Intuitionistic fuzzy sets and interval-valued fuzzy sets: a critical comparison, Proc. Eusflat'03, 2003, 159–163

[6] Deschrijver G, Kerre E (2004) On the relationship between some extensions of fuzzy set theory, Fuzzy Set Syst 133 227–235.

[7] Dubois D, Gottwald S, Hajek P, Kacprzyk J, Prade H (2005) Terminological difficulties in fuzzy set theory - The case of "Intuitionistic Fuzzy Sets", Fuzzy Set Syst, 156 (3) 485-491

[8] Dubois D, Prade H (2008) An introduction to bipolar representations of information and preference, Int J Intell Syst, 23 (8) 866-877

[9] Fodor J, Roubens M (1994) Fuzzy preference modelling and multicriteria decision support, Kluwer Academic, Dordrecht ; Boston

[10] Franco C, Rodríguez T, Montero J. (2010) Information measures over intuitionistic four valued fuzzy preferences. Proceedings IEEE-WCCI Conference, Barcelona, July 18-23, pages 1971-1978

[11] Grattan-Guinness I (1975) Fuzzy membership mapped onto interval and many-valued quantities, Z. Math. Logik Grundlag. Mathe., 22 149-160

[12] Medin D, Schwanenflugel P (1981) Linear separability in classification learning, J. of Experimental Psychology: Human Learning & Memory, 7 355-368

[13] Montero J, Gómez D, Bustince H (2007) On the relevance of some families of fuzzy sets, Fuzzy Set Syst, 158 (22): 2439-2442

[14] Öztürk M, Tsoukiàs A (2007) Modelling uncertain positive and negative reasons in decision aiding, Decision Support Systems, 43 (4): 1512-1526

[15] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) Eurofuse

[16] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) An axiomatic approach to the notion of semantic antagonism, Procs. of the IFSA-AFSS'11, FT104-1/6

[17] Schweizer B, Sklar A (1983) Probabilistic Metric Spaces. North–Holland/Elsevier, New York.

[18] Turunen E, Ozturk M, Tsoukiàs A (2010) Paraconsistent semantics for Pavelka style fuzzy sentential logic Fuzzy Set Syst, 161 (14) 1926-1940

[19] Zadeh LA (1965) Fuzzy Sets, Information and Control, 8 (3) 338-353

[20] Zadeh, LA (1975) The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, Information Sciences 9 (1) 43-80

Abstracción en Lógica Fuzzy

Adolfo R. de Soto

Escuela de Ingenierías Industrial e Informática
Universidad de León, adolfo.rdesoto@unileon.es

Resumen

En este trabajo se muestra la excelente capacidad que ofrecen los sistemas fuzzy para modelizar los procesos de abstracción del pensamiento humano. Para ello se revisa el concepto de abstracción en diversas disciplinas científicas, intentando extraer las características básicas de este fenómeno. Finalmente se hace evidente cómo los sistemas de reglas fuzzy son esencialmente un mecanismo de razonamiento abstracto, característica que no ha sido resaltada suficientemente y que debe ser explotada con más intensidad.

Palabras Clave: Abstracción, Lógica Fuzzy, Granular Computing, Computación con palabras..

"It is this lack of understanding of the way in which humans can abstract in mathematically ill-defined situations and our consequent inability to devise abstracting devices which can perform even remotely as well as the human mind, that lie at the root of many unresolved problems in heuristic programming, pattern recognition and related problem areas."

Lotfi A. Zadeh [21]

1 INTRODUCCIÓN

Muchos son los significados asociados al término "abstracción" y que aparecen de una forma directa o indirecta en cualquier diccionario que se consulte. Expresiones tales como: "ignorar los detalles", "no concreto", "lejos de la experiencia", "no fácilmente comprendido", "una consideración teórica", "oculto", "complejo", "una forma de pensar no basada en una instancia concreta", "una construcción mental", "el proceso de identificar", o "hacer equivalente a través de clasificaciones" y otros que seguramente cualquier persona consideraría habitualmente asociadas a este término. Sin lugar a dudas esa diversidad de aspectos es un reflejo de la importancia que la abstracción tiene en el proceso de conocimiento humano. Y no sólo en las ciencias que propiamente son calificadas de abstractas como las Matemáticas o la Lógica sino en cualquier rama del conocimiento, puesto que la abstracción aparece en los procesos básicos de razonamiento, en los mecanismos de aprendizaje, incluso en los mecanismos de percepción.

Por supuesto, la abstracción ha sido estudiada de una forma u otra por múltiples disciplinas científicas y artísticas tales como la Filosofía, la Psicología, la Pedagogía, la Lingüística, las Ciencias del Conocimiento, las Ciencias de la Computación y la Inteligencia Artificial, las Bellas Artes, entre otras. Todo ello hace muy difícil, sino imposible, intentar llegar a una comprensión global y unificadora del concepto que dé cuenta de las múltiples facetas en que se presenta. Sin embargo, como se indica en [4], decir que la abstracción no es un proceso universal y objetivo puede ser adecuado refiriéndose exclusivamente a la abstracción como una actividad, como un proceso interactivo, que involucra herramientas, lenguaje y procedimientos, pero no se debe obviar que la abstracción es también un proceso neuronal, y que por ello, casi seguro, tiene características universales.

Desde el punto de vista computacional, la abstracción forma parte del currículo básico. En programación se le enseña al alumno desde casi sus primeros pasos de programación a adoptar un nivel adecuado de abstracción acorde con la etapa de análisis, diseño o implementación en el que se encuentre, aunque casi siempre relacionada con la ocultación de información. Sin embargo el mayor interés de este trabajo se centrará en el soporte que la capacidad de abstracción ofrece a los procesos de razonamiento humanos, fácilmente apreciables en el razonamiento por analogía, la generalización, la similaridad, la clasificación, la aproximación, el razonamiento simbólico, etc. En [15] se argumenta que la información abstracta es en muchas ocasiones más útil que la información específica y a menudo el

pilar fundamental en el descubrimiento científico. El conocimiento humano se basaría por tanto en un gran porcentaje en su capacidad de abstracción.

El objetivo de este trabajo es mostrar algunas de esas características universales de la abstracción así como la enorme capacidad que la Lógica Fuzzy presenta para implementar estas características.

2 VISIONES MULTIDISCIPLINARES

Debido a la vasta cantidad de trabajos que de una forma u otra tratan sobre la abstracción y sus procesos asociados, se hace imposible incluirlos todos, ni siquiera la mayor parte. Por eso el trabajo se centrará en ciertos aspectos que se consideran más relevantes para el objetivo general del trabajo. En particular muchos aspectos filosóficos del concepto de abstracción no son tratados, como la dualidad concreto/abstracto, el problema de los universales u otros, a pesar de que su interés intrínseco es importante y que también presentan aspectos que podrían ser interesantes desde el punto de vista computacional. A continuación se muestra por tanto una selección de ciertos trabajos acerca de la abstracción en diferentes disciplinas.

2.1 LA VISIÓN DE PIAGET

En Pedagogía la abstracción ha jugado un papel muy relevante debido a su capacidad de explicación de los procesos de aprendizaje. Especialmente interesante es la visión de Piaget sobre ello [8].

Tres son los tipos de abstracción que distingue Piaget: Abstracción Empírica, Abstracción Pseudo-empírica y Abstracción Reflectiva. La Abstracción Empírica permite derivar conocimiento de las propiedades de los objetos. Ese conocimiento proviene de experiencias externas pero es el resultado de construcciones hechas internamente por el sujeto. Como ejemplo Piaget pone la construcción de universales, el paso de "algunos" a "todos", el paso de lo específico a lo general. La Abstracción Pseudo-empírica es una intermedia entre las otras dos, y permite extraer propiedades de los objetos a través de las acciones que el sujeto ha realizado sobre los objetos. Finalmente la Abstracción Reflectiva es la que produce el sujeto por sí mismo y es completamente endógena. Esta clase de abstracción conduce, según Piaget, a una clase muy diferente de generalización constructiva y resulta en nuevas síntesis a través de las cuales leyes particulares adquieren un nuevo significado. Las tres clases de abstracción no son completamente independientes entre sí.

Piaget considera la Abstracción Reflectiva como la fuerza motora que permite las reconstrucciones involucradas en el camino de: acciones sensor-motoras, representaciones semióticas, operaciones concretas y operaciones formales. Piaget siente que la consciencia del sujeto acerca de acciones físicas o mentales es uno de los resultados de aplicar Abstracción Reflectiva y que ésta permite comprender esas acciones. La importancia que Piaget otorga a este mecanismo se puede evaluar cuando dice que comprendida adecuadamente, la Abstracción Reflectiva es una descripción del mecanismo de desarrollo del pensamiento intelectual, y que aparece en el desarrollo cognitivo a lo largo de la vida de la persona, desde las primeras coordinaciones del niño. El siente que la consciencia del sujeto de sus acciones mentales o físicas es uno de los resultados de aplicar la Abstracción Reflectiva.

De acuerdo a Piaget "hay un proceso que llega ser más evidente a medida de que pasa el tiempo: la construcción de nuevas combinaciones por la conjunción de abstracciones". Él describe cuatro clases diferentes de métodos de construcción de Abstracciones Reflectivas:

1. La aparición de la función semiótica, esto es, la habilidad para usar símbolos, lenguaje, imágenes. Es un método de representación. Piaget llama a esto "interiorización", "traslación de una sucesión de acciones materiales dentro de un sistema de operaciones interiorizadas".
2. Composición o coordinación de dos o más procesos para construir uno nuevo.
3. Encapsulación o conversión de un proceso (dinámico) en un objeto (estático). Las Matemáticas pueden ser vistas así: construir nuevas formas que se apoyan en formas previas e incluirlas como sus contenidos.
4. Cuando un sujeto aprende a aplicar un esquema existente a una colección más amplia de fenómenos: generalizar el esquema, como por ejemplo en el razonamiento por analogía o el razonamiento metafórico.

En [8] se añade a estos métodos la capacidad humana de crear un nuevo proceso a partir de la inversión de otro.

En resumen puede decirse que de acuerdo a Piaget la Abstracción Reflectiva será la capacidad de construir objetos mentales y acciones sobre esos objetos. Piaget utiliza el concepto de "schema" para referirse a una colección más o menos coherente de objetos y procesos.

2.2 ABSTRACCIÓN Y PERCEPCIÓN

En [1], Arnheim, un teórico del arte, comprende por abstracción la capacidad de la mente de concebir algo más allá de cualquier representación concreta, de crear sin ningún soporte basado en lo que es perceptualmente dado o recordado. Sin embargo él mismo considera que en la percepción humana se producen fenómenos de abstracción

Según Arnheim, la percepción humana, o la animal, es un instrumento desarrollado durante la evolución para sobrevivir y para ser alertado del peligro. Esas necesidades se

referían a un tipo o clase de cosas más que a individuos particulares. Es por ello que procesos de abstracción ocurren ya en la percepción. Los elementos de pensamiento en la percepción y elementos perceptuales en el pensamiento son complementarios. Hacen del proceso de conocimiento humano un proceso unitario, que conduce sin solución de continuidad desde la adquisición elemental de información sensorial a las ideas teóricas más genéricas. El punto esencial de este proceso cognitivo es que cada nivel conlleva abstracción. Los trabajos de la escuela de la Gestalt y la enunciación de sus leyes sobre la percepción no harían más que corroborar empíricamente este fenómeno. Por ello la naturaleza y el significado de la abstracción debe ser examinado con cuidado.

2.3 MALAS CONCEPCIONES DEL CONCEPTO DE ABSTRACCIÓN

Varios significados atribuidos a la abstracción son demasiado limitadores del fenómeno. Por ejemplo, el significado negativo que tiene la abstracción, falta de concreción, eliminación de algo, detalles, basada únicamente en palabras. Comúnmente se cree que una criatura animal privada de la capacidad de habla no podría abstraer. Sin embargo, como se ha visto Arnheim apoya el hecho de que procesos de abstracción ya se producen en la percepción.

La abstracción también se ha visto habitualmente como un proceso resultante de la capacidad de generalización humana. Sin embargo la abstracción no puede verse únicamente como la capacidad de reducir muchos a pocos, puesto que si sólo fuera así el proceso adecuado sería comenzar con propiedades o grupos de propiedades que se encuentran en un mayor número de individuos y trabajar gradualmente a aquellos representando menos individuos. Sin embargo los conceptos no se forman así. Estos no están relacionados con el número de individuos a los que se les puede aplicar. Más bien parece que la gente forma conceptos en función de sus intereses. Por otra parte la intensión y la extensión de una clase se presuponen mutuamente: es imposible revisar cada elemento de una clase para extraer la intensión correcta la cuál la define y también se necesita la intensión de la clase para seleccionar elementos de la misma. Otra forma de decirlo es que en orden a generalizar uno debe primero abstraer, pero para poder abstraer uno debe conocer previamente cómo generalizar. Es la percepción y su capacidad para referirse a clases de cosas y no tanto a casos individuales la que rompería el círculo, ello debido a que ciertos mecanismos de abstracción ya suceden en la percepción. Así generalización requiere abstracción.

De modo análogo inducción requiere abstracción. La razón es que si deseamos aplicar un método de inducción, una población debe ser seleccionada. La elección requiere un criterio, es decir, una abstracción de ciertas propiedades las cuáles deben estar presentes en los individuos que pueden ser elegidos.

Hay que tener en cuenta que la abstracción es un concepto relativo, no es una propiedad inherente de una cosa, sino más bien la relación entre la cosa y una persona o comunidad. Considerar el concepto de abstracción únicamente como opuesto al de concreto en el sentido de tener existencia material o física es otra interpretación limitadora. Una imagen, un número, una palabra puede ser algo muy concreto. Un número es algo concreto para un adulto pero no lo es para un niño de 3 años en un experimento de Piaget. Un pictograma chino puede ser totalmente abstracto para un occidental y algo muy concreto para un oriental. Se conoce por reificación al fenómeno de convertir un concepto abstracto en algo concreto. Cualquier fenómeno experimentado por la mente puede adquirir "abstractividad" si es visto como refinamiento de algo más complejo.

Los palabras siempre se han considerado como un mecanismo de abstracción, tradicionalmente se ha visto como una manifestación de la incapacidad de los animales para la abstracción precisamente su ausencia de lenguaje hablado. Los conceptos también son abstracciones pero no el mismo tipo de abstracción que las palabras. La conceptualización verbal no es la única clase de conceptualización en la consciencia humana. De la misma forma existen muchos tipos de abstracción que no son conceptuales sino procedurales, causales, jerárquicas.

2.4 ABSTRACCION: UN CAMINO DE IDA Y VUELTA

La abstracción no es exclusivamente una reducción (de lo concreto a lo abstracto) sino que opera en ambas direcciones [4]. Del mismo modo el razonamiento puede ir de lo particular a lo general y vuelta a lo particular. Por ejemplo, la observación de surcos en Marte (rasgos observados) propicia buscar en el nivel abstracto una explicación (teorías acerca de líquidos) y la adecuación o no de esa teoría a los rasgos observados propicia una vuelta a lo específico (una historia causal acerca de presas de hielo en Marte).

Un marco abstracto no es tanto una reproducción de la realidad, como el establecimiento de un punto de vista que guía nuestro pensamiento. Responder al teléfono no es una actividad guiada por un objetivo en el sentido de resolver un problema. Aún así nosotros estamos abstrayendo, gracias a lo cual podemos darle sentido a una situación concreta y construir un marco por el cual organizar nuestro comportamiento. Se produce un fenómeno de contextualización para lograr que una situación concreta adquiera sentido y que a su vez permita que nuevas acciones puedan resultar de ella. El pensamiento abstracto es una actividad cultural que construye una comprensión de la realidad. La relación entre percepción, acción y concepción es una realización dialéctica a nivel neuronal. La abstracción constriñe nuestro comportamiento y nos ofrece una herramienta para

acometer tareas que llegan a estar definidas dentro de una concepción global de la actividad.

Ciertas abstracciones pueden ser derivadas de instancias simples bajo condiciones adecuadas, una vez abstraídas son impuestas a otras instancias, es decir, son usadas interpretativamente siempre que sean útiles. La cuestión de cuáles son las abstracciones válidas es una cuestión muy problemática. Una simple selección arbitraria de rasgos no tiene porqué producir una abstracción relevante.

2.5 ABSTRACCIÓN EN LAS CIENCIAS DE LA COMPUTACIÓN

Abstracción es un término básico en Ciencias de la Computación y es común hacerlo explícito durante el currículo en los diferentes planes de estudio. Usualmente es comprendida como una ocultación de información, por ejemplo al enseñar programación e insistir en la ocultación de los detalles de implementación de los distintos constructores como pueden ser clases o métodos en la Programación Orientada a Objetos. También es habitual considerar diferentes niveles de abstracción en Computación, tanto hablando del hardware como del software o de las etapas de desarrollo de los mismos. En todos los casos el objetivo principal es reducir la complejidad.

En Computación también se han utilizado distintas técnicas basadas en mecanismos de abstracción para la resolución de problemas. Básicamente la idea principal es representar un problema en una nueva representación que en esencia respeta a la original pero que es en algún sentido más simple [13]. No es de extrañar pues que se hayan investigado múltiples aplicaciones de estas ideas en diferentes campos de la Ciencia de la Computación desde hace ya mucho tiempo. Estos incluyen prueba de teoremas [11, 17], resolución de problemas y planificación [19, 12], verificación de modelos [5], análisis de programas [6] y muchos más. Existe un simposium sobre Abstracción, Reformulación y Aproximación (SARA) que va por su novena edición y en el listado de sus aplicaciones, además de los ya mencionados, aparecen campos como programación automática, modelización cognitiva, programación por restricciones, satisfacción de restricciones, bases de datos, diseño, diagnóstico, juegos, aprendizaje automático, robótica, búsqueda o simulación.

Hay intentos de construir una teoría unificada de la abstracción en computación [10, 14], pero quizá es la teoría de interpretación abstracta desarrollada por P. y R. Cousot [6, 7] la que con más profundidad ha sido estudiada en el campo del análisis estático de programas. En ella, se utilizan las conexiones de Galois [9] como el mecanismo para modelar la abstracción. Las conexiones de Galois son un mecanismo para formalizar el camino de ida y vuelta a través de una abstracción que esencialmente permite aprovecharla para lograr una solución al problema original. En la figura 1 puede verse una abstracción dada por una conexión de Galois. En ella puede verse como el camino de ida y vuelta permite evaluar una aplicación compleja f mediante la composición $\gamma \circ f^* \circ \alpha$ de una aplicación f^* más sencilla gracias a que se establece en el espacio abstracto y a la utilización de los adjuntos superior e inferior de la conexión de Galois. La mayoría de aplicaciones de la abstracción en Computación, al menos en aquellas que se utiliza explícitamente el término abstracción para describirlas, son aplicaciones simbólicas, basadas en diferentes lógicas y/o en estructuras reticulares. De hecho el concepto de aproximación en este contexto (por ejemplo en SARA) es utilizado como el mecanismo por el cual se manipula y usa una abstracción para razonar acerca del problema original. Es decir, el uso de un mecanismo de abstracción permite un mecanismo de aproximación más sencillo al problema. Sin embargo lo recíproco también es válido, cualquier proceso de aproximación conlleva una abstracción, en el sentido de eliminar detalles. Así es posible ver como mecanismos de abstracción la utilización de medidas de agregación, o los mecanismos de multiresolución.

3 ABSTRACCIÓN Y LÓGICA FUZZY

En el campo de la teoría de los conjuntos fuzzy no es habitual encontrar la palabra abstracción para describir alguna de sus características. Sí lo hizo Zadeh en alguno de sus primeros trabajos [23, 21, 2] y sin embargo no aparece apenas posteriormente. Hay que tener en cuenta además que Zadeh en esos trabajos utiliza el término abstracción para referirse al proceso de inferir un fuzzy set a partir de un conjunto de datos inicial y su posterior generalización mediante la aplicación del fuzzy set a nuevos objetos. Tampoco es fácil encontrar un trabajo del ámbito fuzzy en las actas de los simposium de SARA o en relación con el concepto de abstracción en computación.

Sin embargo la teoría de conjuntos fuzzy presenta muchas características que la convierten en un campo especialmente adecuado para aplicar mecanismos de abstracción. Veamos alguno de ellos.

- Los conjuntos fuzzy tienen como primer objetivo representar conceptos borrosos, vagos. El uso de conceptos y de palabras es una de las características del pensamiento humano y que manifiestan más claramente su capacidad de abstracción. Por supuesto la capacidad de representar este tipo de conceptos ha sido profusamente explotada en los sistemas de reglas borrosas o en métodos de clustering fuzzy, por poner algunos ejemplos. Sin embargo el uso de conjuntos fuzzy en ciencias como la Psicología o la Lingüística ha sido minoritario, apenas unos pocos trabajos lo han intentado y rápidamente se han abandonado. Recientemente el trabajo [3] ha mostrado cómo malas concepciones e interpretaciones de la teoría han ejercido

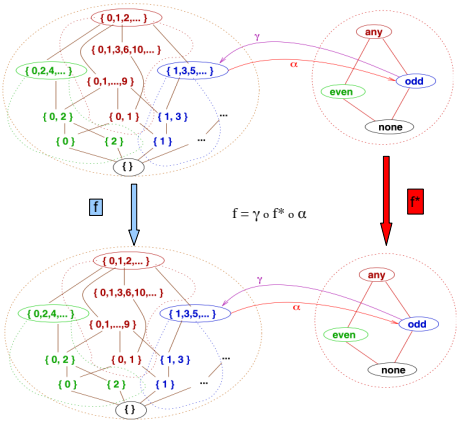


Figura 1: Ejemplo de una abstracción dada por una conexión de Galois

una influencia excesiva y han propiciado el abandono de la línea de trabajo fuzzy en esos campos. Y sin embargo la teoría ofrece algo que nunca antes se había producido, la posibilidad de trabajar con el concepto de forma simbólica (con la palabra) y con su significado dado por una función.

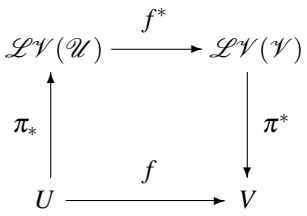


Figura 2: Diagrama que muestra el mecanismo de abstracción en los sistemas de reglas fuzzy

- Los sistemas de reglas fuzzy son un ejemplo cristalino del camino de ida y vuelta que se produce al utilizar mecanismos de abstracción. Como aparece en la figura 2, los sistemas de reglas fuzzy permiten construir una aplicación compleja f mediante la construcción de una aplicación mucho más sencilla f^* definida en los espacios abstractos $\mathscr{LV}(\mathscr{U}), \mathscr{LV}(\mathscr{V})$ de etiquetas lingüísticas perteneciente a las clases de conjuntos fuzzy $\mathscr{F}(\mathscr{U}), \mathscr{F}(\mathscr{V})$ de sus respectivos universos. Las aplicaciones π_*, π^* son respectivamente la fuzzyficación y la defuzzyficación definidas habitualmente en este tipo de sistemas.

  Es gracias a la facilidad en la definición de f^* por la que estos sistemas pueden ser construidos de una forma, en ocasiones, tan sorprendentemente sencilla. Y es que f^* se define entre términos lingüísticos que son conceptos abstractos definidos para los universos U y V.

- La teoría de conjuntos fuzzy es según L.A. Zadeh la base del Granular Computing [22]. Otras teorías como los Rough Sets [16], Retículos Formales de Conceptos [18] o Espacios de Aproximación [20] son técnicas para definir granulaciones y lograr aproximaciones gracias a ellas. En muchas de estas teorías son las relaciones de equivalencia las que permiten definir los gránulos que particionan el universo.

  Un mecanismo básico, y quizá el más evidente, de definir abstracciones es mediante cualquier aplicación epiyectiva, una aplicación que envía muchos a pocos. En la teoría de conjuntos clásica, las relaciones de equivalencia, las aplicaciones epiyectivas y las particiones son conceptos equivalentes, pueden definirse unos a través de los otros. En cambio, en la teoría de conjuntos fuzzy, ésto no es así, particiones fuzzy, relaciones de similaridad o su generalización indistinguibilidades, y aplicaciones epiyectivas no son conceptos equivalentes. Se plantea pues el problema de cuáles son los métodos mejores para definir buenas abstracciones.

4 CONCLUSIONES

En el trabajo se ha pretendido mostrar como la teoría de conjuntos fuzzy contiene en sus fundamentos diversos mecanismos de abstracción y que este hecho no ha sido lo suficientemente destacado ni en el contexto fuzzy ni en el contexto de las técnicas clásicas de abstracción en el campo de la computación. Los numerosos estudios existentes sobre la abstracción pueden ofrecer una fuente valiosa donde descubrir y profundizar en las técnicas propias de la teo-

ría fuzzy y permitir avanzar en el campo del Computing With Words que en este contexto puede interpretarse como Computación con Abstracciones.

Referencias

[1] Rudolfo Arnheim. *Visual Thinking*. University of California Press, 1969.

[2] R. Bellman, R. Kalaba, and L. Zadeh. Abstraction and pattern classification. In George J. Klir and Bo Yuan, editors, *Fuzzy sets, fuzzy logic, and fuzzy systems*, pages 44–50. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1996.

[3] Radim Belohlavek, George Klir, Harold Lewis Iii, and Eileen Way. Concepts and fuzzy sets: Misunderstandings, misconceptions, and oversights. *International Journal of Approximate Reasoning*, 2009.

[4] William J. Clancey. Is abstraction a kind of idea or how conceptualization works? commentary on special issue of cognitive science quarterly, 2000.

[5] Edmund M. Clarke, Orna Grumberg, and David E. Long. Model checking and abstraction. *ACM Trans. Program. Lang. Syst.*, 16:1512–1542, September 1994.

[6] P. Cousot and R. Cousot. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Conference Record of the Fourth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 238–252, Los Angeles, California, 1977. ACM Press, New York, NY.

[7] Patrick Cousot. Abstract interpretation. *ACM Comput. Surv.*, 28:324–328, June 1996.

[8] Ed Dubinsky. Reflective abstraction in advanced mathematical thinking. In David Tall, editor, *Advanced Mathematical Thinking*, volume 11 of *Mathematics Education Library*, pages 95–126. Springer Netherlands, 1991.

[9] M Erné, J Koslowski, A Melton, and GE Strecker. A primer on galois connections. *Annals of the New York Academy of Sciences*, 704(Papers on General Topology and Applications):103–125, 1993.

[10] Fausto Giunchiglia and Toby Walsh. A theory of abstraction. *Artif. Intell.*, 57(2-3):323–389, 1992.

[11] C. Green. Application of theorem proving to problem solving. In *Proc. 1st Int. Joint Conf. Artificial Intelligence*, pages 219–239, 1969. Perhaps only cited because is a classic and because it is very old.

[12] Craig A. Knoblock, Josh D. Tenenberg, and Qiang Yang. Characterizing abstraction hierarchies for planning. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, volume 2, pages 692–697, Anaheim, California, USA, 1991. AAAI Press/MIT Press.

[13] Kim Marriott. *Abstract interpretation: A theory of approximate computation*, volume 1302 of *Lecture Notes in Computer Science*, pages 367–378. Springer-Verlag, 1997.

[14] P. Pandurang Nayak and Alan Levy. A semantic theory of abstractions. In Chris Mellish, editor, *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 196–203, San Francisco, 1995. Morgan Kaufmann. To read.

[15] S. Ohlsson and S. Regan. A function for abstract ideas in conceptual discovery and learning. *Cognitive Science Quarterly*, 1(2):243–277, 2001.

[16] Zdzisław Pawlak. Rough sets. *International Journal of Parallel Programming*, 11:341–356, 1982. 10.1007/BF01001956.

[17] David A. Plaisted. Theorem proving with abstraction. *Artificial Intelligence*, 16(1):47–108, 1981. Perhaps only cited because is a classic.

[18] Rudolf and Wille. Concept lattices and conceptual knowledge systems. *Computers and Mathematics with Applications*, 23(6-9):493 – 515, 1992.

[19] Earl D. Sacerdoti. Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5(2):115–135, 1974. Perhaps only cited because is a classic and because it is very old.

[20] Andrzej Skowron, Roman Świniarski, and Piotr Synak. Approximation spaces and information granulation. In James Peters and Andrzej Skowron, editors, *Transactions on Rough Sets III*, volume 3400 of *Lecture Notes in Computer Science*, pages 91–95. Springer Berlin / Heidelberg, 2005.

[21] L. A. Zadeh. Fuzzy sets and systems. In J. Fox, editor, *System Theory*, pages 29–39. Brooklyn Polytechnic Press, 1965.

[22] L. A. Zadeh. Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Computing*, 2, 1998.

[23] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

OPTIMIZACIÓN DE MODELOS ESTADÍSTICOS Y DIFUSOS PARA EL ANÁLISIS DE SERIES TEMPORALES MEDIANTE EVOLUCIÓN DIFERENCIAL

Christoph Bergmeir[1], **Isaac Triguero**[1], **Francisco Velasco**[2], **José Manuel Benítez**[1]

[1]*Departamento de Ciencias de la Computación e Inteligencia Artificial, CITIC-UGR, Universidad de Granada (España),* {c.bergmeir,triguero,j.m.benitez}@decsai.ugr.es

[2]*Departamento de Lenguajes y Sistemas Informáticos, CITIC-UGR, Universidad de Granada (España),* fvelasco@ugr.es

Resumen

En este trabajo se presenta un procedimiento para ajustar el modelo autorregresivo de transiciones suaves con coeficientes neuronales, propuesto por Medeiros y Veiga (2005, IEEE Trans. NN, 16(1):97–113). El procedimiento de construcción de este modelo sigue un enfoque iterativo basado en test estadísticos. Además, los modelos que se obtienen son interpretables en términos de sistemas basados en reglas difusas. La interpretabilidad de los modelos y una metodología matemáticamente justificada son unas de las propiedades más importantes en una aplicación real. Sin embargo, el ajuste de los parámetros de este método se estima a través de una búsqueda de rejilla en combinación con un algoritmo de búsqueda local tradicional. En este trabajo se propone un procedimiento de ajuste alternativo, basado en un algoritmo de evolución diferencial, con objeto de obtener modelos más precisos. El procedimiento se ha evaluado empíricamente, aplicándolo a distintas series temporales reales. Los resultados indican que se mejora significativamente la precisión de los modelos, haciéndolos competitivos cuando se comparan con otros métodos utilizados habitualmente en este ámbito.

Palabras Clave: series temporales, modelos estadísticos, optimización, evolución diferencial.

1 INTRODUCCIÓN

La predicción y el modelado de series temporales es un campo interdisciplinar que involucra a muchas áreas del conocimiento en cuanto a su aplicación y tiene su génesis en la Estadística. Sin embargo, las técnicas de modelado desarrolladas en el contexto de la Inteligencia Artificial, han encontrado en este campo un área fértil para su adaptación y explotación. Los esquemas estadísticos clásicos para el modelado de series temporales tienen su referencia fundamental en los trabajos de Box y Jenkins [6] en los años 1970. El modelo de referencia es el modelo autorregresivo de medias móviles (autoregressive moving average, ARMA) que se caracteriza por establecer relaciones lineales entre valores retrasados de la serie mediante una regresión lineal. Sin embargo, en la actualidad disponemos de una gran cantidad de información lo que implica que los patrones que se pueden extraer de dichos datos son más complejos y no necesariamente lineales. Nace pues la necesidad de modelar comportamientos no lineales. Uno de los modelos no lineales más conocido es el desarrollado por Tong [17] en el que se permite modelar una serie mediante dos o más regímenes lineales adscritos a zonas disjuntas del espacio de la serie. Se usa una variable de umbral para establecer la transición entre los distintos modelos lineales. En la literatura se encuentran numerosas variantes de esta idea, conocidas genéricamente como modelos autorregresivos (AR) de umbral.

Dentro de estos modelos, se denomina AR de umbral autoexcitativo a aquél en el que la variable de umbral es un valor retrasado de la serie. Por otro lado, considerando que la naturaleza de las series son generalmente continuas, se denotan AR de umbral con transiciones suaves a aquellos modelos que sustituyen el cambio brusco de un modelo lineal a otro por una función de transición suave.

Usando esta teoría, Medeiros y Veiga [10] desarrollaron el AR de umbral coeficientes neuronales (NCSTAR), que utiliza una red neuronal para aprender la variable umbral como una suma ponderada de las entradas a partir de los datos de entrenamiento. Además del modelo, los autores presentaron un procedimiento de construcción de tipo iterativo, basado en técnicas estadísticas.

Es importante destacar que cada uno de estos modelos AR de umbral se puede expresar en términos de un sistema basado en reglas difusas (FRBS) [1, 3, 2, 4]. Esto supone una ventaja clara de cara la interpretabilidad de los mismos.

A pesar de estas ventajas, el modelo NCSTAR se basa en una búsqueda de rejilla en combinación con una búsqueda local como método para optimizar sus parámetros. Éstos pueden ser determinantes durante el proceso incremental del NCSTAR, influyendo especialmente en la precisión del método. Por otro lado, los algoritmos evolutivos [8] han demostrado ser técnicas muy eficientes para abordar problemas de optimización en distintos ámbitos [18]. Concretamente, el algoritmo de evolución diferencial (ED) [16] destaca como uno de los métodos más prometedores en problemas de optimización continua como el que se plantea al optimizar los parámetros del NCSTAR.

El objetivo de este trabajo es el de proponer un procedimiento de aprendizaje y ajuste para NCSTAR basado en un algoritmo evolutivo, con la intención de obtener modelos que preserven la interpretabilidad natural del modelo pero que además sean más precisos.

El resto del trabajo se organiza como sigue: La Sección 2 desarrolla la teoría de modelos de umbral autorregresivo. La Sección 3 describe las características más importantes de la ED. La Sección 4 presenta el algoritmo propuesto que utiliza la evolución diferencial para la construcción de modelos NCSTAR. La Sección 5 muestra los experimentos realizados, y finalmente la Sección 6 concluye el trabajo.

2 MODELOS AUTORREGRESIVOS DE UMBRAL

Definimos el modelo autorregresivo para una serie temporal x, con una función F y una serie e de errores que son independientes e idénticamente distribuidos (i.i.d.), de la siguiente manera:

$$x(t) = F(x(t-1),\ldots,x(t-d)) + e(t) \tag{1}$$

El parámetro de retardo d determina el número valores retrasados que se utilizan como entrada para el método. A partir de esta definición general, se pueden desarrollar distintos modelos para series temporales, según la elección de F.

En una autorregresión lineal, F realiza una combinación lineal de los valores anteriores:

$$x(t) = a_0 + \sum_{i=1}^{d} a_i x(t-i) + e(t) \tag{2}$$

En notación vectorial, con $\mathbf{z}(t) = (1,\ x(t-1),\ x(t-2),\ldots,\ x(t-d))^T$ y $\mathbf{a} = (a_0,\ldots,a_d)$, la Ecuación (2) se convierte en:

$$x(t) = \mathbf{a}\mathbf{z}(t) + e(t) \tag{3}$$

Cuando F es una función no lineal, un método común es la combinación de modelos lineales:

$$x(t) = \sum_{j=1}^{k} f_j(th_j(t))\mathbf{a}_j\mathbf{z}(t) + e(t) \tag{4}$$

Las dos cuestiones que surgen tras esta idea son: (1) qué funciones no lineales f_j se utilizan para combinar, y (2) cómo se compone la función de umbral $th_j(t)$. En la función de umbral se puede, por ejemplo, tener en cuenta las variables exógenas, los valores retrasados, o combinaciones de ambos. En el modelo umbral AR *(threshold AR, TAR)*, las funciones f_j son elegidas para ser funciones índice I_j que cambian entre modelos lineales diferentes, dependiendo del valor actual del umbral $th_j(t)$, usando constantes umbral $c_0\ldots,c_k$ con $-\infty = c_0 < c_1 < \ldots < c_k = \infty$, de la siguiente forma:

$$I_j(th_j(t)) = \begin{cases} 1, & \text{si } th_j(t) \in (c_{j-1}, c_j]; \\ 0, & \text{en caso contrario.} \end{cases}$$

En el modelo TAR auto-excitativo *(self-exciting TAR, SETAR)*, $th_j(t) := x(t-d)$. Como la función de índice provoca cambios bruscos, que podría no ser deseable, otra posibilidad es utilizar la función logística, es decir,

$$f_j(th_j(t)) = (1 + exp(-\gamma_j(th_j(t) - c_j)))^{-1}. \tag{5}$$

Esto define el modelo STAR logístico (LSTAR). Existen otras posibilidades para elegir las funciones f_j y $th_j(t)$. Una de las opciones más relevantes, es una modificación del modelo LSTAR como:

$$th_j(t) = \omega_j \mathbf{w}(t). \tag{6}$$

En este caso $\mathbf{w}(t)$ es un vector que contiene todas las variables que son relevantes para el umbral, es decir, los valores retrasados y/o variables exógenas. En el siguiente, vamos a utilizar $\mathbf{w}(t) = (x(t-1),\ x(t-2),\ldots,\ x(t-d))^T$. Y ω_j es un vector de pesos con $\|\omega_j\| = 1$, que tiene la misma longitud que $\mathbf{w}(t)$. A esto se le llama modelo STAR con coeficientes neuronales, NCSTAR.

El modelo NCSTAR tiene algunas propiedades interesantes. Medeiros y Veiga [10] presentan un método de construcción iterativo basado en tests estadísticos para este tipo de modelo. Aznarte y Benítez [1] demuestran, que los modelos NCSTAR son funcionalmente equivalentes a FRBS TSK aditivos con función de pertenencia logística.

De este modo, con el modelo NCSTAR tenemos un modelo basado en un proceso de construcción iterativo eficiente, y que puede ser interpretado en términos de un sistema basado en reglas difusas.

En el modelo NCSTAR, los parámetros se pueden dividir en parámetros lineales y no lineales. Después de haber fijado los parámetros no lineales, los parámetros lineales se pueden calcular en una solución en forma cerrada. En la versión original de NCSTAR, una combinación de búsqueda en rejilla y de búsqueda local se utiliza para determinar buenas opciones para los parámetros no lineales. Los parámetros no lineales son γ_j, c_j, y ω_j con $j = 1,\ldots,k$ como en las Ecuaciones (5) y (6) con $\omega_{ij} \in [-1,1]$ y $\|\omega_j\| = 1$. Según Aznarte y Benítez [1], estos k regímenes se pueden traducir a k reglas difusas equivalentes. Siempre que durante el proceso iterativo se añada una nueva regla, los valores iniciales de esta regla j se eligen de la siguiente forma [10]:

- ω_j se obtiene aleatoriamente a partir de una distribución uniforme, y se normaliza después para asegurar que tiene longitud uno. Si $\omega_{1j} < 0$, $\omega_j := -\omega_j$. Esto se realiza M veces, para obtener m vectores ω_j^m.
- Para cada ω_j^m, c_j se define como $mediana(\omega_j^m \mathbf{x})$, con $\mathbf{x}$ siendo la serie empotrada.
- Una rejilla de N valores se define para elegir γ_j^n (con $n = 1,\ldots,N$) como los valores $1,\ldots,N$.

Así se generan un total de $N \times M$ soluciones candidatas para la nueva regla difusa. Dado que γ_j, en teoría, es un parámetro que tiene que ser a escala con la serie temporal, la serie se normaliza antes de la construcción de modelos NCSTAR. La regla que, en combinación con las reglas ya presentes, permita una mayor reducción del error es elegida, y el algoritmo de búsqueda local se inicia a partir de esta configuración inicial, para mejorar todos los valores del modelo actual. En la publicación original, se recomiendan los valores $M = 1.000$ y $N = 20$, y el algoritmo de Levenberg-Marquardt [11] se utiliza para la búsqueda local.

3 EVOLUCIÓN DIFERENCIAL

El algoritmo de evolución diferencial (ED) [13] comienza con una población de NP individuos. La población inicial debe cubrir el espacio de búsqueda tanto como sea posible, inicializando los individuos aleatoriamente. Las generaciones siguientes se denotan como: $G = 0,1,\ldots,G_{max}$. En ED es usual denominar a cada individuo como un vector D-dimensional $X_{i,G} = \{x_{i,G}^1,..., x_{i,G}^D\}$, llamado vector objetivo.

Después de la inicialización, el algoritmo ED aplica un operador de mutación para generar un vector mutado $V_{i,G}$ respecto a cada individuo $X_{i,G}$, en la población actual. Para cada vector objetivo $X_{i,G}$, su vector mutado asociado es $V_{i,G}= \{V_{i,G}^1,..., V_{i,G}^D\}$.

En este trabajo nos centramos en el uso de la estrategia *DE/Rand/1* que genera un vector mutado de la siguiente forma:

$$V_{i,G} = X_{r_1^i,G} + F \cdot (X_{r_2^i,G} - X_{r_3^i,G}) \quad (7)$$

Donde los índices r_1^i, r_2^i y r_3^i son enteros mutuamente exclusivos y generados aleatoriamente en el rango $[1,NP]$, y son también diferentes del índice base i. El factor de escala F es un parámetro de control positivo para escalar la diferencia entre vectores.

Tras la fase de mutación, el operador de cruce se aplica para cada par, vector objetivo $X_{i,G}$ y su correspondiente vector mutado $V_{i,G}$, para generar un nuevo vector, denominado vector de "prueba" $U_{i,G}$. Nos centraremos en el uso del operador de cruce binomial, inicializado con los valores del vector objetivo, y en el cual cada componente del vector de prueba se modifica con los valores del vector mutado si al generar un número aleatorio entre 0 y 1, éste es menor o igual que un ratio de cruce RC. Finalmente, debemos decidir qué individuo pasa a la siguiente generación $G+1$. Si el vector de prueba obtiene una solución igual o mejor que el vector objetivo, éste reemplaza su correspondiente vector objetivo en la siguiente generación; en otro caso, el vector objetivo se mantiene en la población.

4 NCSTAR AJUSTADO CON EVOLUCIÓN DIFERENCIAL

La propuesta que se realiza en este trabajo consiste en reemplazar a la combinación de búsqueda local y rejilla del NCSTAR original por un algoritmo de ED. Para ello, el algoritmo ED requiere algunas particularizaciones:

Los individuos de la población del algoritmo evolutivo se forman de vectores $X = \{\gamma_1,\ldots,\gamma_k,c_1,\ldots,c_k,\omega_{11},\ldots,\omega_{d1},\ldots,\omega_{1k},\ldots,\omega_{dk}\}$, que son realizaciones concretas de los parámetros no lineales γ_j, c_j, y ω_j con $j = 1,\ldots,k$ del NCSTAR.

El proceso para la construcción de los modelos es el siguiente. Se aplica el test de linealidad, si éste supone que la serie es lineal, se construye un modelo lineal (un modelo NCSTAR de un régimen es un modelo lineal). De lo contrario, se inicia el proceso iterativo de construcción del modelo. En la iteración k-ésima de este proceso, se construye un modelo NCSTAR de $(k+1)$ regímenes (o reglas difusas) de la siguiente manera:

1. En la población inicial de la ED se incluye un individuo compuesto de las reglas del modelo actual junto con una regla nueva inicializada al azar. El resto de la población inicial se genera aleatoriamente conforme a una distribución uniforme dentro de los respectivos dominios de los parámetros (ver abajo).
2. Los parámetros no lineales para todas las k transiciones se fijan con el algoritmo de optimización (los pa-

rámetros lineales se fijan en cada evaluación de la función de error).

3. Se aplica el test estadístico de linealidad de los residuos.
4. Si el test indica que es necesario añadir una nueva regla, se vuelve al paso (1.).
5. En caso contrario, el método termina.

Es importante acotar adecuadamente los valores de γ_j y c_j. Una regla que sólo es relevante para pocos datos de entrenamiento conduce a problemas numéricos y parámetros lineales ilógicos, lo que puede dar lugar a predicciones inesperadas. Además, si el valor de γ es alto, se pierde la suavidad de la transición y el NCSTAR degenera en un modelo TAR.

Debido a las razones indicadas anteriormente, se aplica las siguientes restricciones sobre los parámetros no lineales:

- El dominio de los γ_j: $\gamma_j \in [0, (max(x) - min(x)) \cdot \gamma_0]$, con x siendo la serie temporal y γ_0 siendo un parámetro del método.
- El dominio de los c_j: Los valores c_j se restringen al intervalo $[min(x), max(x)]$.
- La codificación y el dominio de los ω_j: Con el fin de manejar la restricción $\|\omega_j\| = 1$, ω_j está codificado con coordenadas polares de dimensión n [5], por lo que $\omega_{ij} \in [0, \pi]$ para $i = 1, \ldots, (d-1)$, y $\omega_{dj} \in [0, 2\pi]$.

5 ESTUDIO EXPERIMENTAL Y ANÁLISIS DE RESULTADOS

Para analizar la eficacia del método de ajuste y aprendizaje propuesto, hemos diseñado y ejecutado un estudio experimental. En esta sección se detallan los experimentos que se han llevado a cabo. Se comentan las series temporales utilizadas, los algoritmos con los que se ha comparado, los parámetros de configuración y los resultados obtenidos.

5.1 SERIES TEMPORALES UTILIZADAS

En este estudio, utilizamos datos provenientes de la competición NNGC1[1]. Concretamente, hemos usado las series con mayor frecuencia, es decir, se utilizan los datos semanales, diarios y por horas del NNGC1. Los datos semanales son datos relacionados con la industria del petróleo, como volumen de las importaciones o precios. Las series diarias son medidas de rendimientos de túneles de coches, y los datos horarios son tiempos de llegada de aviones en aeropuertos y de trenes en estaciones de metro.

[1] `http://www.neural-forecasting-competition.com/datasets`

En total, se utilizan 15 series, para las cuales se asegura que son estacionarias y tienen un comportamiento no lineal, mediante la aplicación del test de Dickey-Fuller [15], y un test de linealidad, que es parte del procedimiento de construcción de NCSTAR [10]. Esto se hace porque los modelos *TAR, al igual que los modelos ARMA, sólo son aplicables a series estacionarias [17]. Además, en el caso de una serie lineal, en lugar de un modelo NCSTAR, se construye un modelo AR lineal. En ese caso, no hay parámetros no lineales que puedan estimarse, por lo que el procedimiento de optimización no se ejecuta. Para los experimentos, usamos el 20 % de los datos del final de cada serie como conjunto de prueba. El resto de los datos se utiliza para la construcción de los modelos. Todas las series consideradas están disponibles en el repositorio KEEL-dataset[2].

5.2 ALGORITMOS Y PARÁMETROS

Los experimentos se han llevado a cabo con el lenguaje de programación R [14]. En lugar del algoritmo Levenberg-Marquardt, se utilizó el algoritmo BFGS, que está disponible en R a través de la función `optim`. Para analizar los efectos de los dominios de los parámetros discutidos en la Sección 4, se utiliza aparte del método original, una versión con un algoritmo de optimización con restricciones, el algoritmo L-BFGS-B, que está disponible a través de la misma función `optim` (ver Venables y Ripley [19]). Llamamos a estos dos algoritmos `NCSTAR` y `NCSTAR-BOX`.

La versión del algoritmo que usa ED para la optimización, la llamamos `NCSTAR-DE`. Utilizamos una implementación de ED que está disponible en el paquete `RcppDE` de R (es una reimplementación de la función `DEoptim` [12]).

Los métodos originales fueron utilizados con $N = 140$, y $M = 2.000$. Los $N = 140$ valores de γ se eligen equidistantes dentro del dominio de γ tal como se define en la Sección 4, con $\gamma_0 = 20$ en todos nuestros experimentos. El algoritmo de búsqueda local se ejecuta con un máximo de iteraciones de 20.000. Por lo tanto, en total se ejecuta aproximadamente 300.000 evaluaciones de la función de error para cada regla nueva.

Para realizar una comparación justa, los algoritmos ED se utilizan con la misma cantidad de evaluaciones de la función de error. La ED se utiliza con los parámetros F = 0,5 y RC = 0,5. Aunque en experimentos preliminares se evaluaron otros conjuntos de parámetros, se descubrió que estos valores son una buena elección y ofrecen los mejores resultados para nuestro problema. El tamaño de la población se establece en 30 individuos a lo largo de todos los experimentos, y se calculan 10.000 generaciones (esto produce también un total de 300.000 evaluaciones de la función de error para cada regla nueva).

Además de comparar entre sí distintos modelos NCSTAR,

[2] `http://sci2s.ugr.es/keel/timeseries.php`

se realizó también un estudio comparativo del método propuesto con otros métodos comúnmente empleados para la predicción de series temporales. En concreto, utilizamos máquinas de vectores de soporte ε (SVR), un perceptrón multicapa (MLP) entrenado con el algoritmo BFGS, y un modelo lineal. SVR está disponible en R a través del paquete `e1071`, que implementa un wrapper para la LIBSVM [7]. El MLP entrenado con BFGS está disponible a través del paquete `nnet` [19].

Los parámetros de los métodos se han establecido con 30 series temporales artificiales generadas de un proceso NCSTAR. En primer lugar, una rejilla de parámetros se define empíricamente. Entonces, el conjunto de parámetros de mejor desempeño en todas estas series, respecto al error cuadrático medio *(root mean squared error, RMSE)* en el conjunto de prueba, es elegido. Los parámetros que finalmente se utiliza durante los experimentos son cost = 10, gamma = 0,2, epsilon = 0,1 para la `SVR`, y size = 9, decay = 0,1, maxit = 1.000 para el `MLP`. El modelo `AR` no tiene parámetros libres.

5.3 RESULTADOS OBTENIDOS Y ESTUDIO ESTADÍSTICO

Se construyó un predictor para cada una de las 15 series usando cada uno de los modelos enumerados. Después, se evalúa la calidad de la predicción usando el RMSE sobre el conjunto de prueba. La Tabla 1 muestra los resultados obtenidos. Para cada serie se muestra el RMSE obtenido por cada algoritmo. Además, se ha remarcado en negrita el mejor resultado obtenido en cada serie. Los resultados muestran que `NCSTAR-DE` obtiene por medio un error más pequeño comparado con los otros métodos. Sin embargo, esta medida puede estar afectada por la presencia de *outliers*. Para complementar el estudio, hace falta analizar la frecuencia con que el método obtiene un error más pequeño. Por eso, se van a emplear tests no paramétricos de comparaciones múltiples para contrastar los resultados experimentales.

Específicamente, se ha seleccionado el test de Friedman como método para detectar la existencia de diferencias significativas entre los resultados, y el test de Holm para caracterizar las diferencias encontradas [9][3]. El test de Friedman detecta diferencias significativas ($valorp < 0.005$) sobre los resultados. A partir de los rangos obtenidos, se selecciona a `NCSTAR-DE` como método de control (aquel que obtiene el rango más bajo), y se aplica el test *post-hoc*. La Tabla 2 resume los resultados del estudio estadístico.

En combinación, las dos evaluaciones muestran que `NCSTAR-DE` es el mejor método, tanto en la frecuencia de obtener un error más pequeño que los otros métodos, como en el valor absoluto de este error.

[3]Más información en el sitio web temático de SCI2S sobre *Statistical Inference in Computational Intelligence and Data Mining* `http://sci2s.ugr.es/sicidm/`

Método	Rangos	$p-valor$ de Holm
NCSTAR-DE	1,87	–
NCSTAR-BOX	3,40	0,0383
NCSTAR	3,47	0,0383
MLP	3,80	0,0140
SVR	4,00	0,0072
AR	4,47	0,0007

Tabla 2: Test estadístico no paramétrico

6 CONCLUSIONES

Se ha investigado sobre el uso de ED para optimizar los parámetros en el procedimiento de construcción iterativo del algoritmo NCSTAR, que se puede interpretar como un sistema de reglas difusas. En cada iteración, un test estadístico determina si es necesario añadir más reglas o no. Si se añade una nueva regla, todo el sistema se reajusta usando un método de optimización basado en ED.

La combinación del algoritmo NCSTAR, caracterizado por ser un modelo interpretable, y la ED como método para ajustar sus parámetros ha demostrado generar modelos significativamente más precisos que los métodos originales.

Con el uso de un algoritmo de optimización evolutivo el NCSTAR se hace competitivo con otros procedimientos comúnmente empleados en la predicción de series temporales. Por lo tanto, con NCSTAR-DE, obtenemos un algoritmo que se puede utilizar para crear modelos precisos e interpretables para series temporales.

Debido a la estrecha relación existente entre los modelos TAR, en general, y NCSTAR, en particular y los sistemas basados en reglas difusas es fácil colegir que la propuesta de método de aprendizaje y ajuste presentada en este trabajo podría aplicarse con éxito en el aprendizaje directo de sistemas basados en reglas difusas para el modelado de series temporales.

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto TIN-2009-14575 del Ministerio de Ciencia e Innovación (MICINN). C. Bergmeir disfruta de una beca del Programa de Formación del Profesorado Universitario (FPU) del Ministerio de Educación (MEC). I. Triguero dispone de una beca FPU de la Universidad de Granada.

Referencias

[1] J.L. Aznarte and J.M. Benítez. Equivalences between neural-autoregressive time series models and fuzzy systems. *IEEE Transactions on Neural Networks*, 21(9):1434–1444, 2010.

[2] J.L. Aznarte, J. Manuel Benítez, and J. Luis Castro. Smooth transition autoregressive models and fuzzy

Serie temporal	NCSTAR-DE	NCSTAR	NCSTAR-BOX	AR	MLP	SVR
TS-01	0,193	0,200	0,196	0,184	**0,181**	0,200
TS-02	0,538	0,535	0,538	**0,417**	0,535	0,661
TS-03	0,903	0,912	0,906	**0,813**	1,012	0,938
TS-04	**0,545**	0,617	0,692	0,855	0,551	0,586
TS-05	0,504	**0,493**	0,559	0,851	0,526	0,530
TS-06	0,455	0,486	0,461	0,426	0,463	**0,419**
TS-07	1,042	2,435	1,323	**0,967**	1,075	0,968
TS-08	**0,098**	0,103	0,120	0,472	0,145	0,159
TS-09	**0,176**	0,184	0,188	0,558	0,190	0,216
TS-10	0,120	**0,099**	0,142	0,460	0,155	0,171
TS-11	**0,142**	0,156	0,153	0,519	0,187	0,173
TS-12	**0,221**	0,333	0,236	0,639	0,284	0,320
TS-13	0,144	0,157	**0,132**	0,570	0,203	0,235
TS-14	**0,194**	0,335	0,261	0,579	0,274	0,255
TS-15	**0,153**	0,177	0,214	0,655	0,225	0,215
Promedio	**0,362**	0,482	0,408	0,598	0,400	0,403

Tabla 1: Resultados obtenidos

rule-based systems: Functional equivalence and consequences. *Fuzzy Sets and Systems*, 158(24):2734–2745, 2007.

[3] J.L. Aznarte, M.C. Medeiros, and J.M. Benítez. Linearity testing for fuzzy rule-based models. *Fuzzy Sets and Systems*, 161(13):1836–1851, 2010.

[4] J.L. Aznarte, M.C. Medeiros, and J.M. Benítez. Testing for remaining autocorrelation of the residuals in the framework of fuzzy rule-based time series modelling. *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, 18(4):371–387, 2010.

[5] L. E. Blumenson. A derivation of n-dimensional spherical coordinates. *The American Mathematical Monthly*, 67(1):63–66, January 1960.

[6] G. Box and G. Jenkins. *Time series analysis: Forecasting and control*. Holden-Day, 1970.

[7] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[8] A.E. Eiben and J.E. Smith. *Introduction to Evolutionary Computing*. Springer, Berlin, 2003.

[9] S. García, A. Fernández, J. Luengo, and F. Herrera. Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences*, 180(10):2044–2064, 2010.

[10] M.C. Medeiros and A. Veiga. A flexible coefficient smooth transition time series model. *IEEE Transactions on Neural Networks*, 16(1):97–113, 2005.

[11] J. J. Moré. The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical Analysis: Proceedings of the Biennial Conference*, pages 104–116. Springer, 1978.

[12] K.M. Mullen, D. Ardia, D.L. Gil, D. Windover, and J. Cline. Deoptim: An r package for global optimization by differential evolution. *Journal of Statistical Software*, 40(6):1–26, 2011.

[13] K.V. Price, R.M. Storn, and J.A. Lampinen. *Differential evolution: a practical approach to global optimization*. Natural computing series. Springer, 2005.

[14] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0.

[15] S. E. Said and D. A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71:599–607, 1984.

[16] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[17] H. Tong. *Non-linear Time Series: a dynamical system approach*. Oxford: Clarendon Press, 1990.

[18] I. Triguero, S. García, and F. Herrera. Ipade: Iterative prototype adjustment for nearest neighbor classification. *IEEE Transactions on Neural Networks*, 21(12):1984–1990, 2010.

[19] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer-Verlag, 2002.

SERIES TEMPORALES Y CLUSTERING DIFUSO. CASO DE ESTUDIO: CENTRO DE ATENCIÓN DE LLAMADAS DE EMERGENCIA

Rubén García[1], **Francisco Barrientos**[1], **José M. Benítez**[2], **Gregorio Sáinz**[3,1]

[1]*CARTIF Centro Tecnológico, Parque Tecnológico de Boecillo 205*, {frabar,rubgar}@cartif.es
[2]*Universidad de Granada, DECSAI*, jmbenitez@decsai.ugr.es
[3]*Universidad de Valladolid, E. I. Industriales, Dept. Ingeniería de Sistemas y Automática*, gresai@eii.uva.es

Resumen

En este artículo se describe la utilización de un Sistema Lógico Difuso generado a partir de clustering para resolver con éxito un caso real: la predicción de llamadas entrantes de un centro de atención de llamadas de emergencia. En concreto, se muestra como la selección de variables permite que el sistema lógico difuso se aplique con éxito para predecir series temporales. El rendimiento de esta técnica se compara con regresores bien conocidos: una red neuronal, MLP y modelos estadísticos de Box-Jenkins, AR y ARMA.

En un centro de atención de llamadas de emergencia el objetivo prioritario es la atención al ciudadano. El sistema de predicción de llamadas entrantes generado en este trabajo, permite adoptar mejores decisiones para la planificación de los recursos disponibles. Se busca un equilibrio entre calidad de servicio y eficiencia económica.

Para lograr este objetivo, únicamente se dispone del histórico de llamadas recibidas, que en este trabajo será modelado como una serie temporal.

Palabras Clave: selección de variables, predicción de series temporales, clustering difuso, relief-F, subtractive clustering, emergencias.

1 INTRODUCCIÓN

En la actualidad, los centros de atención de llamadas son habituales en diversas áreas y actividades: compañías, servicios sociales, seguridad, atención médica, etc. Cada centro debe dimensionarse para alcanzar un nivel de servicio de acuerdo a su actividad, priorizando entre aspectos como: nivel de atención, eficiencia en la resolución de incidencias, tiempo de respuesta o recursos económicos disponibles.

Por tanto, uno de los principales problemas que subyacen en la actividad de la gerencia de los centros de atención de llamadas es el de su dimensionamiento. El establecimiento de los recursos, tanto técnicos como humanos, necesarios para dar un servicio de calidad suficiente de acuerdo al nivel de servicio establecido no es un problema trivial. El número de operadores que debe estar disponibles en cada instante está directamente relacionado con el volumen de llamadas que va a ser necesario atender.

Un centro de tamaño medio-grande puede gestionar miles de llamadas al día, que deben ser atendidas en pocos segundos [20]. Uno de los mejores indicadores de calidad del servicio (*QoS*), directamente relacionado con la satisfacción del usuario, es la cantidad de llamadas no atendidas. La mayoría de centros —especialmente los centros de atención de llamadas de emergencia— incluyen un cierto grado de sobredimensionamiento que permite atender picos de demanda puntuales. Debido a los costes de personal, este sobredimensionamiento ha de adecuarse al nivel de servicio deseado. En caso contrario, un excesivo sobredimensionamiento restaría viabilidad económica al sistema.

La calidad del servicio, por tanto, se contrapone a la obtención de beneficios, siendo necesario realizar una planificación de recursos más adecuada que permita alcanzar un mejor compromiso entre costes y servicio. A pesar de que el método clásico empleado para simular el comportamiento dinámico de los centros de atención de llamadas es la teoría de colas [19], en este artículo veremos que las técnicas de predicción de series temporales pueden aplicarse de forma satisfactoria en este contexto [11]. Como veremos en los resultados experimentales, los sistemas difusos pueden aplicarse con éxito en este entorno, permitiendo que el experto pueda adoptar decisiones de gestión con mayor información. La interpretabilidad del sistema lógico difuso aporta, además, un valor añadido al sistema generado.

El resto del artículo se estructura de la siguiente manera. La Sección 2 presenta las características de un centro de atención de llamadas de emergencia. La Sección 3 introduce los conceptos de serie temporal, selección de variables y Sub-

tractive Clustering. Las secciones 4 y 5 presentan la experimentación realizada y sus resultados. Finalmente, se presentan las conclusiones y futuras líneas de investigación.

2 CARACTERÍSTICAS DE UN CENTRO DE ATENCIÓN DE LLAMADAS DE EMERGENCIA

En los centros de atención de llamadas de emergencia se debe prestar especial atención a maximizar la disponibilidad del servicio, evitando pérdida de llamadas y esperas. Aunque siga siendo un aspecto importante, la rentabilidad económica queda relegada a un segundo plano.

Un centro de atención de llamadas de emergencia está diseñado para responder eficientemente a llamadas indviduales, asistiendo a las personas que realizan las llamadas y redirigiéndolas, si es necesario, al proveedor de servicios de emergencia adecuado [18] (e.g. sanidad, policia, bomberos, etc.), tal y como ilustra la Figura 1.

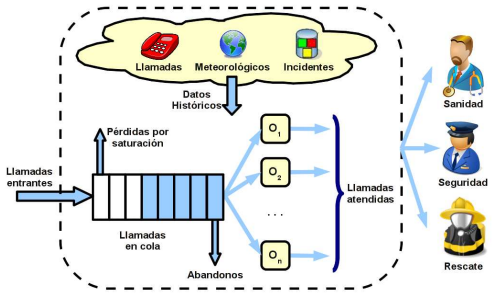


Figura 1: Sistema de atención de llamadas de emergencia.

Para mejorar la gestión de centros de este tipo y en especial la asignación de recursos, en otros trabajos se planteó la necesidad de extraer conocimiento útil a partir del registro histórico de llamadas recibidas [17]. En este trabajo se desea obtener una predicción de las llamadas entrantes del día siguiente a partir del histórico de llamadas recibidas, que será modelado como una serie temporal. Disponer de una predicción precisa es esencial para planificar adecuadamente los recursos humanos que van a ser necesarios.

Como se demostrará en la experimentación, la identificación de los retardos más significativos permite construir mejores modelos. Para identificar esos retardos se emplearán técnicas de selección de variables, concrétamente, métodos de ranking. Una vez identificados esos retardos, se ajustarán regresores basados en técnicas de inteligencia computacional: Sistema Lógico difuso a partir de Clustering y MLP. Por último, se modelará la serie siguiendo un enfoque diferente que no emplea selección de variables: los métodos estadísticos de Box-Jenkins, AR y ARMA.

3 INTELIGENCIA COMPUTACIONAL PARA PREDICCIÓN DE SERIES TEMPORALES

3.1 SERIES TEMPORALES

Una serie temporal (1) se define como una señal que evoluciona a través del tiempo o como un conjunto de valores medidos en intervalos de tiempo sucesivos [10, 12]. El concepto clave de una serie temporal es su evolución temporal.

$$X = \{x_t : t \in T\} \tag{1}$$

donde X es la serie temporal, x_t representa el valor en el instante t y T representa la evolución en el Tiempo.

Según (1), las llamadas diarias recibidas en el centro de atención de llamadas se pueden modelar como serie temporal. Su representación gráfica se observa en la Figura 2.

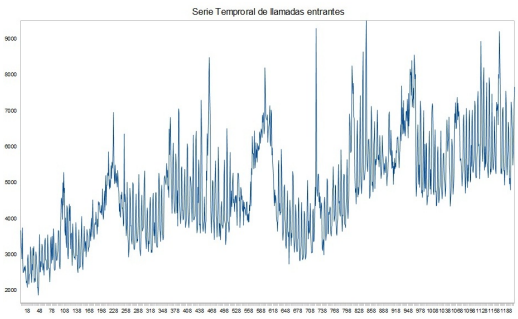


Figura 2: Serie Temporal de llamadas diarias recibidas.

El objetivo del modelado o predicción de series temporales es obtener el siguiente valor de la serie $\hat{x}_{t+1}$ a partir de k valores anteriores x_{t-i}, donde k es desconocido, ver (2).

$$\hat{x}_{t+1} = f(x_t, x_{t-1}, ..., x_{t-k}) \tag{2}$$

La idea subyacente en este proceso implica descubrir el patrón de generación de los datos y extrapolarlo al futuro.

Existen multitud de métodos y publicaciones científicas [13, 6, 16, 7, 11] en las que se construyen modelos para predecir series temporales, aplicadas a diversos dominios como medicina, meteorología, series econométricas, carga energética y automatización, entre otras.

Destacan las aproximaciones estadísticas, como los modelos $AR(p)$ y $ARMA(p,q)$ de Box-Jenkins [1] y las técnicas de inteligencia computacional, como las Redes Neuronales, Máquinas de Vector Soporte, Sistemas Difusos y modelos híbridos.

3.2 Selección de Variables

Los métodos de selección de características o variables, especialmente interesantes en reconocimiento de patrones, buscan obtener un conjunto de variables relevantes para el problema a partir de las disponibles.

En la experimentación se demostrará que la inclusión de una etapa previa de selección de variables que identifique los mejores retardos permite obtener mejores resultados al modelar la serie de llamadas con las técnicas de inteligencia computacional.

Los objetivos generales de los métodos de selección de variables son:

- Mejorar el proceso y el modelo de un problema.
- Mejorar las capacidades predictivas del predictor (respecto a efectividad y rapidez de ejecución).
- Mejor comprensión del proceso que genera los datos.
- Reducir la probabilidad de aparición de la maldición de la dimensionalidad —curse of dimensionality—.

Un enfoque diferente, denominado construcción o extracción de características, obtiene nuevos subconjuntos a partir de combinaciones de las variables originales, perdiendo su significado original y por tanto, su interpretabilidad.

Los algoritmos de selección de características fueron clasificados por Kohavi [2] en métodos de filtro, envoltorio e integrados. Es la clasificación más extendida y utilizada:

Filtro (Filter): Seleccionan las características en un paso aislado del preprocesamiento, sin tener en cuenta los efectos que el subconjunto de variables relevantes ejercen en el rendimiento del algoritmo utilizado para el ajuste. El proceso es independiente de la técnica de ajuste utilizada.

Envoltorio (Wrappers): Emplean un algoritmo de ajuste como caja negra, tratando de asegurar la capacidad predictiva del subconjunto evaluado. Es necesario definir el punto de inicio de la búsqueda, el criterio de evaluación de subconjuntos y la estrategia de generación de nuevos subconjuntos candidatos. La evaluación se realiza a partir de los resultados del ajuste, lo que aporta la ventaja de prueba de los subconjuntos sobre el entorno real. Sin embargo, los wrappers son criticados por su similitud con métodos de fuerza bruta por utilizar métodos de aprendizaje automático para realizar su función.

Integrados (Embedded): Son métodos de aprendizaje automático que realizan una selección de variables durante su ajuste, como la generación de árboles de decisión [4].

Por su parte, Guyon [3] añadió el siguiente grupo:

Valoración de variables (Feature ranking): Son un caso particular de los métodos de filtro que realizan una ordenación de las características empleando una función de relevancia. Su ventaja es la eficiencia tanto computacional —evaluar n características y ordenar sus resultados— como estadística —robustos frente al sobreajuste—. Sin embargo, tienden a obtener variables redundantes y no son capaces de evaluar la capacidad de predicción de subconjuntos de variables. A este tipo de métodos pertenece Relief-F.

3.2.1 Relief-F

Relief-F [9] es un método de ranking basado en la habilidad de una variable para distinguir ejemplos similares. Para ello analiza los vecinos más cercanos de cada ejemplo, considerando acierto o hit si ambos pertenecen a misma clase y fallo o miss si pertenecen a distintas clases. Relief-F valora cada atributo, iterativamente para cada ejemplo, según (3):

$$\begin{aligned} W[f] &= W[f] - diff(f,E_1,H) \\ &+ \sum_{C \neq clase(E_1)} p(c) \cdot diff(f,E_1,M(C)) \end{aligned} \tag{3}$$

donde $W[f]$ representa la valoración del atributo f. H y $M(C)$ son, respectivamente, el Hit o acierto y el Miss o fallo más cercanos de la clase C. E_i representa el ejemplo i-esimo y $diff(f,E_1,E_2)$ valora el grado en el que f es diferente en E_1 y E_2.

Para manejar atributos continuos, Kononenko [8] introdujo la siguiente implementación de $diff$:

$$diff(f,E_1,E_2) = \frac{|valor(f,E_1) - valor(f,E_2)|}{max(f) - min(f)} \tag{4}$$

Como resultado, tras la aplicación de Relief-F se dispondrá de un ranking con la valoración de la capacidad de cada atributo para distinguir las diferentes clases de patrones.

3.3 Predicción de Series Temporales

Para modelar las llamadas recibidas en el centro como una serie temporal en este trabajo se van a aplicar diferentes aproximaciones. Además de utilizar métodos estadísticos más clásicos como los modelos AR y ARMA de Box Jenkins, empleamos una red neuronal MLP y una técnica de generación de un sistema lógico difuso a partir del agrupamiento realizado por Subtractive Clustering.

3.3.1 Subtractive Clustering (SC)

SC [5] es una técnica de agrupamiento derivada de Mountain Clustering [15] que extrae centroides de los propios patrones candidatos con un orden computacional lineal respecto al tamaño del problema —$O(n)$—.

Para identificar los centroides (5), se define una medida de Densidad D_i. El primer centroide x_{C_1} será el patrón con mayor densidad, D_{C_1}. A continuación, se actualiza la medida de densidad para el resto de patrones según (6):

$$D_i = \sum_{j=1}^{n} exp\left(-\frac{\|x_i - x_j\|^2}{\left(\frac{r_a}{2}\right)^2}\right) \quad (5)$$

$$D_i = D_i - D_{C_i} \cdot exp\left(-\frac{\|x_i - x_{C_i}\|^2}{\left(\frac{r_b}{2}\right)^2}\right) \quad (6)$$

Iterativamente, el resto de centroides son identificados hasta alcanzar el número de clusters deseado. SC necesita definir dos parámetros r_a y r_b, que indican, respectivamente, los radios de vecindad y de reducción de la densidad.

Posteriormente, una vez identificados los clusters se crea un Sistema de Inferencia Difuso de tipo Takagi-Sugeno, representando cada centroide x_{C_i} como una regla difusa. Cada centroide es dividido en dos componentes: y_{C_i} —entradas— y z_{C_i} —salidas—. Las funciones de pertenencia de los antecedentes de las reglas se definen en (7):

$$\mu_i = exp\left(-\frac{\|y_i - y_{C_i}\|^2}{\left(\frac{r_a}{2}\right)^2}\right) \quad (7)$$

Por su parte, los consecuentes de las reglas del sistema lógio difuso son funciones lineales de las entradas, obtenidas tras realizar un ajuste por mínimos cuadrados.

4 EXPERIMENTACIÓN

La experimentación realizada en este trabajo, resumida en la Figura 3, consiste en evaluar la capacidad de un sistema lógico difuso para modelar, con una precisión adecuada, la serie de llamadas recibidas. El modelado se realizará empleando subconjuntos de retardos de diferente tamaño, tomando como base el ranking obtenido por Relief-F. La capacidad predictiva del sistema difuso se contrastará con otras técnicas utilizadas para modelar series temporales: MLP, AR y ARMA.

Los subconjuntos de retardos a evaluar se obtendrán a partir del ranking de relevancia obtenido por Relief-F, aplicando diferentes ratios de reducción de la dimensionalidad —PRD (8)—. Esta experimentación demostrará que los regresores se ajustan mejor cuando se emplea un número adecuado de los mejores retardos.

$$PRD = 100 \cdot \left(1 - \frac{p}{q}\right) \quad (8)$$

siendo p y q las dimensiones seleccionada y original.

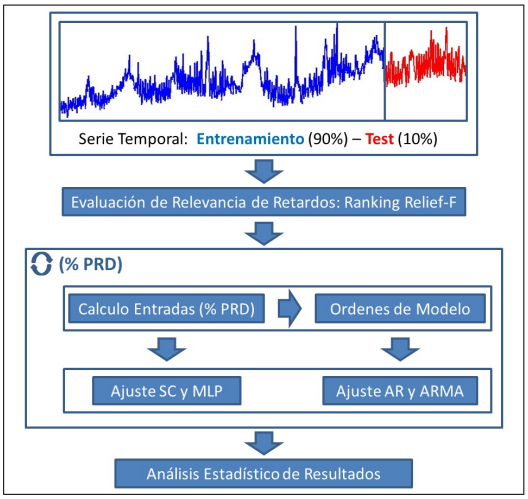


Figura 3: Metodología Experimental

El retardo máximo, de acuerdo a la dirección del centro, se establece en 20 días. A partir de ese retardo, se considera que no hay influencia en las futuras llamadas entrantes.

La Tabla 1 resume la configuración de la experimentación realizada con los diferentes regresores. Subtractive Clustering y MLP trabajarán con las primeras variables identificadas por Relief-F. El número de variables a seleccionar queda determinado por el PDR objetivo.

PRD	Subtractive Clustering Núm. Entradas	MLP Arquitectura Red	$AR(p)$ Orden modelo	$ARMA(p,q)$ Orden modelo
75%	5 (top-5 Relief-F)	5(top-5 Relief-F)-20-1	(5)	(5,5)
70%	6 (top-6 Relief-F)	6(top-6 Relief-F)-20-1	(6)	(6,6)
65%	7 (top-7 Relief-F)	7(top-7 Relief-F)-20-1	(7)	(7,7)
60%	8 (top-8 Relief-F)	8(top-8 Relief-F)-20-1	(8)	(8,8)
55%	9 (top-9 Relief-F)	9(top-9 Relief-F)-20-1	(9)	(9,9)
50%	10 (top-10 Relief-F)	10(top-10 Relief-F)-20-1	(10)	(10,10)
0%	20 (Todas)	20(Todas)-20-1	(20)	(20,20)
Implem.	$genfis2$[1], $substclust$[2] $r_a = 0.06..3,00\,(+0,01)$	$newff$,$train$[3] $epochs = 1000, lr = 0,01, \mu = 0,9$	ar[4]	$armax$[5]

Tabla 1: Configuración de la Experimentación

Además de comparar todos estos resultados entre sí, para completar la experimentación las mejores configuraciones de SC y MLP se compararán frente al ajuste realizado con técnicas estadísticas clásicas empleadas para modelar series temporales: los modelos AR y ARMA de Box Jenkins. El orden de los modelos será el equivalente al número de

[1] subclust + least squares available at: `http://www.mathworks.es/help/toolbox/fuzzy/genfis2.html`

[2] subclust available at: `http://www.mathworks.es/help/toolbox/fuzzy/subclust.html`

[3] mlp train available at: `http://www.mathworks.es/help/toolbox/nnet/ref/train.html`

[4] ar available at: `http://www.mathworks.es/help/toolbox/ident/ref/ar.html`

[5] arma available at: `http://www.mathworks.es/help/toolbox/ident/ref/armax.html`

variables utilizado en los modelos de SC y MLP. De esta manera, además de demostrar que la identificación de los retardos más significativos permiten obtener mejores modelos de predicción, se justificará que la lógica difusa se puede aplicar de forma satisfactoria para modelar series temporales.

5 RESULTADOS

Una vez presentado el marco experimental, en esta sección se muestran los resultados obtenidos en el modelado de la serie de llamadas con todos los regresores: SC, MLP, ARMA y AR.

Configuración		Subtractive Clustering (SC)				
PDR	Vars	r_a	ERP	% RED	ECM	% RED
75%	5/20	1,04	7,94	-6,49	395707,28	-12,51
70%	6/20	1,12	7,47	-0,18	339372,29	3,51
65%	7/20	1,15	7,36	1,37	331235,67	5,82
60%	**8/20**	**1,17**	**7,29**	**2,28**	**329785,08**	**6,23**
55%	9/20	1,35	7,41	0,61	344535,59	2,04
50%	10/20	1,44	7,43	0,46	344731,36	1,98
0%	20/20	1,9	7,46	0	351711,43	0

Tabla 2: Comparativa de Resultados SC

En la Tabla 2 se observan los resultados de la experimentación realizada para los sistemas difusos ajustados a partir del clustering realizado con SC. La tabla muestra cómo la selección de variables ayuda a identificar mejores modelos de la serie. La primera columna indica el PDR planteado, seguido del número de variables identificado por Relief-F para esa reducción y del radio de vecindad r_a que mejores resultados aporta. Finalmente, se muestran dos medidas de error calculadas: el Error relativo porcentual (ERP) y el Error cuadrático medio (ECM). En primer lugar se muestra su valor concreto, seguido del ratio de reducción de ese error (% RED) respecto al ajuste realizado incorporando el máximo de retardos. Se observa que salvo en casos puntuales en los que el sistema lógico difuso se ajusta con muy pocos retardos, el error de predicción es menor al emplear un número razonable de los mejores retardos.

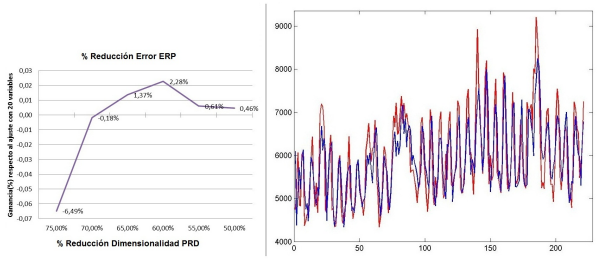


Figura 4: Resultados de la Experimentación con SC.

La parte izquierda de la Figura 4 representa un estudio realizado para los mejores resultados obtenidos con los sistemas lógicos difusos generados con SC. En la gráfica se representa la mejora obtenida con los modelos ajustados con los mejores retardos frente al ajuste realizado con el máximo de retardos. La mejor configuración para predecir la serie con SC se obtiene, por tanto, empleando las mejores 8 variables, es decir, únicamente el 40% de las variables iniciales.

Esta configuración, además de ser más manejable e interpretable, aporta una mejora del 2.28% en el error ERP y del 6.23% en el error ECM frente al ajuste realizado incluyendo el máximo de retardos. La alta precisión de este modelo se puede observar en la parte derecha de la Figura 4, que muestra el comportamiento del modelo frente a nuevos datos independientes no utilizados en el ajuste.

Para completar el estudio, la Tabla 3 muestra una comparativa de los resultados obtenidos en la etapa de test con todos los regresores. Para cada regresor se indica su configuración y los errores ERP y ECM. En la última línea se indica el comportamiento medio de cada regresor.

SC			MLP$(x-20-1)$			$AR(p)$			$ARMA(p,q)$		
Vars.	*ERP*	*ECM*	*ARQ*	*ERP*	*ECM*	p	*ERP*	*ECM*	(p,q)	*ERP*	*ECM*
5	7,94	395707,28	5 – 20 – 1	9,03	498394,76	(5)	9,06	522081,82	(5,5)	8,01	397120,18
6	7,47	339372,29	6 – 20 – 1	8,13	404426,22	(6)	8,46	432165,34	(6,6)	7,92	385867,54
7	7,36	331235,67	**7 – 20 – 1**	**7,45**	**349996,67**	(7)	8,40	423907,01	(7,7)	7,96	385200,43
8	**7,29**	**329785,08**	8 – 20 – 1	7,65	370468,93	(8)	8,12	394123,13	(8,8)	8,11	392400,62
9	7,41	344535,59	9 – 20 – 1	7,67	361296,84	(9)	8,00	391361,27	**(9,9)**	**7,91**	**381335,14**
10	7,43	344731,36	10 – 20 – 1	8,29	424097,26	**(10)**	**7,94**	**385586,24**	(10,10)	7,93	381697,08
20	7,46	351711,43	20 – 20 – 1	7,89	386743,02	(20)	8,04	388742,41	(20,20)	8,16	400912,07
	7,48	348154,10		8,01	399346,24		8,29	389219,01		8,00	419709,60

Tabla 3: Resultados de Experimentación

Como era de esperar, los modelos no lineales superan en rendimiento al modelo lineal AR. Finalmente, SC es la técnica que mejor se comporta, superando en rendimiento a ARMA y MLP que aportan resultados muy similares —a pesar de que ARMA tiene mayor grado de libertad— seguido finalmente por los modelos lineales AR.

Para concluir con la experimentación realizada en este trabajo, se aplican tests estadísticos no paramétricos de Wilcoxon [14] para contrastar que la mejora obtenida de SC respecto al resto de modelos, bajo ambas medidas de error, es estadísticamente relevante. La Tabla 4 muestra los p-valores obtenidos tras comparar ambos tipos de error para cada regresor. A partir de estos p-valores, podemos concluir, para un nivel de confianza del 99%, que la mejora de SC frente al resto de métodos es estadísticamente relevante.

Test de Wilcoxon				H_0: Igualdad de Medias de Error H_1: Error Medio (REG[a]) > Error Medio (SC)			
Comparativas para Error Relativo Porcentual (*ERP*)				Comparativas para Error Cuadrático Medio (*ECM*)			
SC	MLP	$ARMA(p,q)$	$AR(p)$	SC	MLP	$ARMA(p,q)$	$AR(p)$
7,94	9,03	8,01	9,06	395707,28	498394,76	397120,18	522081,82
7,47	8,13	7,92	8,46	339372,29	404426,22	385867,54	432165,34
7,36	**7,45**	7,96	8,40	331235,67	**349996,67**	385200,43	423907,01
7,29	7,65	8,11	8,12	**329785,08**	370468,93	392400,62	394123,13
7,41	7,67	**7,91**	8,00	344535,59	361296,84	**381335,14**	391361,27
7,43	8,29	7,93	**7,94**	344731,36	424097,26	381697,08	**385586,24**
7,46	7,89	8,16	8,04	351711,43	386743,02	400912,07	388742,41
p-valor	0,0078	0,0078	0,0078	p-valor	0,0078	0,0078	0,0078
¿Rechazar H_0?	Si	Si	Si	¿Rechazar H_0?	Si	Si	Si

[a] donde REG será MLP, ARMA o AR, según corresponda

Tabla 4: Resultados Tests Estadísticos

6 CONCLUSIONES

En este trabajo hemos podido modelar, con un nivel de precisión más que adecuado, la serie de llamadas recibidas en un centro de atención de llamadas de emergencia.

A partir de los resultados obtenidos, se ha demostrado que la técnica de generación de reglas difusas a partir del agrupamiento realizado por Subtractive Clustering es adecuada para realizar la predicción de series temporales.

El estudio empírico realizado, certifica la conveniencia de la utilización de una etapa de selección de variables previa al ajuste de modelos para predicción de series de tiempo.

Al modelar la serie con subconjuntos de variables más reducidos, se consigue reducir la complejidad del sistema lógico difuso obtenido, y por lo tanto, mejorar su interpretabilidad.

Por último, tal como era esperado, se ha demostrado que la mejora obtenida en los ajustes realizados por la combinación de selección de variables y SC es estadísticamente relevante frente a la obtenida por modelos equivalentes con el resto de regresores: selección de variables con MLP y métodos estadísticos ARMA y AR.

Referencias

[1] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated, 1990.

[2] R. Kohavi and H. G. John. Wrappers for feature subset selection. *Artificial Intelligence AIJ*, 97:273–324, 1997.

[3] An Introduction to Variable and Feature Selection *Journal of Machine Learning Research*, 3, 1157-1182, 2003.

[4] J. Ross Quinlan. Induction of decision trees. *Mach. Learn.*, 1(1):81–106, 1986.

[5] S.L. Chiu. Fuzzy model identification based on clustering estimation. *Intelligent and Fuzzy Systems*, 2:267–278, 1994.

[6] Hao J. Reyhani N. Ji, Y. and A. Lendasse. Direct and recursive prediction of time series using mutual information. *Lecture Notes in Computer Science*, 3512:1010–1017, 2005.

[7] Mozayani N. Khazaee, P.R. and M.R.J. Motlagh. A genetic-based input variable selection algorithm using mutual information and wavelet network for time series prediction. In *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, number 4811607, pages 2133–2137, 2008.

[8] I. Kononenko. Estimating attributes: Analysis and extensions of relief. *Proc. of the European conf. on Machine Learning*, 1994.

[9] K.; Kira and L.A. Rendell. The feature selection problem: traditional methods and new algorithm. *Proceedings of AIII*, 1992.

[10] O. Nelles. *Nonlinear System Identification*. Springer, 2001.

[11] R. G. Pajares, J. M. Benitez, and G. S. Palmero. Feature selection for time series forecasting: A case study. *Hybrid Intelligent Systems, Eighth International Conference on*, pages 555–560, September 2008.

[12] D. Peña Sánchez de Rivera. *Estadística Modelos y métodos, 1. Fundamentos*. Alianza Editorial, Madrid., 1986.

[13] Hao J. Reyhani N. Ji Y. Sorjamaa, A. and A. Lendasse. Methodology for long-term prediction of time series. *Neurocomputing*, 70:2861–2869, 2007.

[14] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics*, 1:80–83, 1945.

[15] R. Yager and D.P. Filev. Approximate clustering via the mountain method. *IEEE Transactions on Systems Man and Cybern*, 24:1279–84, 1994.

[16] Qiu H. Yan, W. and Y. Xue. Gaussian process for long-term time-series forecasting. In *Proceedings of the International Joint Conference on Neural Networks*, number 5178729, pages 3420–3427, 2009.

[17] F. Barrientos, G. Sáinz: Interpretable knowledge extraction from emergency call data based on fuzzy unsupervised decision tree. *Knowledge-Based Systems* 25, pp. 77–87, 2012.

[18] H. Jasso, W. Hodgkiss, C. Baru, T. Fountain, D. Reich, K. Warner: Using 9-1-1 call data and the space-time permutation scan statistic for emergency event detection. *Government Information Quarterly* 26, pp. 265–274, 2009.

[19] G. Koole, A. Mandelbaum: Queueing models of call centers. An introduction. *Annals of Operations Research* 113, pp. 41–59, 2003.

[20] A. Mandelbaum, O. Garnett, M. Reiman: Designing a call center with impatient customers. *Manufacturing and Service Operations Management* 4, pp. 208–227, 2002.

NEUROFUZZY MODEL OF AN INDUSTRIAL PROCESS, REDUCING COMPLEXITY BY USING PRINCIPAL COMPONENT ANALYSIS

Juan Manuel Escaño, **Carlos Bordons**

Universidad de Sevilla. Escuela Superior de Ingeniería. Avda. de los Descubrimientos s/n. 41092 Sevilla,
{jmesca,bordons}@cartuja.us.es

Abstract

A Neurofuzzy model of a mixing chamber pressure has been proposed. The process is a part of a copper smelter plant. The principal component analysis (PCA) method has been used to reduce the inputs space for a recurrent fuzzy model. The coupling among variables and their mutual influence between themselves, are taken into account by the projection into the PCA axis. The model have been validated with real data from the factory. The validation result shows that the model is suitable for simulation.

Keywords: Neurofuzzy modeling, Industrial process, Principal Component Analysis.

Figure 1: General view of the copper smelter [1]

1 Introduction

The process studied in this work is a part of Atlantic Copper Smelter facilities in Huelva (Spain), whose anual production is around three hundred thousand tons of copper [1]. This plant includes a Flash Furnace and four Pierce-Smith converters, two of them blowing simultaneously. The three currents of gases generated in these processes are mixed in the mixing chamber and sent to three acid plants operating in parallel (see figure 2). It is very important to maintain the gas pressure in the mixing chamber at a desired value, always bellow ambient pressure in order to avoid gas losses to the atmosphere. That pressure depends on other variables of the production line and it is very difficult to get an accurate prediction of it. On the one hand, the causes of the pressure oscillations are hard to detect. Moreover, due that there are different control system in the copper smelter and the acid plant, no clocks synchronization is possible, so there are considerable uncertainties when trying to measure cause-effect delays. A suitable model for prediction one step ahead, has been developed in [1]. Other models have been made in [5] and [6]. All of them are prediction models, because the actual pressure value at time k $P_{MC}(k)$ is used to predict $P_{MC}(k+1)$. Taking into account that the converters operate on a batch mode, while those of the flash furnace and acid plants are continuous, extremely high disturbances both in flow and SO_2 concentration occur at the acid plants inlet due to the converters' operating schedule. The existing control strategy, based on independent single loop PID controllers, is not able to cope with those disturbances [1]. Advanced control schemes should be applied, but it would be interesting to have a model suitable for simulation. In this case, it is difficult to derive a precise mathematical model, based on first principles. Besides, the computation of the solution of models obtained through this methodology may require a large computational effort making them useless for real time tasks like control or optimization. Neurofuzzy modeling, which permits an easy way to derive successful models, is a good alternative which can be employed to overcome such limitations [2],[3],[7] and [9]. Fuzzy Neural Networks (FNN) combine the capability of uncertainty handling in information with learning skill. Recurrent Fuzzy Neural Network (RFNN) have proven to be an excellent choice in order to

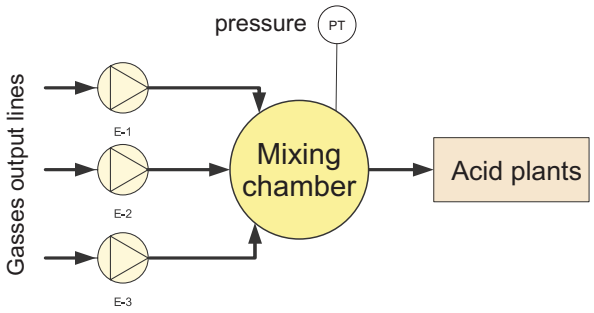


Figure 2: gas mixing. Acid plants

get the dynamics of nonlinear and complex systems. They are systems which have the same advantages than recurrent neural networks [10]. RFNN are also known as *Fuzzy Dynamical Systems* (see figure 3) and extend the application domain of FNN to temporal problems. Feedback allows to capture dynamics and change. Principal Component Anal-

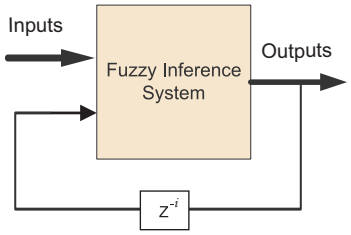


Figure 3: Dynamical Neurofuzzy System

ysis is a well know technique in the field of multivariate methods [8]. The use of PCA reduces the space of input variables and obtain new uncorrelated variables in order to simplify the Fuzzy system. This paper is organized as follows. In section 2, a formulation for dynamic neurofuzzy model is shown. An introduction to Principal Components Analysis is given in section 3. In section 4, the neurofuzzy model for the mixing chamber pressure is presented, giving the validation results. Conclusions are given in section 5.

2 Dynamic Neurofuzzy Model formulation

In the neurofuzzy model proposed by Takagi-Sugeno (TS)[12], the structure of antecedent describes fuzzy regions in the inputs space, and the one of consequent presents non-fuzzy functions of the model inputs. We can use recurrent functions with NARMAX structure (*Nonlinear Auto Regressive Moving Average with eXogenous input*), of the kind:

$$\hat{y}(k+1) = f(y(k),...,y(k-m),u(k),...,u(k-n)) \quad (1)$$

Where u, y are respectively the inputs and outputs of the system, the Neurofuzzy system may be described, for each rule, in the following way:

R_j :
IF x_1 is F_{1j},..., and x_n is F_{nj},
THEN:

$$y_j = g_{j1}x_1 + g_{j2}x_2 + ... + g_{jn}x_n + c_j \quad (2)$$

Where g_{ji} and c_j are constant terms, $X = \begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix}^T$ is the input vector of the neurofuzzy system, F_{ij} is the fuzzy set respective to x_i on the rule j, y_j is the output of the model respective to the operating region associated to the rule. x_i can represent a real input to the system or any other variable, for instance the previous values of inputs or outputs. Thus, we could formulate the consequent (2) like

$$y_j(k) = a_j(z^{-1})y(k-1) + b_j(z^{-1})u(k-d) + c_j \quad (3)$$

Where $a_j(z^{-1}) = a_{1j} + a_{2j}z^{-1} + ... + a_{n_y j}z^{-(n_y-1)}$ and $b_j(z^{-1}) = b_{0j} + b_{1j}z^{-1} + b_{2j}z^{-2} + ... + b_{n_u j}z^{-n_u}$

d is the dead time. If $\mu_{ij}(k)$ is the membership degree of $x_j(k)$ in the fuzzy set F_{ij} and the number of implications or rules is L, the RFNN complete model is described by

$$y(k) = \sum_{j=1}^{L} w_j(k)\left[a_j(z^{-1})y(k-1) + b_j(z^{-1})u(k-d)\right] + \xi(k) \quad (4)$$

Where

$$w_j(k) = \frac{\bar{\mu}_j(k)}{\sum_{j=1}^{L}\bar{\mu}_j(k)}, \qquad \bar{\mu}_j(k) = \prod_{i=1}^{n}\mu_{ij}(k)$$

and

$$\xi(k) = \sum_{j=1}^{L} w_j(k)c_j$$

Rewriting equation (4) as

$$\bar{a}(z^{-1})y(k) = \bar{b}(z^{-1})u(k-d) + \xi(k) \quad (5)$$

Where

$$\bar{a}(z^{-1}) = 1 - \bar{a}_1 z^{-1} - \bar{a}_2 z^{-2} - ... - \bar{a}_{n_y} z^{-n_y} \quad (6)$$

$$\bar{b}(z^{-1}) = 1 - \bar{b}_1 z^{-1} - \bar{b}_2 z^{-2} - ... - \bar{b}_{n_u} z^{-n_u} \quad (7)$$

$$\bar{a}_i = \sum_{j=1}^{L} w_j(k)a_{ij}z^{-i} \quad (8)$$

$$\bar{b}_i = \sum_{j=1}^{L} w_j(k)b_{ij}z^{-i} \quad (9)$$

In [4] a dynamic Neurofuzzy Model is used for simulation. We propose an improvement of that model, using a major number of inputs, including squares of variables, to provide also a non linear dependence for each rule. A PCA has been used both in a model used in [4] as the one proposed here. In the first, the analysis is directed just to get uncorrelated variables, in the second also much more simplification is achieved in the FIS.

3 Principal Components Analysis

Due to coupling between variables and the difficulty to make experimental tests, the use of PCA provides two characteristics: to achieve new uncorrelated variables and condensation of all the information in a smaller space, providing a simpler *Adaptive Neural Fuzzy Model* (ANFIS). It will permit the addition of others input variables, without complexity increasing. The aim of PCA is to reduce system dimension, minimizing missing information. The idea behind PCA is to form a minimum number of new variables to describe the variation of the original data by using linear combinations of the original variables.

Let α be a vector of p variables, $\alpha = \begin{bmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_p\end{bmatrix}$

And $A \in \mathbb{R}^{n\times p}$ a data matrix generated by the p variables, where n is the number of data. The aim of PCA is to find a base $\delta = \begin{bmatrix}\delta_1 & \delta_2 & \dots & \delta_k\end{bmatrix}$ where $k < p$, which defines a new subspace retaining the maximum information of the original data.

$$\begin{aligned}\delta_1 &= w_{11}\alpha_1 + w_{12}\alpha_2 \dots + w_{1p}\alpha_p\\ \delta_2 &= w_{21}\alpha_1 + w_{22}\alpha_2 \dots + w_{2p}\alpha_p\\ &\dots\dots\dots\dots\dots\\ \delta_k &= w_{k1}\alpha_1 + w_{k2}\alpha_2 \dots + w_{kp}\alpha_p\end{aligned}$$

Then the data matrix will be $D \in \mathbb{R}^{n\times k}$.
It is demonstrated that

$$\mu_D = E(D) = E(W^T A) = W^T E(A) \tag{10}$$

And the covariance matrix of D is equal to

$$\Sigma_D = E\{(D-\mu_D)(D-\mu_D)^T\} = W^T \Sigma_A W \tag{11}$$

The goal is to get the maximum data variance in the new axes, that is maximum Σ_D, imposing the orthonormality constraint on it:

$$W^T W = I \tag{12}$$

We have to maximize:

$$W^T \Sigma_A W - \lambda(W^T W - I) \tag{13}$$

deriving and making it equal to zero, we have

$$(\Sigma_A - \lambda I)W = 0 \tag{14}$$

The problem is just one of calculation of eigenvectors of Σ_A. The associated components to greater eigenvalues of Σ_A are the most meaningful to build the data space. They are named *Principal Components* of the system. In order to choose how many principal components to use, a criterium based on the weight of each eigenvalue with respect to the other, can be used:

$$\frac{\sum_{i=1}^{l}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \geq n \tag{15}$$

being n a measurement of desired information.

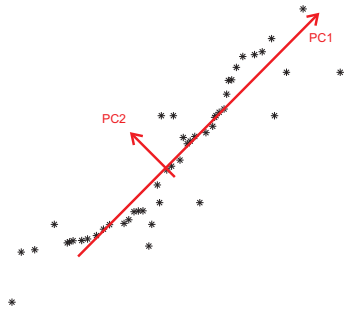


Figure 4: Principal Components

Table 1: List of variables

Description	*Units*	*Type*
Pressure in mixing chamber	mbar	Output
Flow to plant 1	KNm^3/h	Manipulated
Flow to plant 2	KNm^3/h	Manipulated
Flow to plant 3	KNm^3/h	Manipulated
Dilution flow to plant 1	Nm^3/h	Manipulated
Dilution flow to plant 2	Nm^3/h	Manipulated
Dilution flow to plant 3	Nm^3/h	Manipulated
Reference for flash furnace feeding	Tons/h	Disturbance
Flow control valve for flash furnace	%	Disturbance
Fan speed in flash furnace	rmp	Disturbance
Reference for fan speed line 1	rpm	Disturbance
Reference for fan speed line 2	rpm	Disturbance

4 Neurofuzzy Model of Mixing Chamber

After a preliminary study based on some experiments with steps on the variables[1], the evolution of the pressure in the mixing chamber (P_{MC}), is influenced by others that are divided into two groups: control signals and disturbances. In table 1 a brief description of the considered variables is given, whereas in figure 5 a scheme with the situation of each of them is presented.

In [4], the inputs are the current samples of the variables of Table 1 and the previous samples of them, forming a

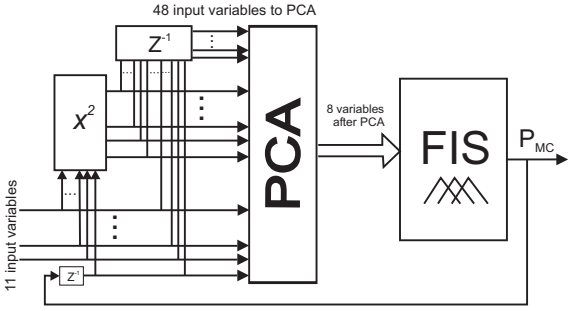


Figure 6: Neurofuzzy scheme proposed

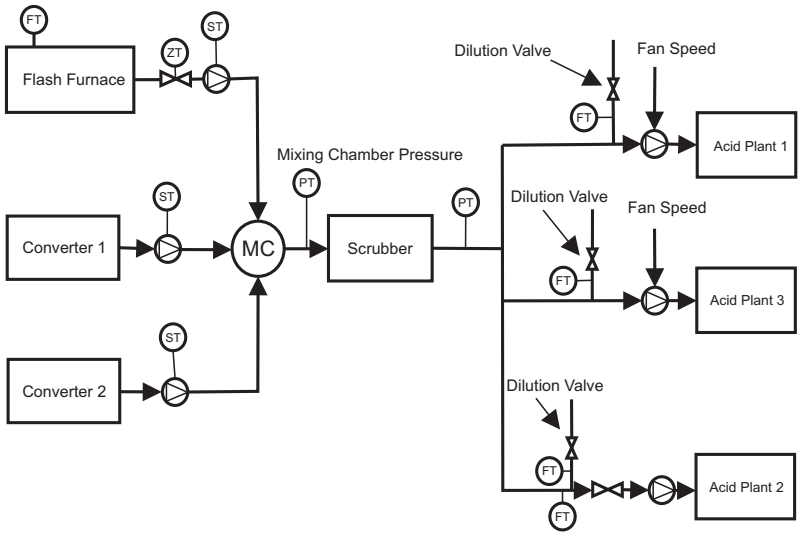


Figure 5: Process, manipulated variables and disturbances

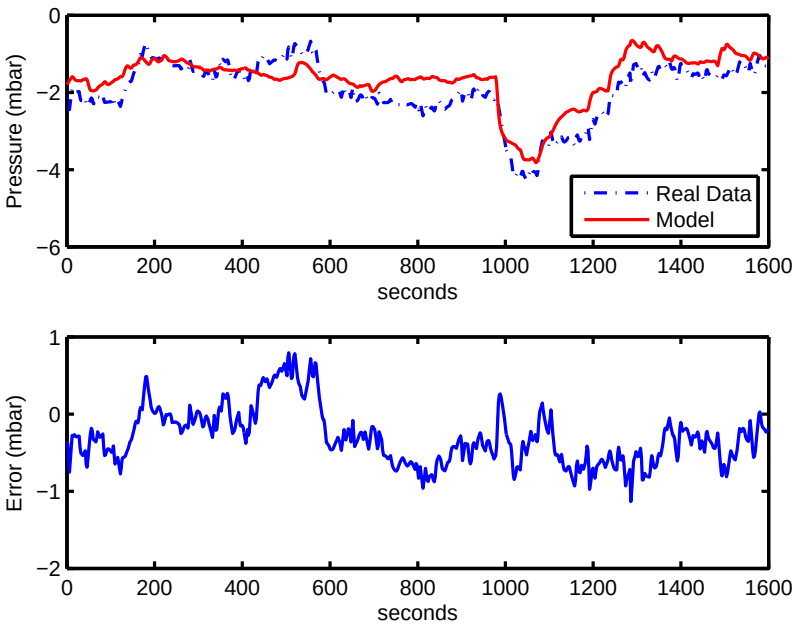


Figure 7: Validation of the model used in [4]

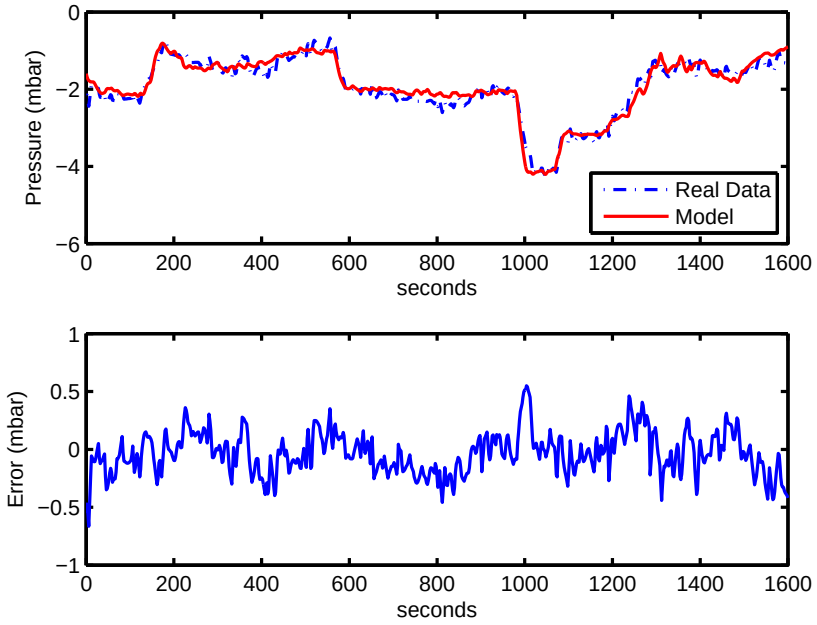


Figure 8: Validation of the proposed model

number of 24 inputs. After a PCA using all the nonzero eigenvalues, 21 components are obtained. As mentioned at the end of the introduction, the addition of entries does not complicate the model when using PCA. For this work, we have taken into account the input variables listed in Table 1 and their squares, to give to the model, quadratic elements of the input variables of the system. Figure 6 presents the scheme followed to obtain the neurofuzzy model. The inputs are the variables presented in table 1, including Pressure in mixing chamber, their squares and the previous samples of all of them. To carry out the PCA, data have been used for nearly three hours of operation, sampled every 2 seconds. With a loss of information from 1%, the new principal components are 7. Using these new uncorrelated variables as inputs to the fuzzy system, and the next sampling pressure as output, an ANFIS is designed, using Subtractive clustering technique [3]. The performance of the models can be seen in figures 7 and 8, where it is validated using a real data set from the process. It is important to note that $P_{MC}(t-1)$ is generated by the models output in the previous sampling. Looking at the figures we see that the error does not grow indefinitely, this fact makes the models, appropriate to simulate the process. In figure 7, the mean error is $0.4099°C$, while the new model proposed in this work, the mean error is $0.1469°C$, obtaining an improvement.

5 Conclusions

A Neurofuzzy model, suitable for simulation, of a mixing chamber pressure has been designed. The PCA method has been used to reduce the inputs space for a recurrent fuzzy model. Using this technique, a previous increment of input has been added, including quadratic elements and previous samplings. The coupling among variables and their mutual influence between themselves is taken into account by the projection into the PCA axis. The model have been validated with real data from the factory. The validation result shows that the model is suitable for simulation. In comparison with other simulation models of the same plant, an improvement has been got.

Acknowledgements

The authors acknowledge J.M. Tejera (Atlantic Copper), for the provision of real data to carry out the study. This work has been funded by the National Plan Project DPI2010-21589-C05-01 of the Spanish Ministry of Science and Innovation and FEDER funds.

References

[1] C. Bordons, M. R. Arahal, E. F. Camacho, J.M. Tejera. “ Energy Saving in a Copper Smelter by means of Model Predictive Control". Identification and Control, Chap. 3. pp. 63-65. Springer London 2007.

[2] Chi-Huang L., Ching-Chih T. “Generalized predictive control using recurrent fuzzy neural networks for industrial processes". Journal of Process Control. 2007

[3] Chiu, S., “Fuzzy Model Identification Based on Cluster Estimation," Journal of Intelligent & Fuzzy Systems, Vol. 2, No. 3, Spet. 1994

[4] J.M. Escaño, C. Bordons, F. Dorado. “Pca Based Pressure Control of a Gas Mixing Chamber". 2009 IEEE Conference on Emerging Technologies & Factory Automation (ETFA 2009), pp. 1-6. Palma de Mallorca. IEEE. 2009.

[5] J.M. Escaño, C. Bordons, A. Nuevo. “Dynamical & non-Dynamical neurofuzzy models of a mixing chamber pressure". Sice Annual Conference. Tokyo. Japón. 2008

[6] J.M. Escaño, C. Bordons. “Modelo neuroborroso de la presión de la cámara de mezcla en una fundición de cobre". XXVIII Jornadas Técnicas de Automática. CEA-IFAC. 2006

[7] Jhy-Shyng, Jang, R. “Neuro-Fuzzy Modeling and Control". Proceedings of the IEEE, Vol 83, No. 3, March 1995

[8] D. E. Johnson. “Applied multivariate methods for data analysis". Brooks Cole Publishing CO and ITP CO. 1998

[9] K. Kiriakidis, “Fuzzy Model-Based Control of Complex Plants". IEEE Transactions On Fuzzy Systems, vol. 6, NO. 4, Nov.1998

[10] Chao-Chee Ku, Kwang Y. Lee, “Diagonal recurrent neural networks for dynamic systems control", IEEE Trans. Neural Networks, vol.6, pp. 144-156, Jan. 1995.

[11] Ching-Hung Lee, Ching-Cheng Teng, “Identification and Control of Dynamic Systems Using Recurrent Fuzzy Neural Networks", IEEE Transactions On Fuzzy Systems, vol.8, no. 4, Aug. 2000

[12] Takagi and Sugeno, “Fuzzy Identification of Systems and Its Applications to Modelling and Control", IEEE Trans. Syst. Man. Cyber., Vol. 15, pp. 116-132, 1985

[13] A. Trabelsi, F. Lafont, M. Kamoun, G. Enea,“Identification of Nonlinear Multivariable Systems by Adaptative Fuzzy Takagi-Sugeno Model". International Journal of Computational Cognition (http://www.YangSky.com/yangijcc.htm). Volume 2. Number 3. Pages 137-153. 2004

MODELO DIFUSO DE PREFERENCIA-AVERSIÓN

Camilo Franco [1], Javier Montero [1], J. Tinguaro Rodríguez [1]

[1] *Facultad de Matemáticas, Universidad Complutense de Madrid, 28040, Madrid*,
francodelosrios@gmail.com, {monty, jtrodrig}@mat.ucm.es

Resumen

Este artículo presenta un modelo de preferencia-aversión (P-A), el cual permite representar el proceso mediante el que un individuo racional (bajo una racionalidad de pérdidas y ganancias) ordena sus alternativas, en base a la agregación independiente de sus percepciones tanto positivas como negativas. Como resultado, se obtiene la estructura P-A, representando los distintos estados epistémicos que bajo este marco pueden ser verificados según distintos grados de intensidad.

Palabras Clave: Estructura de preferencia-aversión, racionalidad de pérdidas y ganancias.

1 INTRODUCCIÓN

En este artículo se presenta el modelo de preferencia-aversión (P-A), examinando el proceso de decisión de un individuo con una racionalidad de pérdidas y ganancias. La intuición detrás de este tipo de racionalidad se sostiene en una primera teoría de decisión subjetiva [13] y en posteriores estudios acerca del comportamiento del individuo frente a situaciones de decisión bajo incertidumbre [8], [15].

Recientemente se ha utilizado el término *bipolar* [1], [7], para referirse a la identificación de dos polos opuestos, los cuales sirven como referencia para definir el significado de un concepto de interés. Estos polos pueden identificarse, por ejemplo, como aquello que es verdadero/falso, o para nuestro caso, como aquello que es ganancia/pérdida, entendiendo las ganancias y pérdidas de manera general, como aquello que se percibe benéfico o nocivo, a diferencia de una aproximación netamente monetaria (como la de [8]).

A continuación se presenta el modelo difuso de preferencia-aversión P-A, haciendo uso de técnicas de lógica difusa para la caracterización del predicado de preferencia [6], [14], [17], y se examina su tipo de bipolaridad de acuerdo con una tipología determinada [1] (siguiendo las pautas marcadas en [11]).

2 MODELO DIFUSO DE PREFERENCIA-AVERSIÓN

2.1. RELACIONES DE PREFERENCIA Y AVERSIÓN

El tipo de información que se obtiene de un conjunto de preferencias es de una gran riqueza semántica, siempre que se asume que pertenecen a un individuo racional o inteligente que pesa sus argumentos en busca de una buena decisión. Precisamente esta riqueza semántica es la que permite expresar las condiciones naturales de subjetividad e inexactitud del conocimiento humano (como se señala en [3]). Por lo tanto, la racionalidad del individuo estructura los datos o la información recibida en forma de preferencias, y esta estructura es isomorfa con las capacidades representativas de la racionalidad individual.

De este modo, la estructura de preferencias contiene una semántica propia, la cual corresponde con el tipo de información y la fuente de la misma, que para el caso es nada menos que la inteligencia humana. Si se acepta que un rasgo fundamental de dicha inteligencia es su capacidad de separar argumentos positivos y negativos sobre cierta opción, la estructura de preferencias debe representar dicha capacidad, guardando su condición isomorfa con la racionalidad humana. Este tipo de racionalidad es la que se denomina racionalidad de pérdidas y ganancias, siguiendo la intuición básica de la Teoría de Perspectivas Acumuladas (CPT, ver [8], [15]).

Esto es, como soporte de la decisión, el orden de preferencia se construye en la medida en que permite enfrentar la incertidumbre sobre las alternativas disponibles, valorando lo negativo en la medida en que es una dimensión de la información sobre la que se tiene un mayor grado de confianza (lo lesivo o nocivo es en principio más fácil de identificar), en oposición a como se valora lo positivo, dimensión ésta sobre la que se puede tener un menor

grado de seguridad (lo deseable es en principio más ambiguo) [13]. Esta diferenciación, junto con las posibilidades representativas de los distintos estados del conocimiento (examinados a partir de las distintas situaciones relacionales básicas de preferencia) y la incertidumbre, son las bases fundacionales de un proceso constructivo de la decisión para ayudar a identificar una buena opción.

Ahora, con el fin de construir un modelo donde es posible representar la independencia entre la verificación de argumentos positivos por un lado, y de argumentos negativos por el otro, lo cual exige por ello mismo un proceso de agregación particular, se introduce el modelo difuso de preferencia-aversión P-A (propuesto inicialmente en [4], y examinado con algún detalle en [5]). De este modo, su metodología obedece al espacio semántico sobre el cual el individuo evalúa sus alternativas y mediante el cual construye sus relaciones básicas de preferencia.

Entonces, el modelo P-A requiere de entrada cuatro intensidades, un par de preferencia y otro par de aversión, a partir de la evaluación independiente de los argumentos positivos y negativos sobre un par de alternativas *x, y* en *X*. Por lo tanto, se definen de entrada cuatro intensidades o grados de pertenencia, los cuales obedecen a sus correspondientes funciones,

$$\mu_{R^+}(x,y),\ \mu_{R^+}^{-1}(x,y): X\times X\to[0,1],$$

$$\mu_{R^-}(x,y),\ \mu_{R^-}^{-1}(x,y): X\times X\to[0,1],$$

representando los grados de intensidad en que un par de elementos $x,y\in X$ verifican las propiedades del respectivo predicado, uno positivo, correspondiente a la relación R^+, y otro negativo, de acuerdo con la relación R^-. De esta forma

$$R^+(x,y)=\left\{\left\langle x,y,\mu_{R^+}(x,y)\right\rangle \middle| x,y\in X\right\}$$

representa el predicado "*x* es al menos tan buena como *y*", tal que $\mu_{R^+}(x,y)\in[0,1]$, y

$$R^-(x,y)=\left\{\left\langle x,y,\mu_{R^-}(x,y)\right\rangle \middle| x,y\in X\right\}$$

representa el predicado "*x* es al menos tan mala como *y*", donde $\mu_{R^-}(x,y)\in[0,1]$.

La inclusión de estas dos relaciones binarias, donde la relación de aversión R^- es positivamente verificada mediante una intensidad independiente de la relación de preferencia R^+, permiten definir seis relaciones básicas, como situaciones fundamentales de preferencia-aversión donde efectivamente los argumentos en contra pueden ser construidos independientemente de los argumentos a favor de una posible decisión. De esta manera se consigue un modelo que representa la racionalidad de pérdidas y ganancias del individuo de manera completa, por medio de una estructura lo suficientemente general y flexible, representando sin pérdida de información los posibles estados del conocimiento de cara a la decisión.

Por lo tanto, las cuatro relaciones de preferencia y aversión débil pueden ser descompuestas en seis relaciones básicas. Estas son, por el lado de la estructura estándar de preferencia débil [2], [10], la preferencia estricta *P*, la indiferencia *I* y la incomparabilidad de preferencia *J*, y por el lado de la aversión débil, la aversión estricta *Z*, la indiferencia negativa *G*, y la incomparabilidad de aversión *H*. La definición nítida de estas relaciones binarias es la siguiente,

$P(x,y)=1$ si y solo si $\mu_{R^+}(x,y)=1$ y $\mu_{R^+}(y,x)=0$,

$I(x,y)=1$ si y solo si $\mu_{R^+}(x,y)=1$ y $\mu_{R^+}(y,x)=1$,

$J(x,y)=1$ si y solo si $\mu_{R^+}(x,y)=0$ y $\mu_{R^+}(y,x)=0$,

$Z(x,y)=1$ si y solo si $\mu_{R^-}(x,y)=1$ y $\mu_{R^-}(y,x)=0$,

$G(x,y)=1$ si y solo si $\mu_{R^-}(x,y)=1$ y $\mu_{R^-}(y,x)=1$,

$H(x,y)=1$ si y solo si $\mu_{R^-}(x,y)=0$ y $\mu_{R^-}(y,x)=0$.

De este modo la estructura básica de preferencia-aversión se puede descomponer en seis situaciones básicas, de acuerdo con la estructura

$$\left\langle R^+,R^-\right\rangle=\left\langle(P,I,J),(Z,G,H)\right\rangle.$$

2.2. MODELO AXIOMÁTICO DE PREFERENCIA-AVERSIÓN

El siguiente es el sistema axiomático en el que se sostiene esta aproximación a las relaciones de preferencia y aversión, denominado modelo P-A [4], [5], el cual extiende un cierto conjunto de axiomas [2], [10] sobre el que se sostiene la estructura de preferencia $R=\langle P,I,J\rangle$.

Axioma IA2: Para todo par de alternativas $x,y\in X$, se tiene que las intensidades de $P(x,y)$, $I(x,y)$ y $J(x,y)$ dependen solamente de los grados de pertenencia de $R^+(x,y)$ y $R^+(y,x)$, denotados por x^+ e y^+ respectivamente, y que las intensidades de $Z(x,y)$, $G(x,y)$ y $H(x,y)$ dependen solamente de los grados de pertenencia de $R^-(x,y)$ y $R^-(y,x)$, denotados por x^- e y^- respectivamente.

De esta forma se afirma la existencia de las funciones continuas $p,i,j:[0,1]^2\to[0,1]$, $z,g,h:[0,1]^2\to[0,1]$, tal que

$$P(x,y)=p\left(x^+,y^+\right),$$

$$I(x,y)=i\left(x^+,y^+\right),$$

$$J(x,y)=j\left(x^+,y^+\right),$$

$$Z(x,y)=z\left(x^{-},y^{-}\right),$$

$$G(x,y)=g\left(x^{-},y^{-}\right),$$

$$H(x,y)=h\left(x^{-},y^{-}\right).$$

Axioma AP2: Las funciones $p\left(x^{+},n\left(y^{+}\right)\right)$, $i\left(x^{+},y^{+}\right)$, $j\left(n\left(x^{+}\right),n\left(y^{+}\right)\right)$, $z\left(x^{-},n\left(y^{-}\right)\right)$, $g\left(x^{-},y^{-}\right)$ y $h\left(n\left(x^{-}\right),n\left(y^{-}\right)\right)$ son no-decrecientes con respecto a ambos argumentos.

Axioma SM2: Las funciones $i\left(x^{+},y^{+}\right)$, $j\left(x^{+},y^{+}\right)$, $g\left(x^{-},y^{-}\right)$ y $h\left(x^{-},y^{-}\right)$ son simétricas.

Entonces, se define el siguiente sistema de ecuaciones en el caso continuo, donde *T*, *S*, *n* son respectivamente una *t*-norma continua, una *t*-conorma continua y una negación estricta [12], tal que

$$S\left(p\left(x^{+},y^{+}\right),i\left(x^{+},y^{+}\right)\right)=x^{+}, \quad (1)$$

$$S\left(p\left(x^{+},y^{+}\right),i\left(x^{+},y^{+}\right),p\left(y^{+},x^{+}\right)\right)=S\left(x^{+},y^{+}\right), \quad (2)$$

$$S\left(p\left(y^{+},x^{+}\right),j\left(y^{+},x^{+}\right)\right)=n\left(x^{+}\right), \quad (3)$$

y de manera análoga se define el siguiente sistema para el caso de la estructura de aversión, donde

$$S\left(z\left(x^{-},y^{-}\right),g\left(x^{-},y^{-}\right)\right)=x^{-}, \quad (4)$$

$$S\left(z\left(x^{-},y^{-}\right),g\left(x^{-},y^{-}\right),z\left(y^{-},x^{-}\right)\right)=S\left(x^{-},y^{-}\right), \quad (5)$$

$$S\left(z\left(y^{-},x^{-}\right),h\left(y^{-},x^{-}\right)\right)=n\left(x^{-}\right). \quad (6)$$

Este sistema de ecuaciones permite construir un orden positivo, basado en las cuatro relaciones $\left\langle p,p^{-1},i,j\right\rangle$, y un orden negativo sobre *X*, basado en las cuatro relaciones $\left\langle z,z^{-1},g,h\right\rangle$; el primero acerca de los atributos positivos y el otro acerca de los atributos negativos de *x* al ser comparado con *y*. Por lo tanto, bajo esta caracterización, la percepción subjetiva negativa sobre la preferencia débil x^{+} se verifica por medio de una identificación directa sobre los atributos negativos, dada por x^{-}, donde $n\left(x^{+}\right)\neq x^{-}$.

A diferencia del modelo estándar de preferencia, donde solo se tiene en cuenta la estructura dada por $R=\left\langle P,I,J\right\rangle$, ahora se cuenta con cuatro valores, dos para la información positiva x^{+} e y^{+}, y dos para la información negativa, x^{-} e y^{-}, los cuales se verifican y descomponen de manera independiente.

2.3. PROPIEDADES Y SOLUCIONES DEL MODELO DE PREFERENCIA-AVERSIÓN

Entonces se tiene que la única combinación posible entre *T*, *S*, y *n*, que permite preservar tanto el principio de contradicción como el del tercero excluido, es la terna de De Morgan con la *t*-norma *T* y *t*-conorma *S* de Lukasiewicz, definidas unívocamente a partir de un determinado automorfismo φ [2], [10], [12],

$$T(x,y)=\varphi^{-1}\left(\max\left\{\varphi(x)+\varphi(y)-1,0\right\}\right),$$

$$S(x,y)=\varphi^{-1}\left(\min\left\{\varphi(x)+\varphi(y),1\right\}\right),$$

$$n(x)=\varphi^{-1}\left(1-\varphi(x)\right),$$

donde *n* es una negación de tipo fuerte.

En general, tomando de ahora en adelante la terna de De Morgan $\left\langle T,S,n\right\rangle_{\varphi}$ de tipo fuerte, se identifican las siguientes cotas para $\left\langle p,i,j\right\rangle_{\varphi}$ [2], tal que se satisfacen las ecuaciones (1) y (3),

$$T\left(x^{+},n\left(y^{+}\right)\right)\le p\left(x^{+},y^{+}\right)\le\min\left\{x^{+},n\left(y^{+}\right)\right\},$$

$$T\left(x^{+},y^{+}\right)\le i\left(x^{+},y^{+}\right)\le\min\left\{x^{+},y^{+}\right\},$$

$$T\left(n\left(x^{+}\right),n\left(y^{+}\right)\right)\le j\left(x^{+},y^{+}\right)\le\min\left\{n\left(x^{+}\right),n\left(y^{+}\right)\right\},$$

y para $\left\langle z,g,h\right\rangle_{\varphi}$, las respectivas cotas tal que se satisfacen las ecuaciones (4) y (6) (la demostración para el caso de aversión se sigue directamente del caso de preferencia [2]),

$$T\left(x^{-},n\left(y^{-}\right)\right)\le z\left(x^{-},y^{-}\right)\le\min\left\{x^{-},n\left(y^{-}\right)\right\},$$

$$T\left(x^{-},y^{-}\right)\le g\left(x^{-},y^{-}\right)\le\min\left\{x^{-},y^{-}\right\},$$

$$T\left(n\left(x^{-}\right),n\left(y^{-}\right)\right)\le h\left(x^{-},y^{-}\right)\le\min\left\{n\left(x^{-}\right),n\left(y^{-}\right)\right\}.$$

También se tiene que para toda solución $\left\langle p,i,j\right\rangle_{\varphi}$ de (1) y (3), se cumplen las siguientes propiedades [2], donde $x^{+},y^{+}\in[0,1]$:

$$T\left(p\left(x^{+},y^{+}\right),p\left(y^{+},x^{+}\right)\right)=0,$$

$$T\left(p\left(x^{+},y^{+}\right),i\left(x^{+},y^{+}\right)\right)=0,$$

$$T\left(p\left(x^{+},y^{+}\right),j\left(x^{+},y^{+}\right)\right)=0,$$

$$T\left(i\left(x^+,y^+\right),j\left(x^+,y^+\right)\right)=0,$$

$$S\left(S\left(p\left(x^+,y^+\right),p\left(y^+,x^+\right)\right),S\left(i\left(x^+,y^+\right),j\left(x^+,y^+\right)\right)\right)=1.$$

y para toda solución $\langle z,g,h\rangle_\varphi$ de (4) y (6), se cumplen las siguientes propiedades, donde $x^-,y^-\in[0,1]$:

$$T\left(z\left(x^-,y^-\right),z\left(y^-,x^-\right)\right)=0,$$

$$T\left(z\left(x^-,y^-\right),g\left(x^-,y^-\right)\right)=0,$$

$$T\left(z\left(x^-,y^-\right),h\left(x^-,y^-\right)\right)=0,$$

$$T\left(g\left(x^-,y^-\right),h\left(x^-,y^-\right)\right)=0,$$

$$S\left(S\left(z\left(x^-,y^-\right),z\left(y^-,x^-\right)\right),S\left(g\left(x^-,y^-\right),h\left(x^-,y^-\right)\right)\right)=1.$$

De esta manera, existe una solución para el sistema (1)-(3), si y sólo si las funciones *p*, *i*, *j* se definen de la siguiente forma, donde *i* y *j* son mutuamente excluyentes [2],

$$p\left(x^+,y^+\right)=\min\left\{x^+,n\left(y^+\right)\right\},$$

$$i\left(x^+,y^+\right)=T\left(x^+,y^+\right),$$

$$j\left(x^+,y^+\right)=T\left(n\left(x^+\right),n\left(y^+\right)\right).$$

También se tiene la solución para las ecuaciones (1) y (3), tal que *P* es asimétrica, donde [2],

$$p\left(x^+,y^+\right)=T\left(x^+,n\left(y^+\right)\right),$$

$$i\left(x^+,y^+\right)=\min\left\{x^+,y^+\right\},$$

$$j\left(x^+,y^+\right)=\min\left\{n\left(x^+\right),n\left(y^+\right)\right\}.$$

Además, la solución para *p*, *i* y *j* que permite su verificación simultánea, tal que

$$T^*\left(p\left(x^+,y^+\right),p\left(y^+,x^+\right)\right)=T^*\left(i\left(x^+,y^+\right),j\left(y^+,x^+\right)\right),$$
$$S^*\left(p\left(x^+,y^+\right),p\left(y^+,x^+\right),i\left(x^+,y^+\right),j\left(y^+,x^+\right)\right)=1,$$

para una *t*-norma T^* y una *t*-conorma S^* multiplicativas, es la siguiente [16],

$$p\left(x^+,y^+\right)=T^*\left(x^+,n\left(y^+\right)\right),$$

$$i\left(x^+,y^+\right)=T^*\left(x^+,y^+\right),$$

$$j\left(x^+,y^+\right)=T^*\left(n\left(x^+\right),n\left(y^+\right)\right).$$

Del mismo modo se pueden caracterizar las siguientes soluciones para el sistema (4)-(6), análogamente a como se ha hecho para el sistema (1)-(3), para una terna de De Morgan de tipo fuerte. De esta manera, la solución para el sistema (4)-(6) de $\langle z,g,h\rangle_\varphi$ es la siguiente,

$$z\left(x^-,y^-\right)=\min\left\{x^-,n\left(y^-\right)\right\},$$

$$g\left(x^-,y^-\right)=T\left(x^-,y^-\right),$$

$$h\left(x^-,y^-\right)=T\left(n\left(x^-\right),n\left(y^-\right)\right).$$

También se tiene que la solución para las ecuaciones (4) y (6), tal que *Z* es asimétrica, es la siguiente,

$$z\left(x^-,y^-\right)=T\left(x^-,n\left(y^-\right)\right),$$

$$g\left(x^-,y^-\right)=\min\left\{x^-,y^-\right\},$$

$$h\left(x^-,y^-\right)=\min\left\{n\left(x^-\right),n\left(y^-\right)\right\}.$$

Además, la solución para $\langle z,g,h\rangle_\varphi$, que permite la coexistencia de las cuatro situaciones de aversión, tal que

$$T^*\left(z\left(x^-,y^-\right),z\left(y^-,x^-\right)\right)=T^*\left(g\left(x^-,y^-\right),h\left(x^-,y^-\right)\right),$$
$$S^*\left(z\left(x^-,y^-\right),z\left(y^-,x^-\right),g\left(x^-,y^-\right),h\left(x^-,y^-\right)\right)=1,$$

es tal que,

$$z\left(x^-,y^-\right)=T^*\left(x^-,n\left(y^-\right)\right),$$

$$g\left(x^-,y^-\right)=T^*\left(x^-,y^-\right),$$

$$h\left(x^-,y^-\right)=T^*\left(n\left(x^-\right),n\left(y^-\right)\right).$$

De esta manera, el modelo IA2-SM2 permite la representación tanto del predicado de preferencia y de sus relaciones básicas, como del de aversión y su respectiva descomposición, a partir de los grados de intensidad correspondientes a los datos de entrada x^+, y^+ y x^-, y^-, por medio de las funciones $\langle p,p^{-1},i,j\rangle$ y $\langle z,z^{-1},g,h\rangle$, respectivamente.

Dicha representación admite una total y perfecta separabilidad entre sus componentes, tal como ocurre con las relaciones de preferencia débil, identificándose las relaciones básicas de preferencia estricta, indiferencia e incomparabilidad, o con las relaciones de aversión débil, de donde se identifican las de aversión estricta, indiferencia e incomparabilidad negativas. Esto es consecuencia de que para ambos órdenes, se tiene una partición perfecta del conjunto de alternativas, de manera que,

$$p\left(x^+,y^+\right)+i\left(x^+,y^+\right)+p\left(y^+,x^+\right)+j\left(y^+,x^+\right)=1$$

y

$$z\left(x^{-},y^{-}\right)+g\left(x^{-},y^{-}\right)+z\left(y^{-},x^{-}\right)+h\left(y^{-},x^{-}\right)=1.$$

2.4. LA ESTRUCTURA COMPLETA DE PREFERENCIA-AVERSIÓN

El modelo IA2-SM2 de preferencia-aversión (P-A), incorpora un tipo de racionalidad que enmarca las alternativas en términos de pérdidas y ganancias, siguiendo la intuición principal de la CPT, asignando dos valores diferentes e independientes para expresar preferencia y aversión débiles (se advierte que en la CPT se agregan las pérdidas y ganancias de manera dependiente, por medio de un funcional acumulativo como lo es la integral de Choquet [7], [15]). Entonces es posible ordenar distintas piezas de información de acuerdo con la fuerza o intensidad de los atributos positivos y negativos de las opciones disponibles.

De esta forma, el espacio evaluativo de este modelo corresponde con el tipo bipolar bivariado [1], [7], [11], donde los aspectos positivos y negativos pueden ser simultáneamente verificados. Por lo tanto, la bipolaridad del significado del predicado de preferencia-aversión $R=\left\langle R^{+},R^{-}\right\rangle$ se mide por medio de dos escalas independientes.

En este marco, la ignorancia tiene un papel primordial, pues representa el estado a partir del cual las dos escalas pueden ser conjuntamente examinadas. Tal estado representa la situación en que no se cuenta con información alguna sobre las alternativas, representando el punto de partida para todo proceso de aprendizaje sobre el problema de decisión [9].

Una especificación conjunta del orden de preferencia y del de aversión genera una estructura combinada de dieciséis posibles situaciones para representar del estado de conocimiento del individuo, conformada por diez relaciones compuestas, *PZ*, *PA*, *PG*, *PH*, *IZ*, *IG*, *IH*, *JZ*, *JG* y *JH*. Estas relaciones se encuentran definidas en la tabla 2, para todo par de alternativas $x,y\in X$.

Tabla 1: La estructura completa de preferencia-aversión.

$R=R^{+},R^{-}$	$Z(x,y)$	$Z(y,x)$	$G(x,y)$	$H(x,y)$
$P(x,y)$	$PZ(x,y)$	$PA(x,y)$	$PG(x,y)$	$PH(x,y)$
$P(y,x)$	$PA(y,x)$	$PZ(y,x)$	$PG(y,x)$	$PH(y,x)$
$I(x,y)$	$IZ(x,y)$	$IZ(y,x)$	$IG(x,y)$	$IH(x,y)$
$J(x,y)$	$JZ(x,y)$	$JZ(y,x)$	$JG(x,y)$	$JH(x,y)$

De esta manera, examinando la combinación entre las relaciones básicas del modelo P-A, se tienen diez relaciones compuestas que generan un sistema relacional combinado de preferencia y aversión. Dicho sistema representa una graduación entre diversas situaciones de decisión, partiendo de argumentos independientes sobre las pérdidas y ganancias y luego verificando las situaciones obtenidas mediante la agregación de ambas dimensiones, de acuerdo con la intersección, por medio de la *t*-norma T^{M}=min, de las relaciones básicas de preferencia y aversión.

Entonces, los distintos estados epistémicos representados por cada una de estas relaciones constituyen un marco representativo de la racionalidad del individuo y sus posibilidades frente a una decisión. De la preferencia estricta se obtienen cuatro posibles relaciones, *PZ*, *PA*, *PG* y *PH*. La primera es una relación de máximo conflicto, denominada *incomparabilidad por ambivalencia* (*PZ*), dado que coexisten dos valoraciones antagónicas de preferencia y aversión estrictas. La segunda, en cambio, es una relación de *preferencia estricta fuerte* (*PA*), donde se tiene preferencia estricta y aversión estricta inversa, lo cual hace que la preferencia estricta sea aún más fuerte debido a que la otra alternativa es peor.

Este análisis permite identificar que *x* puede ser mejor y a la vez peor que *y*, caso en el cual no hay una decisión fácil debido a la ambivalencia natural de ciertas situaciones, o que *x* es mejor e *y* es peor, lo cual hace aún más nítida y fácil la elección de *x*. La teoría ayuda a entender, bajo un enfoque relacional, porqué el resultado final puede ser decisivo o no.

Siguiendo con la descripción del sistema relacional combinado, la relación de *preferencia pseudo-estricta* (*PG*) refleja una situación en que se verifica preferencia estricta aunque las dos alternativas son igual de malas, reflejando un cierto grado de insatisfacción conflictiva, mientras que la relación de *preferencia semi-estricta* (*PH*) representa el caso en el cual se tiene preferencia estricta pero no se sabe nada acerca de los atributos negativos de las opciones, pues existe un grado de incomparabilidad por ignorancia sobre la relación de aversión débil.

Por el lado de la indiferencia positiva, su combinación con la aversión estricta genera una relación compuesta de *aversión pseudo-estricta* (*IZ*), pues son igual de buenas pero una es peor que la otra. En cuanto a la combinación entre la indiferencia positiva y la negativa, el resultado es el de la *indiferencia fuerte* (*IG*), pues ambas alternativas son igual de buenas y de malas, es decir, son muy similares entre sí. Y por la combinación entre indiferencia e incomparabilidad sobre la aversión débil, se obtiene la *semi-indiferencia* (*IH*), pues no se sabe nada de los atributos negativos de *x* e *y* pero sí se sabe que son igual de buenas.

Por último, se tienen las relaciones compuestas de incomparabilidad por ignorancia sobre la preferencia débil *JZ*, *JG* y *JH*. La primera relación es de *aversión semi-estricta* (*JZ*), la cual representa el caso en el cual se tiene aversión estricta pero no se sabe nada acerca de los atributos positivos de las opciones, pues existe un grado de incomparabilidad por ignorancia sobre la relación de preferencia

débil. La segunda es una relación de *semi-indiferencia negativa* (*JG*), donde no se sabe nada acerca de los atributos positivos entre *x* e *y* pero sí se sabe que son igual de malas, reflejando una inconformaidad absoluta sobre las alternativas disponibles. Finalmente se obtiene la relación de absoluta ignorancia, denominada *incomparabilidad por ignorancia* (*JH*), donde el individuo no cuenta con ningún tipo de información relevante.

3 CONCLUSIÓN

El modelo IA2-SM2 y su estructura P-A permite identificar de manera fiable las distintas situaciones o estados cognitivos del individuo racional frente a un problema de decisión. Como atributo principal, esta aproximación permite la agregación independiente de pérdidas y ganancias, o de percepciones negativas y positivas acerca de las alternativas de *X*. De este modo, se mantiene la estructura estándar de preferencia, y se incorpora la correspondiente estructura de aversión, con el fin de examinarlas conjuntamente.

Para una investigación futura, se planea examinar con mayor detalle las relaciones entre esta propuesta y otros modelos de preferencia, con el fin de determinar sus similitudes y diferencias, identificando un espacio semántico común, a partir del cual poder formular y solucionar el problema de decisión en base al tipo de información disponible. De esta manera, se puede entender mejor el proceso mediante el cual un individuo racional ordena sus alternativas e intenta encontrar la mejor de éstas.

Agradecimientos

Esta investigación ha sido parcialmente subvencionada por el Gobierno de España, proyecto TIN2009-07901.

Referencias

[1] D. Dubois, H. Prade. An introduction to bipolar representations of information and preference. International Journal of Intelligent Systems, 23 866-877 (2008)

[2] J. Fodor, M. Roubens. Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht (1994)

[3] C. Franco, J. Montero. Organizing information by fuzzy preference structures –Fuzzy preference semantics. Proceedings ISKE Conference, Beijing, November 15-16 (2010), pages 135-140

[4] C. Franco, J. Montero, J.T. Rodríguez. Aggregation weights for a preference-aversion model. Proceedings World Conference on Soft Computing, San Francisco, May 23-26 (2010), paper 199

[5] C. Franco, J. Montero, J.T. Rodríguez. Partial comparability and preference-aversion models. ISKE2011 (2011)

[6] J. Goguen. The logic of inexact concepts. Synthese 19, 325-373 (1969)

[7] M. Grabisch, Ch. Labreuche. A decade of application of the Choquet and Sugeno integrals in multicriteria decision aid. A Quarterly Journal of Operations Research 6, 1-44, (2008)

[8] D. Kahneman, A. Tversky. Prospect theory: an analysis of decision under risk. Econometrica 47, 263-291 (1979)

[9] J. Montero, D. Gómez, H. Bustince. On the relevance of some families of fuzzy sets. Fuzzy Sets and Systems 158, 2429-2442 (2007)

[10] J. Montero, J. Tejada, C. Cutello. A general model for deriving preference structures from data. European Journal of Operational Research 98, 98-110 (1997)

[11] J.T. Rodríguez, C. Franco, J. Montero. On the relationship between bipolarity and fuzziness. Proceedings EUROFUSE Conference, Régua, Portugal, September 21-23 (2011)

[12] B. Schweizer, A. Sklar. Probabilistic Metric Spaces. North-Holland, Amsterdam (1983)

[13] A. Smith. The Theory of Moral Sentiments. Cambridge University Press, Cambridge (2002)

[14] E. Trillas. On a model for the meaning of predicates (A naïve approach to the genesis of fuzzy sets). In: Views of Fuzzy Sets and Systems from Different Perspectives (ed.: Rudolf Seising), Springer, 175-205 (2009)

[15] A. Tversky, D. Kahneman. Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297-323 (1992)

[16] B. Van de Walle, B. De Baets, E. Kerre. Characterizable fuzzy preference structures. Annals of Operations Research 80, 105-136 (1998)

[17] L.A. Zadeh. Fuzzy sets. Information and Control 8, 338-353 (1965)

Fuzzy quadratic programming solving a portfolio selection problem

Carlos Cruz[1], **Ricardo C. Silva**[2], **José Luis Verdegay**[1]

[1]*Department of Computer Science and Artificial Intelligence, University of Granada, E-18071, Granada, Spain*,
{carloscruz,verdegay}@decsai.ugr.es
[2]*Institute of Science and Technology, Federal University of São Paulo, rua Talim, 330, 12231-280, São José dos Campos, SP, Brazil*,
ricardo.coelho@unifesp.br

Abstract

Although quadratic mathematical programming can be applied to general linear programming problems, it can also be applied in a special class of nonlinear programming. As ambiguity and vagueness are natural and ever-present in real-life situations requiring solutions, it makes perfect sense to attempt to address them using fuzzy quadratic programming problems. This way of problem modeling is applied in an increasing variety of practical fields especially those with portfolio selection problems. In this work we present a novel approach that slve quadratic programming problems under different kind of uncertainties in its data. Finally, a portfolio selection problem with uncertain data is formulated and solved as a fuzzy quadratic programming problem, in order to show the efficiency of the method presented in this work.

Keywords: Fuzzy logic, Fuzzy quadratic optimization, Portfolio selection problem.

1 INTRODUCTION

The classical portfolio selection problem was formulated by Markowitz [9] as a quadratic programming problem, based on the time series of the return rate. Quadratic programming refers to a special class of nonlinear programming but can also be viewed as a generalization of the linear programming problems in which the objective function is quadratic and the constraints are linear.

The Markowitz model is a well-established method which provides good results. However, it is also well known that the sample average is not always the best option for describing a data set. Besides, all too often, the decision maker cannot formulate all the data precisely [5] and uncertainty is an inherent feature of most business environments. Decisions must often be made with uncertain and incomplete knowledge about future circumstances.

Fuzzy methodology allows the incorporation of uncertainty into databases and also the incorporation of subjective characteristics into the models, these being important elements for establishing different estimations of risk and expected return. This done, the imperfect knowledge of the returns on the assets and the uncertainty involved in the behavior of financial markets may then also be introduced by means of fuzzy quantities and/or fuzzy constraints.

With this in mind, the objective of this paper is to show an application relating to the quadratic topic and to outline some soft computing models used to solve fuzzy quadratic mathematical programming problems.

The paper is organized as follows: Section 2 presents a brief description of a fuzzy quadratic programming problem some methods that were developed to solve fuzzy quadratic programming problems with uncertainties in the relationships or coefficients; Section 3 demonstrate that a fuzzy quadratic programming problem can be transformed in a parametric quadratic problem and the solutions obtained by traditional techniques is a fuzzy solution acceptable; Section 4 show a logistic problem that selects securities to generate a portfolio; Finally, conclusions are presented in Section 5.

2 FUZZY QUADRATIC PROGRAMMING

As it is well known Quadratic Programming, QP, represents a quadratic objective function and linear constraints. Assume $\mathbf{x}$ belongs to $\mathbb{R}^n$ space. The $n \times n$ matrix $\mathbf{H}$ is symmetric, and $\mathbf{c}$ is any $n \times 1$ vector. This set of problems can be formalized in the following form:

$$\begin{array}{ll} \text{Minimize} & f(\mathbf{x}) = \mathbf{c^t x} + \frac{1}{2}\mathbf{x^t H x} \\ \text{subject to} & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array} \quad (1)$$

where $\mathbf{v}^t$ indicates the vector transpose of $\mathbf{v}$. The notation $\mathbf{Ax} \leq \mathbf{b}$ means that every entry of the vector $\mathbf{Ax}$ is less than or equal to the corresponding entry of the vector $\mathbf{b}$.

There are several classes of problems that are naturally expressed as quadratic problems. Examples of such problems can be found in game theory, engineering modeling, design and control, problems involving economies of scale, facility allocation and location problems, problems in microeconomics, as portfolio analysis, structural analysis, VLSI design, discrete-time stabilization, optimal and fuzzy control, finite impulse response design, optimal power flow, economic dispatch among others.

Moreover there are some cases that the parameters of the real-world problems are seldom known exactly and have to be estimated by the decision maker. Therefore the application of the fuzzy logic is also an way to describe mathematically these vagueness. The uncertainties can be found in the relation, constants, decision variables or in all parameters of the problem. Some authors have applied the Soft Computing methodologies to quadratic programming. In [1] is described a method to solve a quadratic programming problem with fuzzy coefficients, which divides the original fuzzy quadratic problem into two classical quadratic problems. An extended model to the fuzzy quadratic problem described in [2] uses the approach proposed by Zimmermann [15, 16]. This approach solves a classic problems that is not violated and another problem that is totally violated. In [6, 7] is discussed as transforming a fuzzy quadratic programming problem in a two-level mathematical programming problem for finding the bounds of the fuzzy objective values. In [11] is proposed a parametric approach which provides a fuzzy solution to the problem, that involves all particular-punctual-solutions provided by their method.

3 A GENERAL MODEL FOR FUZZY QUADRATIC PROGRAMMING

In this section will be presented some proposed methods to solve quadratic programming problems with fuzzy parameters. The fuzzy parameters can be in the coefficients of the set of constraints or the costs of the objective function and in the relation of the constraints.

A quadratic programming problems with fuzzy coefficients and relation can be formulated on the following way

$$\begin{aligned} \min \quad & f(\mathbf{x}) = \tilde{\mathbf{c}}^{\mathbf{t}}\mathbf{x} + \tfrac{1}{2}\mathbf{x}^{\mathbf{t}}\tilde{\mathbf{H}}\mathbf{x} \\ \text{s.t.} \quad & \tilde{\mathbf{A}}\mathbf{x} \lesssim \tilde{\mathbf{b}} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{2}$$

where $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{b}}$ are an $n \times 1$ vector and an $m \times 1$ vector, respectively, with fuzzy numbers, $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{A}}$ are respectively an $n \times n$ matrix and an $m \times n$ matrix with fuzzy components. The vagueness are represented for membership functions defined by decision maker.

3.1 PARAMETRIC IDEA TO SOLVE A FUZZY QUADRATIC PROGRAMMING

The fuzzy parameters of the quadratic problem (2) are defined with fuzzy nature, that is, some violations in the accomplishment of such objective function and restrictions are permitted. Therefore this fuzzy parameters can be determined by the decision maker.

It is clear that each membership function will give the membership (satisfaction) degree such that any $\mathbf{x} \in \mathbb{R}^n$ accomplishes the corresponding fuzzy objective function and constraint upon which it is defined. These membership functions can be formulated as follows

$$\mu_i : \mathbb{R} \to (0,1], \;\; i \in I$$

where μ is an linear membership function (and formally can also be a non linear one), and $\mathbf{I}$ is the set that contains all fuzzy parameters.

In order to solve this problem in a two-phase method, first let us define for each fuzzy constraint, $i \in I$

$$X_i = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \tilde{A}_i\mathbf{x} \lesssim \tilde{b}_i, \mathbf{x} \geq 0 \right\}.$$

If $\mathbf{X} = \bigcap_{i \in I} X_i$ then the former fuzzy quadratic problem can be addressed in a compact form as

$$\min \{ f(\mathbf{x}) \mid \mathbf{x} \in \mathbf{X} \}$$

It is clear that $\forall \alpha \in (0,1]$, an α-cut of the fuzzy constraint set will be the classical set

$$X(\alpha) = \{ \mathbf{x} \in \mathbb{R}^n \mid \mu_X(\mathbf{x}) \geq \alpha \}$$

where $\forall \mathbf{x} \in \mathbb{R}^n$,

$$\mu_X(\mathbf{x}) = \inf \mu_i(\mathbf{x}(1-\alpha)), i \in I$$

Hence an α-cut of the i-th constraint will be denoted by $X_i(\alpha)$. Therefore, if $\forall \alpha \in (0,1]$,

$$S(\alpha) = \{ \mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = \min \; f(\mathbf{y}), \, \mathbf{y} \in X(\alpha) \}$$

the fuzzy solution to the problem will therefore be the fuzzy set defined by the following membership function

$$S(\mathbf{x}) = \begin{cases} \sup\{\alpha : \mathbf{x} \in S(\alpha)\} & \mathbf{x} \in \bigcup_\alpha S(\alpha) \\ 0 & \text{otherwise.} \end{cases}$$

Provided that $\forall \alpha \in (0,1]$,

$$X(\alpha) = \bigcap_{i \in I} \{ \mathbf{x} \in \mathbb{R}^n \mid (Ax)_i \leq r_i(\alpha), \mathbf{x} \geq 0, \mathbf{x} \in \mathbb{R}^n \}$$

with $r_i(\alpha) = b_i + d_i(1-\alpha)$. The operative solution to the former problem can be found, α-cut by α-cut. Thus, the fuzzy quadratic programming problem was parameterized at the end of the first phase. In the second phase the parametric quadratic programming problem is solved for each of the different α values using conventional quadratic programming techniques.

The obtained results for each α value generate a set of solutions $S(\mathbf{x})$ and then the Representation Theorem can be used to integrate all these specific alpha-solutions. It is then demonstrated that the outlined solution to the parametric method is a valid solution to the fuzzy quadratic problem.

3.2 QUADRATIC PROGRAMMING PROBLEMS UNDER UNCERTAIN ENVIRONMENT

The uncertainties can be in different ways in a mathematical programming problem, in particular case, as it is presented in Problem (2). It will be shown in this section different approaches to solve quadratic programming problems with imprecise coefficients and relations.

3.2.1 Methods to solve quadratic programming problems with fuzzy relations

By using the parametric idea describes above a general method to solve quadratic programming problems with fuzzy relations is presented here as described in [**?**]. This method is an extension of method that were developed to solve fuzzy linear programming problems.

$$\begin{array}{ll} \min & \mathbf{c^t x} + \frac{1}{2}\mathbf{x^t H x} \\ \text{s.t.} & (Ax)_i \leq b_i + d_i(1-\alpha), i \in I \\ & x_j \geq 0, j \in J, \alpha \in (0,1]. \end{array} \tag{3}$$

where d_i is the violation permitted for each one constraint i, $\mathbf{H}$ and $\mathbf{A}$ are an $n \times n$ matrix and an $m \times n$ matrix with real numbers, and $\mathbf{c}$ and $\mathbf{b}$ are an n vector and m vector with real numbers.

3.2.2 Methods to solve quadratic programming problems with fuzzy coefficients in the objective function

A multiobjective approach is developed in [4] to solve fuzzy nonlinear programming problems. This approach will use to solve Problem (2) with only fuzzy costs. The fuzzy solution is obtained by transforming a fuzzy nonlinear programming problem into a parametric multiobjective nonlinear programming problem in which the parameter $\alpha,\ \beta_i,\ \gamma_{ij} \in [0,1]$ is treated as new decision variable. Besides the decision variable α, they also consider $n(n+1)+1$ new decision variable $\beta_j,\ j = 1,\ldots,p$, to transform the intervals $I_j(\alpha) = [h_j^{-1}(1-\alpha), g_j^{-1}(1-\alpha)]$ and $I_{ij}(\alpha) = [h_{ij}^{-1}(1-\alpha), g_{ij}^{-1}(1-\alpha)]$ into functions of the form $z_j(\alpha,\beta_j) = h_j^{-1}(1-\alpha) + \beta_j(g_j^{-1}(1-\alpha) - h_j^{-1}(1-\alpha))$ and $z_{ij}(\alpha,\gamma_{ij}) = h_{ij}^{-1}(1-\alpha) + \gamma_{ij}(g_{ij}^{-1}(1-\alpha) - h_{ij}^{-1}(1-\alpha))$.

Consequently, the objective is to minimize $f(\mathbf{x},\alpha,\beta_1,\ldots,\beta_n,\gamma_{11},\ldots,\gamma_{nn})$ and, in the same time, to maximize and minimize $\alpha,\ \beta_i,\ \gamma_{ij},\ i,j = 1,\ldots,n$, simultaneously. Therefore, the multiobjective quadratic programming problem is stated as follows:

$$\begin{array}{ll} \min & \left[\sum_{i=1}^n \left((c_\alpha^L)_i + \beta_i\left((c_\alpha^U)_i - (c_\alpha^L)_i\right)\right)x_i + \right. \\ & \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \left((q_\alpha^L)_{ij} + \gamma_{ij}\left((q_\alpha^U)_{ij} - (q_\alpha^L)_{ij}\right)\right)x_i x_j, \\ & \alpha, 1-\alpha, \beta_1, 1-\beta_1, \ldots, \beta_n, 1-\beta_n, \\ & \left. \gamma_{11}, 1-\gamma_{11}, \ldots, \gamma_{nn}, 1-\gamma_{nn}\right] \\ \text{s.t.} & \mathbf{Ax} \leq \mathbf{b},\ \mathbf{x} \geq \mathbf{0}, \\ & \alpha,\beta_j,\gamma_{ij} \in [0,1],\ i,j = 1,\ldots,n \end{array} \tag{4}$$

where the optimal solution to quadratic programming problem is composed of the solutions with maximum values for a function $f(\mathbf{x},\alpha,\beta_1,\ldots,\beta_n,\gamma_{11},\ldots,\gamma_{nn})$ for each value of the parameters $\alpha,\ \beta_k,\ \gamma_{kp},\ k,p = 1,\ldots,n$.

3.2.3 Methods to solve quadratic programming problems with fuzzy coefficients in the set of constraints

By using the multiobjective approach described previous, an approach to solve quadratic programming problems with fuzzy coefficients in the constraints set can be formulated into a multiobjective quadratic programming problem in which the parameter $\alpha,\beta_{ij} \in [0,1],\ i = 1,\ldots,m, j = 1,\ldots,n+1$ is treated as new decision variable.

Consequently, by minimizing $f(\mathbf{x})$ and by maximizing and minimizing $\alpha,\ \beta_{ij},\ i = 1,\ldots,m,\ j = 1,\ldots,n+1$, simultaneously. Therefore, the multiobjective quadratic programming problem is stated as follows:

$$\begin{array}{ll} \min & \left[\mathbf{c^t x} + \frac{1}{2}\mathbf{x^t Q x}, \lambda, \alpha, 1-\alpha, \beta_{11}, 1-\beta_{11}, \ldots, \right. \\ & \left. \beta_{m,n+2}, 1-\beta_{m,n+2}\right] \\ \text{s.t.} & \sum_{j=1}^n \left((a_\alpha^L)_{ij} + \beta_{ij}\left((a_\alpha^U)_{ij} - (a_\alpha^L)_{ij}\right)\right)x_j \leq (b_\alpha^L)_i + \\ & + \beta_{i,n+1}\left((b_\alpha^U)_i - (b_\alpha^L)_i\right) \\ & \mathbf{x} \geq \mathbf{0}, \alpha, \beta_{ij} \in [0,1],\ i = 1,\ldots,m,\ j = 1,\ldots,n+1 \end{array} \tag{5}$$

where it considers $m(n+1)$ new decision variables α and $\beta_{ij},\ i = 1,\ldots,m,\ j = 1,\ldots,n+1$, to transform the intervals $I_{ij}(\alpha) = [h_{ij}^{-1}(1-\alpha), g_{ij}^{-1}(1-\alpha)]$ into functions of the form $z_{ij}(\alpha,\beta_{ij}) = h_{ij}^{-1}(1-\alpha) + \beta_{ij}(g_{ij}^{-1}(1-\alpha) - h_{ij}^{-1}(1-\alpha))$.

4 PORTFOLIO SELETION PROBLEM

The classical portfolio selection problem was formulated by Markowitz [9] as a quadratic programming problem, based on the time series of the return rate. The mean-variance approach decides the best investing rate to each

of investments, which minimizes the risk or the variance of the profits to affirm the least rate of the expected return rate, ρ, which a decision maker expects.

This classical mean-variance model is a well-established method which provides good results. However, it is also well known that the sample average is not always the best option for describing a data set. Fuzzy methodology allows to incorporate uncertainty into databases and also to incorporate subjective characteristics into the models, which are basic aspects for establishing different estimations of risk and expected return, described in [1, 3, 8, 12, 13, 14]. Then the imperfect knowledge of the returns on the assets and the uncertainty involved in the behaviour of financial markets may also be introduced by means of fuzzy quantities and/or fuzzy constraints.

In order to show the performance of our method, we used the set of historical data shown in Table 1 introduced by Markowitz. The columns 2-4 represent United States Steel, General Motors, Firestone and Sharon Steel securities data, respectively. The returns on the nine securities, during the years 1945-54, are presented in Table 1.

This example will consider performances of portfolios with respect to "return" defined. This assumes that a dollar of realized or unrealized capital gains is exactly equivalent to a dollar of dividends, no better and no worse. This assumption is appropriate for certain investors, for example, some types of tax-free institutions. Other ways of handling capital gains and dividends, which are appropriate for other investors, can be viewed in [10].

Therefore, this problem with the use of fuzzy sets can be formulated in the following form:

$$\begin{aligned} \min \quad & \mathbf{x^t}\tilde{\sigma}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x^t}\tilde{\mathbf{E}}(\mathbf{R}) \gtrsim \tilde{\rho} \\ & \mathbf{1x} = 1 \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{6}$$

where $\mathbf{x}$ is an n vector that represents the percentage of money invested in asset, i.e., the proportion of total investment funds devoted to each security; $\mathbf{E}(\mathbf{R})$ is the average vector of returns over m periods because $\mathbf{R} = [r_{ij}]$ is an $m \times n$ matrix that represents the random variables of the returns of asset varying in m discrete times; ρ is a parameter representing the minimal rate of return required by an investor; and $\sigma = [\sigma_{ij}^2]$ is a covariance $n \times n$ matrix between returns of asset

Here we show the results obtained for the approaches by the fuzzy quadratic programming methods introduced in Sub-section 3.2. All simulations described in this work were solved by using the linear approach of the h and g membership functions. In Tables 2-4 were presented the solution of the real-world portfolio selection problem to the different kind of imprecise data.

Table 2 presents the results of Problem (6) with only fuzzy relation. This problem was solved by using the approach presented in Sub-sub-section 3.2.1 to different values of $\alpha \in [0,1]$.

Table 2 presents the results of Problem (6) with only fuzzy costs. This problem was solved by using the approach presented in Sub-sub-section 3.2.2 to different values of $\alpha \in [0,1]$.

Table 2 presents the results of Problem (6) with only fuzzy coefficients in the constraints set. This problem was solved by using the approach presented in Sub-sub-section 3.2.3 to different values of $\alpha \in [0,1]$.

The tests were all performed on a PC with two 2.26GHZ Intel® Core(TM) 2 Duo processors, 4GB RAM running Windows Vista operational system. All the problems presented in this work were resolved using **fmincon** and **fminimax** functions to solve constraint programming problems of *ToolBox Optimization* of MATLAB® 7.4.0 program.

5 CONCLUSIONS

Fuzzy quadratic programming problems are of utmost importance in an increasing variety of practical fields because real-world applications inevitably involve some degree of uncertainty or imprecision, for example in the logistics management.

One of these problems is the portfolio selection problem, where the imperfect knowledge of the returns on the assets and the uncertainty involved in the behaviour of financial markets may be introduced by means of fuzzy quantities and/or fuzzy constraints.

In this paper, it was proposed a parametric approach that solve a fuzzy quadratic problem. The vagueness can be in the set of constraints and/or costs of the objective function. The set of optimal solutions obtained by one parametric approach to each $\alpha \in [0,1]$ constructs the fuzzy solution by using the Representation Theorem. The portfolio selection presented in this work was solved by three parametric approaches. Firstly, it was formulated with vagueness in the relation of the constraints set that the set of optimal solutions forms a front that represents the possible fuzzy solutions. Then, this problem is formulated with fuzzy costs in the objective function and the set of optimal solutions of parametric approach represents a interval of optimal values of the objective function. Finally, the portfolio problem is formulated with fuzzy coefficients in the constraints set and the set of optimal solutions also constructs an optimal front.

Acknowledgements

The authors want to thank the support provided by the agency FAPESP with project number 2010/51069-2 and

Table 1: Fuzzy portfolio selection problem

	#1	#2	#3	#4		#1	#2	#3	#4
Year	U.S.S.	G.M.	Frstn.	S.S.	Year	U.S.S.	G.M.	Frstn.	S.S.
1945	0.419	0.216	0.229	0.578	1950	0.732	0.305	0.650	0.327
1946	-0.078	-0.272	-0.126	0.289	1951	0.021	0.195	-0.131	0.333
1947	0.169	0.144	0.009	0.184	1952	0.131	0.390	0.175	0.062
1948	-0.035	0.107	0	0.114	1953	0.006	-0.072	-0.084	-0.048
1949	0.133	0.321	0.223	-0.222	1954	0.908	0.715	0.756	0.185

Table 2: Results of the first phase of the parametric approach of 3.2.1.

α	**x**	**Solution**
0.0	[0.1424 ; 0.4028 ; -0.0000 ; 0.4548]	0.0335
0.1	[0.2314 ; 0.3521 ; 0.0000 ; 0.4165]	0.0379
0.2	[0.3205 ; 0.3014 ; -0.0000 ; 0.3782]	0.0431
0.3	[0.4095 ; 0.2507 ; 0.0000 ; 0.3398]	0.0490
0.4	[0.4985 ; 0.2000 ; -0.0000 ; 0.3015]	0.0557
0.5	[0.5875 ; 0.1493 ; 0.0000 ; 0.2632]	0.0631
0.6	[0.6766 ; 0.0986 ; -0.0000 ; 0.2248]	0.0713
0.7	[0.7656 ; 0.0479 ; -0.0000 ; 0.1865]	0.0803
0.8	[0.8535 ; 0.0000 ; 0.0000 ; 0.1465]	0.0900
0.9	[0.9218 ; 0.0000 ; 0.0000 ; 0.0782]	0.1007
1.0	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1124

Table 3: Results of the first phase of the parametric approach of 3.2.2.

α	**x**	**Solution**
0.0	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1012
0.1	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1023
0.2	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1034
0.3	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1046
0.4	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1057
0.5	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1068
0.6	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1079
0.7	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1091
0.8	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1102
0.9	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1113
1.0	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1124

Table 4: Results of the first phase of the parametric approach of 3.2.3.

α	x	**Solution**
0.0	[0.0909 ; 0.4321 ; -0.0000 ; 0.4770]	0.0313
0.1	[0.1773 ; 0.3829 ; 0.0000 ; 0.4398]	0.0352
0.2	[0.2653 ; 0.3328 ; -0.0000 ; 0.4019]	0.0398
0.3	[0.3549 ; 0.2817 ; -0.0000 ; 0.3633]	0.0453
0.4	[0.4463 ; 0.2297 ; 0.0000 ; 0.3240]	0.0517
0.5	[0.6116 ; 0.0000 ; 0.0000 ; 0.3884]	0.0611
0.6	[0.6844 ; -0.0000 ; 0.0000 ; 0.3156]	0.0683
0.7	[0.7586 ; 0.0000 ; 0.0000 ; 0.2414]	0.0770
0.8	[0.8342 ; 0.0000 ; 0.0000 ; 0.1658]	0.0872
0.9	[0.9114 ; 0.0000 ; -0.0000 ; 0.0886]	0.0990
1.0	[0.9901 ; 0.0000 ; -0.0000 ; 0.0099]	0.1124

the Spanish projects TIN2008-06872-C04-04 and P07-TIC-02970.

References

[1] E. AMMAR and H. A. KHALIFA. Fuzzy portfolio optimization a quadratic programming approach. *Chaos, Solitons and Fractals*, 18:1045–1054, 2003.

[2] C. R. BECTOR and S. CHANDRA. *Fuzzy mathematical programming and fuzzy matrix games*, volume 169 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin, 2005.

[3] M. INUIGUCHI and J. RAMIK. Possibilistic linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. *Fuzzy Sets and Systems*, 111(1):3–28, 2000.

[4] F. JIMÉNEZ, J.M. CADENAS, G. SÁNCHEZ, A.F. GÓMEZ-SKARMETA, and J.L. VERDEGAY. Multi-objective evolutionary computation and fuzzy optimization. *International Journal of Approximate Reasoning*, 43(1):59–75, 2006.

[5] G. J. KLIR. Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like. *Fuzzy Sets and Systems*, 24:141–160, 1987.

[6] S. T. LIU. Quadratic programming with fuzzy parameters: A membership function approach. *Chaos, Solitons and Fractals*, 40(1):237–245, 2009.

[7] S. T. LIU. Solving quadratic programming with fuzzy parameters based on extension principle. In *FUZZ-IEEE 2007*, pages 1–5, London, UK, 2007. IEEE International Conference Fuzzy Systems.

[8] T. LEON, V. LIERN, and E. VERCHER. Viability of infeasible portfolio selection problems: A fuzzy approach. *European Journal of Operational Research*, 139:178–189, 2002.

[9] H. M. MARKOWITZ. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[10] H. M. MARKOWITZ. *Portfolio Selection: Efficient Diversification of Investments*. Blackwell Publisher, Massachusetts, second edition, 1991.

[11] C. CRUZ, R. C. SILVA, and J. L. VERDEGAY. Extending and relating different approaches for solving fuzzy quadratic problems. *Fuzzy Optimization and Decision Making*, 10:193–210, 2011.

[12] H. TANAKA, P. GUO, and B. TÜRKSEN. Portfolio selection based on fuzzy probabilities and possibity distributions. *Fuzzy Sets and Systems*, 111:387–397, 2000.

[13] E. VERCHER, J. D. BERMÚDEZ, and J. V. SEGURA. Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets Syst.*, 158(7):769–782, 2007.

[14] J. WATADA. Fuzzy portfolio selection and its applications to decision making. *Tatra Mountains Mathematics Publication*, 13:219–248, 1997.

[15] H. J. ZIMMERMANN. Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems*, 1:45–55, 1978.

[16] H. J. ZIMMERMANN. Fuzzy mathematical programming. *Computers & Operations Research*, 10(4):291–298, 1983.

DECISIÓN MULTICRITERIO BORROSA DE LA DISTRIBUCIÓN POTENCIAL DE ABIES PINSAPO BOISS CON TECNOLOGÍAS GIS

Verónica Bonis Martín , Luis Garmendia Salvador, Alfonso Garmendia Salvador

vbonismartin@gmail.com, lgarmend@fdi.ucm.es, algarsal@upvnet.upv.es

Resumen

El pinsapo (*Abies pinsapo* Boiss.) o abeto español es una de las especies vegetales más singulares en la Península Ibérica catalogada como en peligro de extinción. En el presente trabajo se calcula su distribución potencial con técnicas borrosas en las provincias de Cádiz y Málaga y en particular en sus tres principales localizaciones (Sierra de Grazalema, Sierra de las Nieves y Sierra Bermeja). Para ello se utilizan como indicadores la precipitación anual acumulada, altitud, orientación, precipitación mínima y "Compound Topographic Index" (índice para representar la capacidad del suelo para retener agua). Las técnicas de decisión multicriterio borrosas utilizadas muestran un resultado considerablemente diferente al análisis booleano y estadístico, pero todos señalan una elevada potencialidad de desarrollo en zonas situadas fuera de las principales figuras de protección y una ocupación actual mucho menor que su área potencial.

Palabras clave : *Abies pinsapo*, distribución potencial, decisión multicriterio, lógica borrosa.

1 INTRODUCCIÓN

Debido a los cambios drásticos del clima o al abandono de los usos tradicionales del suelo muchas especies han visto reducida su distribución y su pronóstico de supervivencia a largo plazo es muy bajo. Es importante conocer las áreas de expansión naturales de la especie, para ser capaces de tomar decisiones a cerca de su conservación.

Existen varios métodos para estimar el área potencial de distribución de las especies basados en Sistemas de Información Geográfica (SIG) que permiten la modelización del nicho ecológico utilizando registros de las especies georreferenciados y variables ambientales que producen representaciones formales del espacio ecológico (nicho) sobre el espacio geográfico (distribución). En este caso nos centraremos en las técnicas multicriterio, y especialmente en las técnicas que incluyen la lógica difusa .

Una de las especies más emblemáticas de la península muy afectada por estos cambios es *Abies pinsapo* que ve su distribución restringida a sólo tres zonas del Sur peninsular con apenas 2000 hectáreas de extensión .

Los métodos multicriterio se definen como un conjunto de técnicas orientadas a asistir en los procesos de toma de decisiones [1]. Éstos permiten el análisis de las propiedades entre las alternativas de selección que se proponen, que deben ser completas, operacionales y fáciles de descomponer en partes más pequeñas y no redundantes [2].

Estas técnicas pueden ser divididas en *compensatorias* (aditivas o basadas en la aproximación al punto ideal, como la TOPSIS- Technique for Order Preference by Similarity to Ideal Solution, AIM- Aspiration-level Interactive Method- o MDS- Multi-Dimensional Scaling ; *no compensatorias* o *difusas,* basadas en el concepto de la pertenencia al conjunto borroso.

La decisión multicriterio borrosas es útil para procesar información con zonas de transición gradual, como pueden ser las clasificaciones de elementos lo cual la hace una potente herramienta para la representación geográfica de entidades con límites imprecisos, ejecución y análisis basados en SIG, ya que es capaz de manejar la incertidumbre acerca de los datos geográficos, reglas de decisión o representación de conceptos imprecisos [1].

El artículo está organizado de la siguiente forma; las subsecciones siguientes describen la especie objetivo de estudio y el conjunto de herramientas utilizado para el análisis. La sección 2 explica el proceso de selección de variables y asignación de función de pertenencia difusa. En la sección 3 se desarrolla el método para la obtención de la distribución potencial de pinsapo para pasar en las secciones 4 y 5 a exponer los resultados y principales conclusiones derivadas de los mismos.

1.1 ESPECIE OBJETIVO: ABIES PINSAPO

El pinsapo es uno de los endemismos vegetales más singulares del Península Ibérica. Es considerada una especie relicta, siendo la única manifestación de abetos capaz de sobrevivir en un ambiente mediterráneo. Puede medir hasta 30 metros de altitud y desarrollar troncos de hasta 1 metro de diámetro.

Las condiciones ambientales que exige la especie están muy definidas, el pinsapo es capaz de sobrevivir exclusi-

vamente en regiones de clima mediterráneo húmedo e hiperhúmedo con precipitaciones mayores de 950 mm anuales,llegando a ocupar zonas de hasta 3000 mm anuales como excepcionalmente se registran en la Sierra de Grazalema.

Sin embargo es capaz de soportar los periodos de sequía estival refugiándose en orientaciones de umbría (exposición Norte) soportar los periodos estivales más secos y cálidos en los meses de julio a agosto refugiándose en laderas de elevada pendiente orientadas al Norte, donde los efectos del estiaje se ven amortiguados, siempre en torno a los 1000 y 1800 metros sobre el nivel del mar [3],[4],[5].

Debido a todo lo descrito anteriormente, la distribución actual del *Abies pinsapo* es muy restringida, formando tres masas importantes ubicadas entre las provincias de Cádiz y Málaga. En la provincia de Málaga , dentro de las Sierra de las Nieves y Sierra Bermeja y en la provincia de Cádiz en la Sierra de Grazalema

Hoy en día los pinsapares se encuentran protegidos por la Ley 2/89 de 18 de Julio de 1989 en la que se aprueba el Inventario de Espacios Naturales de Andalucía, creándose las figuras del Parque Natural de Sierra de Grazalema (1984) y Sierra de las Nieves (1989) y el Paraje Natural de Los Reales de Sierra Bermeja (1989).

En el ámbito europeo, los pinsapares han sido incluidos en la Directiva 92/43 de conservación de hábitats naturales y de la fauna y flora silvestres.

Y como especie, el pinsapo se encuentra protegido por el Decreto 104/94 que establece el Catálogo Andaluz de Especies de la Flora Silvestre Amenazadas, catalogada "en peligro de extinción" [6].

Recientemente se aprueba el Plan de Recuperación y Conservación del pinsapo, cuyo objeto es eliminar las amenazas que le afectan y poder alcanzar un grado de conservación adecuado que le permita ser descatalogado de "en peligro de extinción".

Una de las medidas que se proponen en dicho acuerdo es la mejora y mantenimiento del hábitat en zonas de presencia actual de la especie, así como aquellas potenciales de ser ocupadas, apartado sobre el que profundiza el presente trabajo.

Figura 1: El abeto español (Abies pinsapo Boiss.)

1.2 ARCGIS- MÓDULO FUZZY

Es posible realizar el análisis introduciendo la lógica difusa para la asignación del grado de adecuación de cada localización y cálculo superficies con mayor adecuación para la supervivencia del pinsapo a partir de la manipulación de datos en un Sistema de Información Geográfica. En el módulo *Spatyal Analyst* de ArcGIS 10.0 podemos encontrar diferentes herramientas relacionadas con la implementación de la lógica borrosa para la superposición de rásteres, como *Fuzzy Membership* que permite asignar a un ráster una función de pertenencia borrosa que el usuario define mediante el tipo de función y una serie de parámetros de entrada. Ésta herramienta dispone de un conjunto de funciones para realizar la asignación .

Para realizar la combinación usaremos el módulo *Fuzzy Overlay* que permite la conexión de varios criterios mediante distintos operadores de conjunción, disjuncion y agregación borrosos.

2 SELECCIÓN DE VARIABLES

La selección de variables es quizás uno de los puntos más críticos en el estudio. La distribución del pinsapo responde a una serie de variables muy relacionadas entre sí. Por ejemplo, la disponibilidad de agua se ve modificada por factores como la orografía del territorio que introduce variaciones locales o la altitud que hace que al ascender disminuya la presión atmosférica y la temperatura y por consiguiente reduciendo el déficit hídrico.

Basándonos en la bibliografía y características de la especie consideramos una serie de variables :

Climáticas (temperatura del mes más cálido, precipitación anual acumulada)

Topográficas (altitud, Compound Topographic Index y orientación)

El CTI (Compound Topographic Index) es un índice que mide la capacidad del suelo para retener agua, es función de la pendiente y el área subsidiaria de cada punto.

El primer paso es la asignación de la función de pertenencia, que depende del conjunto de datos que tratemos y la función de transformación que mejor represente a cada variable:

Precipitación anual acumulada *(mm)* : función Alta difusa, con parámetros 950 como punto medio y 4 de amplitud (los valores mayores de 950 tienen más posibilidades de pertenecer al conjunto) .

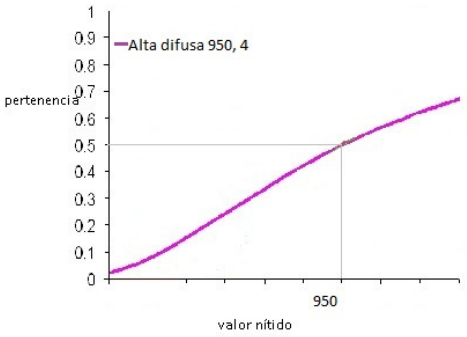


Figura 2: Conjunto borroso 'alta difusa' seleccionado en el módulo Spatyal Analyst de ArcGIS 10.0 para las variable "Precipitación anual acumulada", con parámetros 950 como punto medio y 4 de amplitud "Precipitación del mes más seco

Precipitación del mes más seco *(mm)*: función Alta difusa, parámetros 7 como punto medio y 3 de amplitud.

Figura 3: Conjunto borroso 'alta difusa' para las variable "Precipitación del mes más seco", con parámetros 7 como punto medio y 3 de amplitud.

Altitud *(m)*: los valores entre 1000 y 1800 metros son los más adecuados para la supervivencia de la especie, para modelar este comportamiento usamos la función Gaussiana difusa, que asignará mayores valores de pertenencia a los terrenos con altitudes de en torno 1400 m, la adecuación disminuirá tanto hacia la derecha como hacia la izquierda.

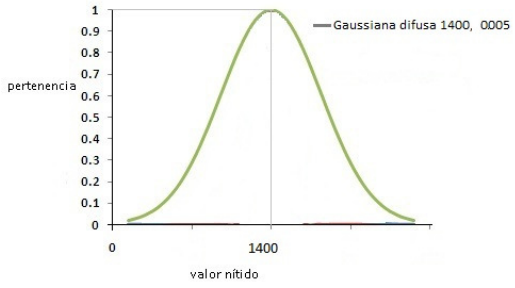


Figura 4: Conjunto borroso de tipo 'gaussiano' seleccionado para las variables "Altitud idonea", con punto medio de 1400 m y 0,0005

Compound Topographic Index *:* Alta difusa para los valores mayores de 7, amplitud 3.

Orientación *(grados):* en el caso de la orientación, tendremos que subdividir la asignación en dos partes, ya que tanto los valores en torno a 11º como los valores en torno a 349º son óptimas para la supervivencia del pinsapo (orientación Norte). Volvemos a usar la función Gaussiana difusa y el operador OR para obtener los conjuntos favorables de ambas.

3 OBTENCIÓN DE LA DISTRIBUCIÓN POTENCIAL CON LÓGICA BORROSA

Para la obtención de la distribución borrosa de abeto, se lleva a cabo una superposición borrosa, la cual cuantifica la relación de pertenencia del fenómeno a conjuntos específicos. El resultado será la posibilidad de que cada celda pertenezca a cada conjunto definido por los criterios:

-Factor pluviométrico

Combinamos los criterios precipitación anual acumulada y precipitación del mes más seco mediante el operador OR cuyo resultado serán los valores de pertenencia máximos para cualquiera de los dos criterios (valores de precipitación anual máximos y valores de precipitación del mes más seco máximos).

-Factor topográfico

Para tener en cuenta los criterios topográficos tendremos que añadir las variables altura y orientación al modelo, el compound topographic index (CTI) en este caso no se tiene en cuenta ya que no aporta mucha información para el resultado final.

-Zonas más aptas para el crecimiento de <u>Abies pinsapo</u>:

Los resultados de las operaciones anteriores se asocian mediante el operador AND que extraerá los valores de mínimo denominador común para la pertenencia del factor pluviométrico y el topográfico (es decir, obtendremos con un valor mayor de posibilidad de pertenencia al conjunto favorable –más cerca de 1- los valores más altos de precipitación anual acumulada y del mes más seco pero que a la vez sean los más altos para la orientación y altura).

La siguiente figura resume el método utilizado, denominado superposición borrosa, para calcular el mapa de adecuación borrosa.

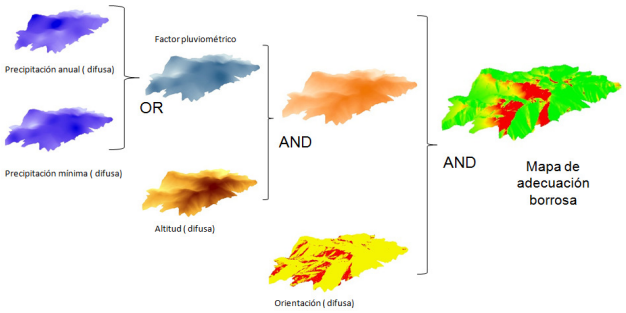


Figura 5: Método de superposición borrosa y mapa de adecuación borrosa

4 RESULTADOS Y COMPARACIÓN CON MÉTODOS BOOLEANOS Y PROBABILÍSTICOS

Los datos disponibles se analizaron mediante diversos métodos cuyos resultados se compararon con los obtenidos con la lógica difusa. Los resultados que se muestran en la tabla 1 son bastante diferentes para cada método, la suma lineal ponderada es la que aporta los valores más altos, mientras que el análisis booleano es el que proporciona unos resultados globales menores.

Tabla 1: Resultados totales para la distribución potencial de *A. pinsapo* por cada uno de los métodos considerados.

	Nivel de adecuación	Nº de píxeles	Superficie (ha)
Análisis booleano	Apta	636702	6367,2
Suma lineal ponderada	Mala	6804360	68043,6
	Media	5030960	50309,6
	Buena	2724109	27241,09
	Excelente	1384200	13842
Lógica y superposición borrosa	Óptima	971149	9711,49

Al utilizar la Suma Lineal Ponderada los resultados para las adecuaciones Buena y Media son mucho menores que el resultado de zonas de potencialidad muy baja para el desarrollo de la especie.

En el desglose de superficie potencial por las localizaciones principales (tabla 2) se repite la misma regla siendo la Suma Lineal Ponderada la que arroja mayores resultados en los tres casos.

Tabla 2: Superficie potencial y superficie ocupada de *A.pinsapo* en los Parques Naturales de Sierra de Grazalema y Sierra de las Nieves y en el Paraje Natural de Sierra Bermeja

Para los tres métodos la zona que presenta mayor potencialidad es el Parque Natural de Sierra de las Nieves, que

	A.B	SLP	LB	Sup. ocupada
P.N. Sª Grazalema	2228,39	5398,1	1862,33	1579,68
P.N. Sª Nieves	1541,23	3297,23	2371,83	2272,46
PJ. N. Sª Bermeja	65,5	247,82	150,84	67,64

es también el lugar con mayores poblaciones de pinsapo en la actualidad con una superficie ocupada de en torno 2000 ha. Por detrás de éste está el Parque Natural de Sierra de Grazalema con alrededor de 1500 ha y por último el Paraje Natural de Sierra Bermeja con menos de 100 ha.

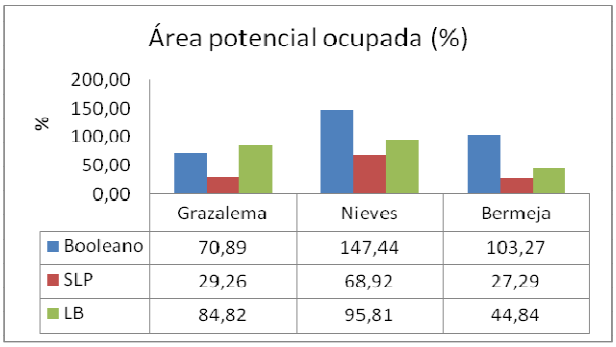


Figura 6: Superficie potencial ocupada de A. pinsapo (%) respecto al área potencial para cada localización especificada para cada métodos de cálculo (Análisis Booleano, Suma Lineal Ponderada-SLP- y Lógica Borrosa –LB-).

En la figura 6 se representa el tanto por ciento de superficie potencial ocupada en cada localización por cada método respecto al área potencial total. Excepto para el análisis booleano (que parece que subestima la superficie) ; en todos los casos la superficie potencial es mayor que la ocupada. L os valores de tanto por ciento de ocupación para el método de SLP en todos los casos es menor ya que éste método arroja valores más altos para la superficie potencial. De nuevo la lógica borrosa da resultados con valores entre los dos anteriores.

5 COMPARACIÓN DE RESULTADOS Y CONCLUSIONES

El apéndice muestra la distribución potencial del pinsapo utilizando métodos de decisión booleanos, estadísticos en la sierra de Grazalema (Cádiz), Sierra de las nieves (Málaga) y sierra Vermeja (Málaga) borrosos.

5.1 CONCLUSIONES ACERCA DEL USO DE DIFERENTES MÉTODOS MULTICRITERIO

Según los resultados el empleo de Análisis Booleano subestima las áreas potenciales de distribución, probablemente porque asume que todos los criterios son exactamente igual de importantes y restrictivos, mientras que la Suma Lineal Ponderada sobreestima los valores finales ya que es posible que alguno de los criterios tenga un peso excesivo sobre el resto, en este caso la altura, minimizando el de otras como la precipitación anual o mínima. El uso de la asignación de funciones de pertenencia y superposición borrosa proporciona unos resultados finales entre los valores de los dos métodos anteriores, asumiendo entonces que es el que produce resultados más realistas.

5.2 CONCLUSIONES ACERCA DE LA DISTRIBUCIÓN DEL *ABIES PINSAPO*

Dadas las características de la especie, el estudio pronosticaba buenos resultados *a priori* con el uso de variables como la altura, orientación o precipitación anual acumulada, hecho que se corrobora al comprobar que las diferentes técnicas han incorporado de manera correcta la presencia actual de la especie.

A pesar de la incertidumbre en la elección de las variables, el empleo de varios métodos para el cálculo de la superficie potencial nos permite la comparación entre ellos pudiendo afirmar con seguridad que el área de distribución potencial del pinsapo, a pesar de sus exigentes necesidades, es mucho mayor que su área de distribución actual. El pinsapo ocupa un relativamente pequeño porcentaje de superficie respecto a su superficie potencial de desarrollo, como ocurre con otras especies de árboles entre los que se encuentra el *Abies alba* [7]. La explicación de este comportamiento limitado en la dispersión a zonas cercanas con elevada potencialidad para el crecimiento de la especie podría responder a muchas y desconocidas causas, como por ejemplo como ya hemos comentado la limitada capacidad dispersiva de la especie.

Por otra parte una gran proporción de la superficie potencial de pinsapo está fuera de las figuras de protección (Parques de Sierra de Grazalema y Sierra de las Nieves y Paraje Natural de Sierra Bermeja). Éstas zonas no protegidas deberían ser tenidas en cuenta para posibles repoblaciones o ampliaciones de las zonas protegidas, así como la consideración de un área de estudio mayor que no esté reducida a las provincias de Cádiz y Málaga y que incluyera todo el extremo sur de la Península y la zona de levante e incluso el Norte de Marruecos donde probablemente las condiciones sean favorables para el desarrollo de la especie.

5 REFERENCIAS

[1] Barredo Cano, J. I., & Gómez Delgado, M. (2005). *Sistemas de Información Geográfica y evaluacion multicriterio en la ordenación del territorio.(2ªEd.)* (p.2312). RA-MA.

[2] Malczewski, J. (1999). *GIS and multicriteria decision analysis* (p. 392). New York: J. Wiley & Sons.

[3] Arista M., F .Herrera, S. T. (1997). *Biología del pinsapo*. *Flora* (p. 162). Sevilla: Consejería de Medio Ambiente. Junta de Andalucía.

[4] Linares, J C ; Carreira, J. A. (2006). El pinsapo , abeto endémico andaluz . O , ¿ Qué hace un tipo como tú en un sitio como éste ? *Ecosistemas*, *3*, 21.

[5] Cerrillo, R. M. N., Fernández, A. L., Oyonarte, P. B., Martínez, C. C., Quintanilla, J. L., Cancio, A. F., Álvarez, J. R. G., et al. (2006). Aproximación a la definición del habitat fisiográfico del Abies pinsapo Boiss . en Andalucía Introducción. *Investigación Agraria: Sistemas y Recursos Forestales. Fuera de serie*, 137-152.

[6] Blanca, G., Cabezudo, B., Hernández-Bermejo, J. E., Molero, J., Muñoz, J., & Valdés, B. (2000). *Libro Rojo de la Flora Silvestre Amenazada de Andalucía.* Consejería de Medio Ambiente Junta de Andalucía.

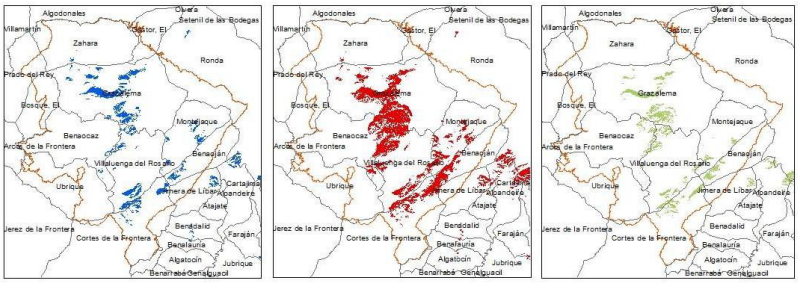
[7] Svenning, J.-C., & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, *7*(7), 565-573.

[8] Zadeh, L. A. (1965). "*Fuzzy sets*". Proceedings of the IEEE (Information and Control), Vol.: 8. 338-353.

6 APÉNDICE

La siguientes figuras muestra la distribución potencial del pinsapo utilizando métodos de decisión booleanos, estadísticos y borrosos en la sierra de Grazalema (Cádiz), Sierra de las nieves (Málaga) y sierra Bermeja (Málaga)

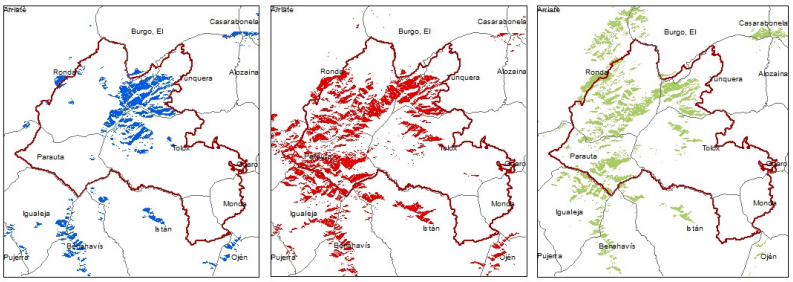
Distribución potencial de "Abies pinsapo" en el Parque Natural de Sierra de Grazalema (Cádiz)



Límite municipal
Parque Natural Sierra de Grazalema
Superficie potencial de "A.pinsapo" (Análisis Booleano)
Superficie potencial de "A.pinsapo" (Suma Lineal Ponderada)
Superficie potencial de "A.pinsapo" (Lógica Borrosa)

	Análisis booleano		Suma Lineal Ponderada		Lógica Borrosa	
Municipio	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)
Cortes de la Frontera	25323	253,23	56307	563,07	25727	257,27
Villaluenga del Rosario	43597	435,97	114356	1143,56	38189	381,89
Benaoján	14258	142,58	36091	360,91	17636	176,36
Benaocaz	7325	73,25	63727	637,27	14805	148,05
Montejaque	20108	201,08	51792	517,92	10076	100,76
Grazalema	11069	110,69	214036	2140,36	79537	795,37
Zahara	1393	13,93	1498	14,98	156	1,56

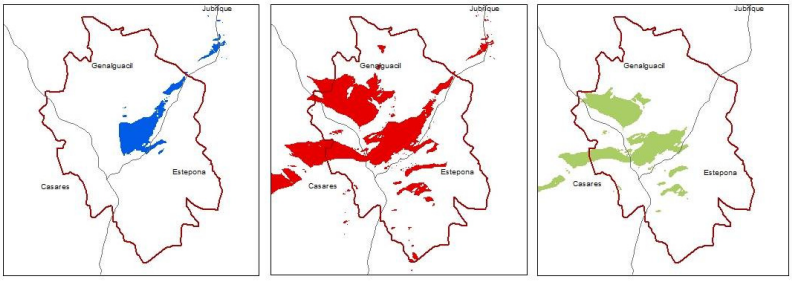
Distribución potencial de "Abies pinsapo" en el Parque Natural de Sierra de las Nieves (Málaga)



Límite municipal
Parque Natural Sierra de las Nieves
Superficie potencial de "A.pinsapo" (Análisis Booleano)
Superficie potencial de "A.pinsapo" (Suma Lineal Ponderada)
Superficie potencial de "A.pinsapo" (Lógica Borrosa)

	Análisis booleano		Suma Lineal Ponderada		Lógica Borrosa	
Municipio	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)
Ronda	38911	389,11	132818	1328,18	116364	1163,64
Yunquera	68787	687,87	63569	635,69	48962	489,62
Tolox	43463	434,63	34085	340,85	16484	164,84
Parauta	758	7,58	96479	964,79	54441	544,41

Distribución potencial de "Abies pinsapo" en el Paraje Natural de Sierra Bermeja (Málaga)



Límite municipal
Paraje Natural de Sierra Bermeja
Superficie potencial de "A.pinsapo" (Análisis Booleano)
Superficie potencial de "A.pinsapo" (Suma Lineal Ponderada)
Superficie potencial de "A.pinsapo" (Lógica Borrosa)

	Análisis booleano		Suma Lineal Ponderada		Lógica Borrosa	
Municipio	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)	Nº pixeles	Superficie (ha)
Genalguacil	6131	61,31	19183	191,83	11538	115,38
Estepona	424	4,24	2964	29,64	1436	14,36

STANDAR GOAL PROGRAMMING WITH FUZZY HIERARCHIES: A SECUENTIAL APPROACH

Mariano Jiménez[1], Mar Arenas-Parra[2], Amelia Bilbao-Terol[2]

[1]*Universidad del País Vasco*, mariano.jimenez@ehu.es
[2]*Universidad de Oviedo*, {mariamar, ameliab}@uniovi.es

Abstract

In standard goal programming (GP) the most widely used way to introduce different importance between goals are the preferential weights or the pre-emptive priorities. However in many cases the decision maker (DM) is not able to establish a precise preference structure. Even in the case of precise weights the solution does not match necessarily the relative weights or in the case of precise pre-emptive priority the result could be very restrictive. In order to overcome the aforementioned drawbacks, in this paper, the normalized unwanted deviations are interpreted in terms of degree of achievement of the targets and, using this approach, fuzzy binary relations are used to model the relative importance between goals. We apply our method to problems where there is a "natural" clustering between goals of the same class (i.e. with a common top-level criterion). We address this situation by solving two phases, in the first one each class is handled separately taking into account the hierarchy between their goals and in the second phase we integrate the results of the first phase and the imprecise hierarchy between the top level criteria. A numerical example illustrates the proposed procedure

Keywords: Goal programming, Fuzzy binary relation, Imprecise hierarchy

1 INTRODUCTION

Let a multi-objective programming (MOP) problem, with k objective functions $f_i(x) = c_i x$, $i = 1,...,k$, be

$$\begin{aligned} &\text{Max} \quad z(x) = \left(f_1(x), f_2(x),..., f_k(x) \right) \\ &\text{s.t.} \quad x \in X \end{aligned} \tag{1}$$

where x and c_i, i=1,…,k are n-vectors of real numbers and X is the feasible set. For the sake of simplicity we suppose that all the objectives should be maximized. However, the proposed method is easily extendable to the case including some minimizing objectives.

In (MOP) it is unlikely that all objectives will simultaneously achieve their optimal value.

Goal programming (GP) is a satisfying technique for the solution of multi-objective problems in which it is admitted that the decision maker (DM) is capable of determining an aspiration level for each objective. Thus the feasible solutions with minimal distance to the aspiration levels are looked for. The objective function is called achievement function and it is a mathematical expression of the unwanted deviation between each objective value and its corresponding aspiration level [5]. One of the most widely used variant of GP is the weighted GP (WGP) in which the achievement function list the unwanted deviation variables, each weighted according to valorization [4]. However the setting of the weights could be a difficult task [1]. Moreover, even in the case in which the DM is able to determine the weights, we have to point out that the relation between the changes in weights and the changes in decision space is not smooth. Especially in the linear case where the solution jumps from one extreme point of the feasible set to another. So small change in weights may lead to a large change in decisions space. Similarly large change in weights can lead to no change in decisions space because they can be insufficient to move to a new extreme point [3].

Due to the previous drawbacks, in order to assure that the preferences of the DM are reflected in the solution, the pre-emptive priorities approach, in which each goal is set in a predefined level, can be used. But it may be difficult to set absolute goal hierarchy. Even in the case of a DM may find absolute goal priority levels easier than precise weights for the goals, we have to point out that pre-emptive priorities implies infinite trade-off between goals placed in different levels, which leaves to a high achievement for the goals placed in the high priority levels and a too very low achievement for those situated in secondary priority levels, which may lead to unsatisfactory results.

In order to overcome the aforementioned drawbacks, Aköz and Petrovic [1] proposed to express the preferences of DM in an imprecise way through fuzzy binary relations. But they only developed their method to fuzzy goal programming models. In this paper we show how the Aköz and Petrovic's method can be adapted in order to solve a standard GP model. In addition we apply this approach, in a sequential manner, to a complex situation

in which the goals are grouped, attending to a common criterion, in several top classes, for example, economic or environmental, and, in its turn, there are internal priorities inside each one of the top classes.

2 THE FUZZY BINARY RELATIONSHIPS

We represent the aspiration level of *i*-th goal by m_i. As it is well known, in order to avoid solutions biased towards the goals with high aspiration levels, the unwanted deviations are normalized in the achievement function dividing them through by a constant pertaining to the corresponding objective. We interpret the concept of deviation in terms of degree of achievement of the aspiration levels; which can be obtained simply by counter-balancing the absolute deviation by the size of the exceeding interval. Without loss of generality assume that the unwanted deviation is the negative one n_i, if *i*-th goal is more important than *j*-th goal, it should be reasonable that, if possible, the optimal solution verifies that the achievement degree of *i*-th goal is greater than the achievement degree of *j*-th goal, that is to say

$$\frac{n_i^*}{m_i - \underline{f}_i} \le \frac{n_j^*}{m_j - \underline{f}_j} \tag{2}$$

where $\underline{f}_i$ is a tolerance threshold established by the DM for *i*-th goal that must verify:

$$0 \le \frac{n_i^*}{m_i - \underline{f}_i} \le 1 \tag{3}$$

In case of the DM does not define a tolerance threshold, $\underline{f}_i$ should be the anti-ideal value of *i*-th goal.

In order to graduate the intensity of the preferences we make more flexible the inequalities (2). Following the ideas of Aköz and Petrovic [1] we propose that the DM expresses their preferences imprecisely by the following fuzzy binary relations: $\tilde{R}_1(i,j)$ "*i*-th goal is slightly more important than *j*-th goal"; $\tilde{R}_2(i,j)$ "*i*-th goal is moderately more important than *j*-th goal" or $\tilde{R}_3(i,j)$ "*i*-th goal is significantly more important than *j*-th goal", which are modeled with the following membership functions that represent the satisfaction degree of the DM preferences,

$$\mu_{\tilde{R}_1(i,j)} = \begin{cases} \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} + 1; & \text{if } -1 \le \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} \le 0 \\ 1; & \text{if } 0 \le \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} \le 1 \end{cases}$$

$$\mu_{\tilde{R}_2(i,j)} = \frac{\dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} + 1}{2}; \text{if } -1 \le \frac{n_j}{m_j - \underline{f}_j} - \frac{n_i}{m_i - \underline{f}_i} \le 1$$

$$\mu_{\tilde{R}_3(i,j)} = \begin{cases} 0; & \text{if } -1 \le \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} \le 0 \\ \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i}; & \text{if } 0 \le \dfrac{n_j}{m_j - \underline{f}_j} - \dfrac{n_i}{m_i - \underline{f}_i} \le 1 \end{cases}$$

Whose graphical representation can be seen in figures 1, 2 and 3.

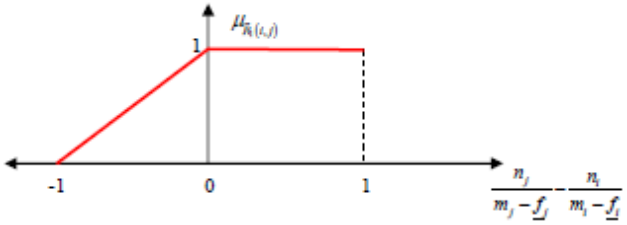


Figure 1. $\tilde{R}_1(i,j)$: *i*-th goal is slightly more important than *j*-th goal.

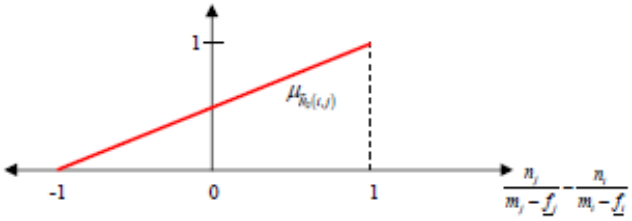


Figure 2. $\tilde{R}_2(i,j)$: *i*-th goal is moderately more important than *j*-th goal.

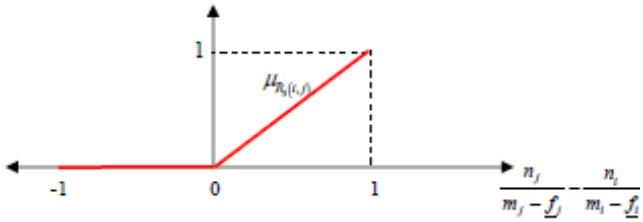


Figure 3. $\tilde{R}_3(i,j)$: *i*-th goal is significantly more important than *j*-th goal.

3 ALGEBRAIC FORMULATION OF A GOAL PROGRAMMING MODEL WITH FUZZY HIERARCHIES

In order to incorporate the preference structure of the DM we built a GP model in which the objectives are the aspiration levels achievement and the imprecise preferences fulfilment. Thus, extending the ideas of Aköz and Petrovic [1] from fuzzy goal programming to standard GP, we propose an achievement function that is a convex combination of the sum of normalized goals achievement and the sum of satisfaction degrees of the fuzzy binary relations. Therefore the following GP with fuzzy hierarchies (GPFH) model is formulated:

$$Max \quad \lambda\sum_{k=1}^{K}\left(1-\frac{n_k}{m_k-\underline{f}_k}\right)+$$

$$+(1-\lambda)\left[\sum_{i=1}^{K}\sum_{j=1}^{K}\left(a_{ij}\mu_{\tilde{R}_1(i,j)}+b_{ij}\mu_{\tilde{R}_2(i,j)}+c_{ij}\mu_{\tilde{R}_3(i,j)}\right)\right]$$

s.t

$$f_k(\boldsymbol{x})+n_k-p_k=m_k \quad k=1,...,K,$$

$$\frac{n_j}{m_j-\underline{f}_j}-\frac{n_i}{m_i-\underline{f}_i}+1\geq\mu_{\tilde{R}_1(i,j)};\ \text{for all}\ (i,j)\ \text{such as}\ a_{ij}=1$$

$$\frac{\frac{n_j}{m_j-\underline{f}_j}-\frac{n_i}{m_i-\underline{f}_i}+1}{2}\geq\mu_{\tilde{R}_2(i,j)};\ \text{for all}\ (i,j)\ \text{such as}\ b_{ij}=1$$

$$\frac{n_j}{m_j-\underline{f}_j}-\frac{n_i}{m_i-\underline{f}_i}\geq\mu_{\tilde{R}_3(i,j)};\ \text{for all}\ (i,j)\ \text{such as}\ c_{ij}=1$$

$$\boldsymbol{x}\in X$$

$$0\leq\mu_{\tilde{R}_l(i,j)}\leq 1, \quad l=1,2,3$$

$$n_k,p_k\geq 0, \qquad n_k\times p_k=0, \quad k=1,\ldots,K$$

Where:

λ ($0 \leq \lambda \leq 1$) is a parameter which allows to obtain different solutions, more or less balanced between the goals achievement ($\lambda = 1$) and the satisfaction of preference relations ($\lambda = 0$). As λ decreases, the relative preference relations have more importance and consequently solutions that satisfy them in a better amount are obtained. a_{ij}, b_{ij} y c_{ij} are binary parameters associated respectively to the relations R_1, R_2 y R_3, which take value equal to 1 if the corresponding relation is defined between the *i*-th and *j*-th goals and 0 otherwise.

This model provides a more flexible decision support that the weighted GP and the lexicographical GP, since the fuzzy preferences that we have defined can be regarded as mutually exclusive priorities relaxed, which at the same time quantify the relative importance that the Decision Maker assigns to the different goals.

4 A SEQUENTIAL GOAL PROGRAMMING MODEL WITH FUZZY HIERARCHIES

Let a GP problem be in which the goals are grouped in several classes, attending to top-level criteria. Assume that the DM in not able to establish a precise hierarchy scheme between them, it is possible to solve the problem by means of a sequence of GPFH models. Each GPFH problem would be modeled attending to its internal hierarchies. In the final GPFH model there appears a problem in which vague hierarchies inter-classes are considered. In order to do more comfortable the decision process we propose the following two phase solution procedure with fuzzy hierarchies:

Phase I. Consider independently each class. Ask the DM for an imprecise hierarchy structure between the goals included in each one. Formulate, and solve, the corresponding GPFH model, one for each class.

Suppose we have two classes of goals of different nature: G_I and G_{II}. Then we have to solve two GPFH problems:

I) A GPFH problem for G_I with achievement function AF_I, and

II) A A GPFH problem for G_{II} with achievement function AF_{II}.

Phase II. Formulate, and solve, a general GPFH in which the first phase solutions are aggregated and the imprecise hierarchy structure between the top-level criteria are considered.

In this model, for the first phase's objectives we consider as aspiration levels and imprecise hierarchy relations those that the Decision Maker had set in this phase. The aspiration levels to the first phase's achievement functions are the optimum values obtained in this phase and they are denoted with *.

Therefore the following model is formulated:

$$Max \quad \lambda\left(1-\frac{n_I}{AF_I^*-\underline{AF}_I}+1-\frac{n_{II}}{AF_{II}^*-\underline{AF}_{II}}\right)+$$

$$+(1-\lambda)\left(\sum_{i,j=I}^{II}a_{ij}\mu_{\tilde{R}_1(i,j)}+b_{ij}\mu_{\tilde{R}_2(i,j)}+c_{ij}\mu_{\tilde{R}_3(i,j)}\right)$$

s.t.

$$\left.\begin{array}{l} AF_I+n_I-p_I=AF_I^* \\ AF_{II}+n_{II}-p_{II}=AF_{II}^* \\ \dfrac{n_j}{AF_j^*-\underline{AF}_j}-\dfrac{n_i}{AF_i^*-\underline{AF}_i}+1\geq\mu_{\tilde{R}_1(i,j)} \quad if\ a_{ij}=1 \\ \dfrac{\dfrac{n_j}{AF_j^*-\underline{AF}_j}-\dfrac{n_i}{AF_i^*-\underline{AF}_i}+1}{2}\geq\mu_{\tilde{R}_2(i,j)} \quad if\ b_{ij}=1 \\ \dfrac{n_j}{AF_j^*-\underline{AF}_j}-\dfrac{n_i}{AF_i^*-\underline{AF}_i}\geq\mu_{\tilde{R}_3(i,j)} \quad if\ c_{ij}=1 \\ 0\leq\mu_{\tilde{R}_l(i,j)}\leq 1, \quad l=1,2,3, \quad i,j\in\{I,II\} \\ n_I,p_I\geq 0, n_{II},p_{II}\geq 0, n_I\times p_I=0, n_{II}\times p_{II}=0 \end{array}\right\}$$

$$
\left.\begin{array}{l}
f_{k_I}(\boldsymbol{x}) + n_{k_I} - p_{k_I} = m_{k_I} \quad k_I = 1,...,K_I \\
\dfrac{n_{j_I}}{\left|m_{j_I} - \underline{f}_{j_I}\right|} - \dfrac{n_{i_I}}{\left|m_{i_I} - \underline{f}_{i_I}\right|} + 1 \geq \mu_{\tilde{R}_1(i_I, j_I)} \qquad \text{if } a_{i_I j_I} = 1 \\
\dfrac{\dfrac{n_{j_I}}{\left|m_{j_I} - \underline{f}_{j_I}\right|} - \dfrac{n_{i_I}}{\left|m_{i_I} - \underline{f}_{i_I}\right|} + 1}{2} \geq \mu_{\tilde{R}_2(i_I, j_I)} \qquad \text{if } b_{i_I, j_I} = 1 \\
\dfrac{n_{jI}}{\left|m_{j_I} - \underline{f}_{j_I}\right|} - \dfrac{n_{iI}}{\left|m_{i_I} - \underline{f}_{i_I}\right|} \geq \mu_{\tilde{R}_3(i_I, j_I)} \qquad \text{if } c_{i_I, j_I} = 1 \\
A_I \boldsymbol{x} \leq b_I, \quad \boldsymbol{x} \geq 0 \\
0 \leq \mu_{\tilde{R}_l(i_I, j_I)} \leq 1, \quad l = 1,2,3 \\
n_{k_I}, p_{k_I} \geq 0, \quad n_{k_I} \times p_{k_I} = 0, \quad k_I = 1,\ldots,K_I
\end{array}\right\} (G_I)
$$

$$
\left.\begin{array}{l}
f_{k_{II}}(\boldsymbol{x}) + n_{k_{II}} - p_{k_{II}} = m_{k_{II}} \quad k_{II} = 1,...,K_{II} \\
\dfrac{n_{j_{II}}}{\left|m_{j_{II}} - \underline{f}_{j_{II}}\right|} - \dfrac{n_{i_{II}}}{\left|m_{i_{II}} - \underline{f}_{i_{II}}\right|} + 1 \geq \mu_{\tilde{R}_1(i_{II}, j_{II})} \qquad \text{if } a_{i_{II} j_{II}} = 1 \\
\dfrac{\dfrac{n_{j_{II}}}{\left|m_{j_{II}} - \underline{f}_{j_{II}}\right|} - \dfrac{n_{i_{II}}}{\left|m_{i_{II}} - \underline{f}_{i_{II}}\right|} + 1}{2} \geq \mu_{\tilde{R}_2(i_{II}, j_{II})} \qquad \text{if } b_{i_{II}, j_{II}} = 1 \\
\dfrac{n_{jII}}{\left|m_{j_{II}} - \underline{f}_{j_{II}}\right|} - \dfrac{n_{iII}}{\left|m_{i_{II}} - \underline{f}_{i_{II}}\right|} \geq \mu_{\tilde{R}_3(i_{II}, j_{II})} \qquad \text{if } c_{i_{II}, j_{II}} = 1 \\
A_{II} \boldsymbol{x} \leq b_{II}, \quad \boldsymbol{x} \geq 0 \\
0 \leq \mu_{\tilde{R}_l(i_{II}, j_{II})} \leq 1, \quad l = 1,2,3 \\
n_{k_{II}}, p_{k_{II}} \geq 0, \quad n_{k_{II}} \times p_{k_{II}} = 0, \quad k_{II} = 1,\ldots,K_{II}
\end{array}\right\} (G_{II})
$$

The following example illustrates the proposed procedure.

5 ILLUSTRATIVE EXAMPLE

The illustration of the proposed sequential GPFH method is done using a modification of representative example considered in the literature [1,2,6]. The following problem with five goals and four crisp system constraints was given:

$$
\begin{aligned}
&\text{Goal 1:} \quad 4x_1 + 2x_2 + 8x_3 + x_4 \leq 35 \\
&\text{Goal 2:} \quad 4x_1 + 7x_2 + 6x_3 + 2x_4 \geq 100 \\
&\text{Goal 3:} \quad x_1 - 6x_2 + 5x_3 + 10x_4 \geq 120 \\
&\text{Goal 4:} \quad 5x_1 + 3x_2 + 2x_4 \geq 70 \\
&\text{Goal 5:} \quad 4x_1 + 4x_2 + 4x_3 \geq 40 \\
&\text{s.t.} \\
&\quad 7x_1 + 5x_2 + 3x_3 + 2x_4 \leq 98 \\
&\quad 7x_1 + x_2 + 2x_3 + 6x_4 \leq 117 \\
&\quad x_1 + x_2 + 2x_3 + 6x_4 \leq 130 \\
&\quad 9x_1 + x_2 + 6x_4 \leq 105 \\
&\quad x_i \geq 0, \quad i = 1,...,4
\end{aligned}
$$

Attending to top-level criteria, we group the goals in two principal top-levels denoted *Group I* (goals 1, 2 and 4) and *Group II* (goals 3 and 5).

The tolerance thresholds for the five goals are:

$$\underline{f}_1 = 55,\ \underline{f}_2 = 40,\ \underline{f}_3 = 70,\ \underline{f}_4 = 30,\ \underline{f}_5 = 10$$

Phase I: we apply the GPHF model to the goals that compose the *Group I* and the *Group II*:
First a GPHF model for *Group I* is formulated**,** by setting:

- goal 1 is *moderately more important than* goal 2, which implies that b_{12}=1, that is to say, the $\tilde{R}_2(1,2)$ relation is considered

and

- goal 2 is *moderately more important than* goal 4: $\tilde{R}_2(2,4)$

the parameter for balancing goals achievement and preference relation satisfaction take the following value: $\lambda_I = 0.5$

Next we formulate a GPHF model for *Group II*, in which we consider two possible scenarios, by setting:

- goal 3 is *moderately more important than* goal 5: $\tilde{R}_2(3,5)$

or

- goal 3 is *significantly more important than* goal 5: $\tilde{R}_3(3,5)$

and $\lambda_{II} = 1$

Phase II: we solve a GPFH with imprecise hierarchy structure between the top-level criteria (*Group I* and *Group II*). We consider several possible scenarios, by setting:

- *Group I* is *moderately more important than Group II*: $\tilde{R}_2(I, II)$

or

- *Group II* is *moderately more important than Group I*: $\tilde{R}_2(II, I)$

or

- *Group II* is *slightly more important than Group I*: $\tilde{R}_1(II, I)$.

Therefore the following GPFH model is formulated:

$$
Max \quad \lambda \sum_{k=I}^{II} \left(1 - \frac{n_k}{AF_k^*}\right) + (1-\lambda)\left[\sum_{i=I}^{II}\sum_{j=I}^{II}\left(a_{ij}\mu_{\tilde{R}_1(i,j)} + b_{ij}\mu_{\tilde{R}_2(i,j)}\right)\right]
$$

s.t

$$
\begin{aligned}
&4x_1 + 2x_2 + 8x_3 + x_4 + n_1 - p_1 = 35 \\
&4x_1 + 7x_2 + 6x_3 + 2x_4 + n_2 - p_2 = 100 \\
&x_1 - 6x_2 + 5x_3 + 10x_4 + n_3 - p_3 = 120 \\
&5x_1 + 3x_2 + 2x_4 + n_4 - p_4 = 70 \\
&4x_1 + 4x_2 + 4x_3 + n_5 - p_5 = 40
\end{aligned}
$$

$$\lambda_I\left(1-\frac{p_1}{20}+1-\frac{n_2}{60}+1-\frac{n_4}{40}\right)+(1-\lambda_I)\left[\mu_{\tilde{R}_2(1,2)}+\mu_{\tilde{R}_2(2,4)}\right]+n_I-p_I=AF_I^*$$

$$\lambda_{II}\left(1-\frac{n_3}{50}+1-\frac{n_5}{30}\right)+(1-\lambda_{II})\left[b_{35}\mu_{\tilde{R}_2(3,5)}+c_{35}\mu_{\tilde{R}_3(3,5)}\right]+n_{II}-p_{II}=AF_{II}^*$$

$$\frac{n_{II}}{AF_{II}^*}-\frac{n_I}{AF_I^*}+1\geq\mu_{\tilde{R}_1(I,II)},\quad if\ a_{I,II}=1$$

$$\frac{n_I}{AF_I^*}-\frac{n_{II}}{AF_{II}^*}+1\geq\mu_{\tilde{R}_1(II,I)},\quad if\ a_{II,I}=1$$

$$\frac{\frac{n_{II}}{AF_{II}^*}-\frac{n_I}{AF_I^*}+1}{2}\geq\mu_{\tilde{R}_2(I,II)},\quad if\ b_{I,II}=1$$

$$\frac{\frac{n_I}{AF_I^*}-\frac{n_{II}}{AF_{II}^*}+1}{2}\geq\mu_{\tilde{R}_2(II,I)},\quad if\ b_{II,I}=1$$

$$\frac{\frac{n_2}{60}-\frac{p_1}{20}+1}{2}\geq\mu_{\tilde{R}_2(1,2)}$$

$$\frac{\frac{n_4}{40}-\frac{n_2}{60}+1}{2}\geq\mu_{\tilde{R}_2(2,4)}$$

$$\frac{\frac{n_5}{30}-\frac{n_3}{50}+1}{2}\geq\mu_{\tilde{R}_2(3,5)},\quad if\ b_{35}=1$$

$$\frac{n_5}{30}-\frac{n_3}{50}\geq\mu_{\tilde{R}_3(3,5)},\quad if\ c_{35}=1$$

$$\boldsymbol{x}\geq 0$$

$$0\leq\mu_{\tilde{R}_l(i,j)}\leq 1,\quad l=1,2,3,\quad i,j\in\{1,\ldots,5,I,II\}$$

$$n_k,p_k\geq 0,\quad n_k\times p_k=0,\quad k=I,II$$

where the aspiration levels for each objective of this problem, AF_I^* and AF_{II}^*, are the solutions obtained for the achievement functions of the two GPFH models solved in Phase I.

If we compare the cases A and B (see Table 1), the sensitivity of the solutions with respect to the parameter λ is shown, we obtain for case A (λ=0.3) better results for *Group I goals*, which are moderately more important, than for *Group II goals*. We have the same conclusion if we compare cases C and D, the two most important goals of the *Group I* are achieved in case C.

If we compare cases A and C, that is to say by changing the R(3,5) from R_3 to R_2, the main goal of *Group II* (G_3) get worse.

If we solve the same problem by permuting the imprecise hierarchies between the top-levels criteria (cases D and E) we can observe in case E that all the goals of e *Group II* are achieved while the main goal of *Group I* does not reach its aspiration level.

If we compare cases E and F, we can observed that if we relax the importance of *Group II* over *Group I* (by changing the R(II,I), from R_2 to R_1) then the value for goal 5 get worse and the most important goal of *Group I*, goal 1, improves.

As it is observed on table 1, the decision-maker may obtain different solutions by expressing relative importance relations among goals linguistically and adjusting only a single parameter, instead of searching for the best weights scheme like in WGP variant

Table 1: Results obtained for various scenarios and several λ values.

	CASE A λ=0,3	**CASE B λ=0,9**	**CASE C λ=0,3**	**CASE D λ=0,5**
R(1,2)	R2	R2	R2	R2
R(2,4)	R2	R2	R2	R2
R(3,5)	R3 (c_{35}=1)	R2 (R3)	R2 (b_{35}=1)	R2
R(I,II)	R2 ($b_{I,II}$=1)	R2	R2	R2
$\mathbf{x_1}$	0,00	0,00	0,09	0,00
$\mathbf{x_2}$	8,46	7,48	10,12	9,55
$\mathbf{x_3}$	0,25	0,47	0,00	0,00
$\mathbf{x_4}$	16,09	16,25	14,42	15,91
$\mathbf{G_1}$	35,00	35,00	35,00	35,00
$\mathbf{G_2}$	92,89	87,72	100,00	98,64
$\mathbf{G_4}$	57,56	54,95	59,62	60,45
$\mathbf{G_3}$	111,39	120,00	83,56	101,82
$\mathbf{G_5}$	34,83	31,82	40,82	38,18
$\mathbf{AF_I}$	1,14	1,20	1,06	1,07
$\mathbf{AF_{II}}$	2,34	2,27	2,73	2,42

Table 1 (cont.)

	CASE E λ=0,5	**CASE F λ=0,5**
R(1,2)	R2	R2
R(2,4)	R2	R2
R(3,5)	R2	R2
R(II,I)	R2($b_{II,I}$=1)	R1($a_{II,I}$=1)
$\mathbf{x_1}$	0,00	0,00
$\mathbf{x_2}$	8,29	7,67
$\mathbf{x_3}$	1,71	0,76
$\mathbf{x_4}$	16,12	16,22
$\mathbf{G_1}$	46,38	37,60
$\mathbf{G_2}$	100,53	90,64
$\mathbf{G_4}$	57,11	55,44
$\mathbf{G_3}$	120,00	120,00
$\mathbf{G_5}$	40,00	33,69
$\mathbf{AF_I}$	1,51	1,27
$\mathbf{AF_{II}}$	2,00	2,21

If we solve the same problem by permuting the imprecise hierarchies between the top-levels criteria (cases D and E) we can observe in case E that all the goals of *Group II* are achieved while the main goal of *Group I* does not reach its aspiration level.

If we compare cases E and F, we can observed that if we relax the importance of *Group II* over *Group I* (by changing the R(II,I), from R_2 to R_1) then the value for goal 5 get worse and the most important goal of *Group I*, goal 1, improves.

As it is observed on table 1, the decision-maker may obtain different solutions by expressing relative importance relations among goals linguistically and adjusting only a single parameter, instead of searching for the best weights scheme like in WGP variant

6 CONCLUSIONS

In this paper a new method for incorporating the preference structure of the DM to a standard GP model is introduced. This method is based in the approach recently proposed by Aköz and Petrovich [1] for fuzzy goal programming problems. It makes more comfortable the definition of preferences by allowing expressing it in a linguistic way. The relative preferences between goals are incorporated directly to the achievement function, which implies that the solution and the preference scheme matches better than in the WGP approach, in which the solution does not necessarily correspond to the relative weights preferences. In addition the rigidity of the lexicographic GP is avoided by relaxing the preemptive priorities through fuzzy relationships. It is shown how the above properties make the proposed procedure useful to handle complex decisional problems with several levels of criteria.

Acknowledgements

The authors wish to gratefully acknowledge financial support from the Spanish Ministry of Education, project ECO2011-26499 and the University of Oviedo, project UNOV-11-MA-11.

References

[1] O. Aköz, D. Petrovic: A fuzzy goal programming method with imprecise goal hierarchy. *European Journal of Operational Research*, vol. 181, pp. 1427-1433, 2007.

[2] L.-H. Chen, F.-C. Tsai: Fuzzy goal programming with different importance and priorities. *European Journal of Operational Research*, vol. 133, pp. 548–556, 2001.

[3] D. Jones: A practical weight sensitivity algorithm for goal and multiple objective programming. *European Journal of Operational Research* 213, pp.238-245, 2011.

[4] O. Ketani, B. Aouni, J.M. Martel: The double role of the weight factor in the goal programming model. *Computers & Operations Research* 31, pp.1833-1845, 2004.

[5] C. Romero: A general structure of achievement function for a goal programming model. *European Journal of Operational Research* 153, pp.675-686, 2004.

[6] R.N. Tiwari, S. Dharmar, J.R. Rao: Fuzzy goal programming – An additive model. *Fuzzy Sets and Systems* 24, pp. 27–34, 1987.

IMPLICACIONES RESIDUADAS EN EL CONJUNTO DE NÚMEROS BORROSOS DISCRETOS

J. Vicente Riera[1], **Joan Torrens**[1]

[1]*Departamento de Matemáticas y Informática, Universidad de las Islas Baleares, Ctra. de Valldemossa, Km.7.5, 07122 Palma de Mallorca, España*, {jvicente.riera,dmijts0}@uib.es

Resumen

En este artículo se estudian implicaciones residuadas en el conjunto de números borrosos cuyo soporte es un conjunto de números naturales consecutivos. Se proporciona una construcción específica de estas implicaciones y se presentan algunos ejemplos, demostrando en particular que dicha construcción generaliza el caso de implicaciones residuadas intervalo valoradas. Finalmente, se investigan las propiedades más usuales obteniendo una estructura de retículo residuado en el conjunto de números borrosos discretos.

Palabras Clave: Implicación borrosa, R-implicación, número borroso discreto, retículo residuado.

1 INTRODUCCIÓN

Las funciones de implicación borrosas juegan un papel fundamental en lógica borrosa, control borroso y razonamiento aproximado. Como generalizaciones de implicaciones en la lógica clásica, no sólo son utilizadas para modelar los condicionales borrosos, sino también en procesos de inferencia a través de las reglas de modus ponens y modus tollens [3]. Es bien conocido que en lógica borrosa habitualmente se utiliza como escala de valoración el intervalo unidad $[0,1]$. Sin embargo, en muchas aplicaciones sólo se usan un número finito de valores e incluso en algunos casos éstos son de tipo cualitativo. Por ello, algunas extensiones de la lógica borrosa y de los conjuntos borrosos han sido desarrolladas para mejorar el modelado de variables lingüísticas y la propia incertidumbre. Entre ellos, los subconjuntos borrosos intervalo valorados [27], conjuntos intuicionistas de Atanassov [2] (equivalentes a los anteriores [12]), subconjuntos borrosos tipo 2 que incluyen los anteriores (ver [19] y sus referencias). Para todas estas extensiones los conectivos lógicos , t-normas, t-conormas, negaciones e implicaciones borrosas han sido estudiadas en detalle. Recientemente, algunas generalizaciones de estas aproximaciones han sido introducidas: *conjuntos borrosos n-dimensionales* (ver [4, 23]) y *hesitant fuzzy sets* (ver [24]).

Por otra parte, muchos autores han estudiado en los últimos años diferentes aproximaciones basadas en operaciones definidas sobre una cadena finita L, usualmente llamadas operaciones discretas. Esta aproximación permite evitar interpretaciones numéricas y consecuentemente, los procesos de fuzzificación y defuzzificación resultan innecesarios. Muchas de estas operaciones han sido estudiadas en este marco, por ejemplo, t-normas y t-conormas discretas fueron caracterizadas en [18], uninormas y nulnormas en [15], medias ponderadas [14] y también implicaciones discretas (ver [16] y [17]). En concreto, se demuestra en [18] que sólo el número de elementos de la cadena finita L es relevante cuando se estudian operaciones monótonas sobre L, y así, la cadena finita utilizada en muchos casos de los mencionados trabajos es la más simple $L = \{0,1,\ldots,n\}$. Como en el caso del intervalo unidad [0,1], los 4 modos usuales de definir funciones de implicación sobre la cadena finita L son los llamados R, S, QL, y D-implicaciones (ver [16] y [17]).

Recientemente, una generalización de esta aproximación ha sido investigada. En ella se propone la posibilidad de extender operaciones monótonas definidas sobre L al conjunto de números borrosos discretos que tienen por soporte un subconjunto de números naturales consecutivos contenidos en L. Concretamente, el concepto de número borroso discreto fue introducido en [25] como un subconjunto borroso sobre $\mathbb{R}$ con soporte discreto y propiedades análogas a las de número borroso. Es bien conocido que las operaciones aritméticas y las operaciones reticulares entre números borrosos están definidas a partir del principio de extensión de Zadeh (ver [13]). Sin embargo, en general, para números borrosos discretos este método falla y diferentes aproximaciones han sido propuestas ([5, 6, 7] y [26]). En particular, fue demostrado [7, 8] que el conjunto, $\mathscr{A}_1^L$, de

números borrosos discretos cuyo soporte es un subconjunto de números naturales consecutivos contenidos en L, es un retículo acotado distributivo. Por tanto es natural estudiar operaciones monótonas definidas sobre $\mathscr{A}_1^L$ equipadas con el orden reticular usual. Un método es el mencionado anteriormente, extendiendo operaciones monótonas definidas sobre L a operaciones monótonas definidas en el conjunto $\mathscr{A}_1^L$. Este método fue utilizado para el caso especial de t-normas y t-conormas discretas en [8], para el caso de medias ponderadas discretas en [10], y para funciones de agregación en general en [22]. En todos estos trabajos, se proponen interesantes aplicaciones de las mencionadas operaciones sobre $\mathscr{A}_1^L$, por ejemplo en problemas de toma de decisiones, evaluación subjetiva, etc.

Además, funciones de implicación definidas sobre L fueron también extendidas a $\mathscr{A}_1^L$ en [9, 21]. En estos trabajos, los autores demuestran que (S,N), QL y D-implicaciones sobre L producen también (S,N), QL y D-implicaciones sobre $\mathscr{A}_1^L$, respectivamente. Sin embargo, las R-implicaciones no fueron incluidas en este estudio a pesar de que este tipo de funciones de implicación son las más extensivamente estudiadas. El interés por las R-implicaciones es debido fundamentalmente a sus *buenas* propiedades basadas principalmente en el principio de residuación o ley de residuación:

$$T(x,y) \leq z \quad \Longleftrightarrow \quad I(x,z) \geq y.$$

Tal propiedad es esencial en la teoría de retículos residuados ya que estas estructuras algebraicas son inherentes a lógicas borrosas y subestructuras lógicas (ver [20]).

En este artículo queremos estudiar R-implicaciones definidas sobre $\mathscr{A}_1^L$. Desafortunadamente, en este caso la extensión de R-implicaciones sobre L produce implicaciones sobre $\mathscr{A}_1^L$ que en general no son R-implicaciones, y por ello es necesaria una construcción específica para este tipo de funciones. Se presentan diferentes ejemplos demostrando que esta construcción, en particular, generaliza el caso de implicaciones residuadas intervalo valoradas (ver [1]). También demostramos que con esta nueva construcción las propiedades más usuales de las R-implicaciones son también ciertas sobre $\mathscr{A}_1^L$, incluyendo la propiedad de residuación, el principio de intercambio y el modus ponens. Finalmente, a partir de dichas propiedades, se obtiene que $\mathscr{A}_1^L$ con la t-norma $\mathscr{T}$ y su implicación residuada $\mathscr{I}_{\mathscr{T}}$ es un retículo residuado acotado.

2 PRELIMINARES

2.1 IMPLICACIONES DISCRETAS

Sea $(\mathscr{P}, \leq)$ un retículo acotado, donde 0 denotará el mínimo y 1 el máximo.

Definición 2.1 *[11] Una función de implicación I sobre $(\mathscr{P}, \leq)$ es un operador binario $I : \mathscr{P} \times \mathscr{P} \to \mathscr{P}$ decreciente en la primera variable, creciente en la segunda y que satisface las siguientes condiciones frontera $I(0,0) = 1, I(1,1) = 1$ y $I(1,0) = 0$.*

Nota 1 *[11] Notar que para cualquier función de implicación I sobre $(\mathscr{P}, \leq)$ se verifica la propiedad*

$$I(0,\alpha) = I(\alpha,1) = 1, \text{ para todo } \alpha \in \mathscr{P}$$

llamada principio de absorción.

Un caso particular bien conocido de conjuntos ordenados acotados son las cadenas finitas $L = \{0,\ldots,n\} \subset \mathbb{N}$. Operadores definidos sobre L se llaman usualmente operaciones discretas y han sido estudiadas por muchos autores (ver [11, 14, 15, 18]). En estos trabajos se utiliza la siguiente condición, generalmente usada en el caso discreto como el equivalente a la propiedad de continuidad en $[0,1]$):

Definición 2.2 *Una función $f : L \to L$ se dice suave si satisface la condición $\mid f(x) - f(x-1) \mid \leq 1$ para todo $x \in L$ con $x \geq 1$.*

Definición 2.3 *Una operación binaria $F : L^2 \to L$ se dice suave cuando cada una de sus secciones, vertical y horizontal ($F(x,.)$ y $F(.,y)$, respectivamente) son suaves.*

T-normas y t-conormas suaves discretas han sido caracterizadas en [18]. En este trabajo, también se demuestra que existe una y solo una negación fuerte sobre L que viene dada por la expresión

$$N(x) = n - x \quad \text{para todo} \quad x \in L. \tag{1}$$

Las cuatro maneras habituales de definir funciones de implicación sobre L han sido investigadas en [16] y [17]. Esto es, R-implicaciones, S-implicaciones, QL-implicaciones y D-implicaciones. En este artículo sólo utilizaremos R-implicaciones sobre L derivadas de t-normas discretas suaves. Para un amplio estudio de sus propiedades ver [16].

2.2 NÚMEROS BORROSOS DISCRETOS

En esta sección, recordamos las principales definiciones y conceptos sobre números borrosos discretos que utilizaremos después. Por subconjunto borroso, entendemos una función $A : \mathbb{R} \to [0,1]$. Para cada subconjunto borroso A, $A^\alpha = \{x \in \mathbb{R} : A(x) \geq \alpha\}$ para cada $\alpha \in (0,1]$ denotará el α-conjunto de nivel (o α-corte). Por $supp(A)$, denotaremos el soporte de A, es decir, el conjunto $\{x \in \mathbb{R} : A(x) > 0\}$. Por A^0, entenderemos la clausura del conjunto $supp(A)$.

Definición 2.4 *[25] Un subconjunto borroso A de $\mathbb{R}$ con función de asociación $A : \mathbb{R} \to [0,1]$ será llamado* número borroso discreto *si su soporte es finito, es decir, existen $x_1,\ldots,x_n \in \mathbb{R}$ con $x_1 < x_2 < \ldots < x_n$ tales que $supp(A) = \{x_1,\ldots,x_n\}$, y existen números naturales s,t con $1 \leq s \leq t \leq n$ tales que:*

1. *$A(x_i)$=1 para cada número natural i con $s \leq i \leq t$ (*núcleo*)*
2. *$A(x_i) \leq A(x_j)$ para cada número natural i, j con $1 \leq i \leq j \leq s$*
3. *$A(x_i) \geq A(x_j)$ para cada número natural i, j con $t \leq i \leq j \leq n$*

Nota 2 *Si el subconjunto borroso A es un número borroso discreto entonces su soporte coincide con su clausura, es decir, $supp(A) = A^0$.*

A partir de ahora, el conjunto de números borrosos discretos, cuyo soporte es un subconjunto de números naturales consecutivos contenido en L, será denotado por $\mathscr{A}_1^L$. Sean pues $A, B \in \mathscr{A}_1^L$ y sean $A^\alpha = [x_1^\alpha, x_p^\alpha]$, $B^\alpha = [y_1^\alpha, y_k^\alpha]$ sus α-conjuntos de nivel respectivamente. Para cada $\alpha \in [0,1]$, consideremos los siguientes conjuntos,

$$\text{mín}(A,B)^\alpha = \{z \in L \mid \text{mín}(x_1^\alpha, y_1^\alpha) \leq z \leq \text{mín}(x_p^\alpha, y_k^\alpha)\} \quad (2)$$

$$\text{máx}(A,B)^\alpha = \{z \in L \mid \text{máx}(x_1^\alpha, y_1^\alpha) \leq z \leq \text{máx}(x_p^\alpha, y_k^\alpha)\} \quad (3)$$

Entonces tenemos la siguiente

Proposición 1 *[5] Existen dos únicos números borrosos discretos, que serán denotados por $MIN(A,B)$ y $MAX(A,B)$, tales que para cada $\alpha \in [0,1]$, sus α-conjuntos de nivel son respectivamente,* $\text{mín}(A,B)^\alpha$ *y* $\text{máx}(A,B)^\alpha$.

El siguiente resultado es cierto en $\mathscr{A}_1^L$, pero no se verifica en general en el conjunto de números borrosos discretos (ver [7]).

Teorema 1 *[7] $(\mathscr{A}_1^L, MIN, MAX)$ es un retículo acotado distributivo donde $1_n \in \mathscr{A}_1^L$ (el único número borroso discreto que tiene por soporte el conjunto unitario $\{n\}$) y $1_0 \in \mathscr{A}_1^L$ (el único número borroso discreto que tiene por soporte el conjunto unitario $\{0\}$) representan el máximo y el mínimo, respectivamente.*

En [8], t-normas y t-conormas suaves sobre L fueron extendidas a t-normas y t-conormas sobre $\mathscr{A}_1^L$.

Teorema 2 *[8] Sea $T(S)$ una t-norma(t-conorma) suave sobre L y sea*

$$\begin{aligned} \mathscr{T}(\mathscr{S}) : \mathscr{A}_1^L \times \mathscr{A}_1^L &\to \mathscr{A}_1^L \\ (A,B) &\longmapsto \mathscr{T}(\mathscr{S})(A,B) \end{aligned}$$

la extensión de la t-norma(t-conorma) $T(S)$ sobre $\mathscr{A}_1^L$, donde $\mathscr{T}(\mathscr{S})(A,B)$ es el número borroso discreto cuyos α-cortes son los conjuntos

$$\{z \in L \mid T(S)(\text{mín}A^\alpha, \text{mín}B^\alpha) \leq z \leq T(S)(\text{máx}A^\alpha, \text{máx}B^\alpha)\}$$

para cada $\alpha \in [0,1]$. Entonces $\mathscr{T}(\mathscr{S})$ es una t-norma(t-conorma) sobre el conjunto acotado $\mathscr{A}_1^L$.

Nota 3 *Como comentamos antes sólo existe una única negación fuerte sobre la cadena finita L dada por la ecuación (1). Además, se demuestra en [9] que esta función de negación sobre L puede ser extendida a $\mathscr{A}_1^L$ vía*

$$\begin{aligned} \mathscr{N} : \mathscr{A}_1^L &\longrightarrow \mathscr{A}_1^L \\ A &\longmapsto \mathscr{N}(A) \end{aligned}$$

donde $\mathscr{N}(A)$ es el único número borroso discreto cuyos α-cortes son los conjuntos

$$\{z \in L \mid N(\text{máx}A^\alpha) \leq z \leq N(\text{mín}A^\alpha)\}$$

para cada $\alpha \in [0,1]$.

Definición 2.5 *Consideremos la función de implicación I sobre la cadena finita L. La operación binaria sobre $\mathscr{A}_1^L$ definida de la siguiente manera*

$$\begin{aligned} \mathscr{I} : \mathscr{A}_1^L \times \mathscr{A}_1^L &\longrightarrow \mathscr{A}_1^L \\ (A,B) &\longmapsto \mathscr{I}(A,B) \end{aligned}$$

será llamada la extensión de la función de implicación discreta *I* a *$\mathscr{A}_1^L$, siendo $\mathscr{I}(A,B)$ el número borroso discreto cuyos α-cortes son los conjuntos*

$$\{z \in L \mid I(\text{máx}A^\alpha, \text{mín}B^\alpha) \leq z \leq I(\text{mín}A^\alpha, \text{máx}B^\alpha)\}$$

para cada $\alpha \in [0,1]$.

Teorema 3 *[21] Sea I una función de implicación sobre L. Entonces* la extensión de la función de implicación discreta *I, $\mathscr{I}$ es una función de implicación sobre $\mathscr{A}_1^L$.*

3 FUNCIONES DE IMPLICACIÓN RESIDUADAS SOBRE $\mathscr{A}_1^L$

En adelante N denotará la única negación fuerte $N(x) = n - x$ sobre L y $\mathscr{N}$ su extensión a $\mathscr{A}_1^L$ (ver nota 3). Similarmente, si T, S y I representan, una t-norma, t-conorma y una implicación sobre L, entonces denotaremos por $\mathscr{T}, \mathscr{S}$ y $\mathscr{I}$ sus extensiones a $\mathscr{A}_1^L$, respectivamente. Recordar que si I es una (S,N)-implicación construida a partir de una t-conorma S y la negación N, entonces su correspondiente extensión $\mathscr{I}$ es la (S,N)-implicación sobre $\mathscr{A}_1^L$ construida a partir de la t-conorma $\mathscr{S}$ y la negación $\mathscr{N}$, y lo mismo ocurrre para QL y D-implicaciones (ver [21]).

Desafortunadamente, esto no es cierto para R-implicaciones. Consideremos una t-norma T sobre L y sea $\mathscr{T}$ su extensión sobre $\mathscr{A}_1^L$. La R-implicación obtenida a partir de $\mathscr{T}$ tiene la siguiente expresión

$$\mathscr{I}_{\mathscr{T}}(A,C) = \sup\{B \in \mathscr{A}_1^L \mid \mathscr{T}(A,B) \preceq C\}, \quad (4)$$

donde $\preceq$ indica el orden en el retículo $\mathscr{A}_1^L$.

Si consideramos I_T la R-implicación sobre L obtenida a partir de T, entonces su extensión sobre $\mathscr{A}_1^L$ no se corresponde con la R-implicación sobre $\mathscr{A}_1^L$ dada por la ecuación (4), como veremos en el siguiente ejemplo.

Ejemplo 1 *Consideremos la cadena finita* $L = \{0,1,2,3,4,5,6,7\}$. *Tomemos la t-norma mínimo sobre L y su implicación residual que viene dada por*

$$I_{\text{mín}}(x,y) = \begin{cases} 7 & si \quad x \leq y \\ \\ y & si \quad x > y \end{cases}$$

Denotemos en este caso por MIN la extensión de la t-norma mínimo, y por $\widetilde{I_{\text{mín}}}$ *la extensión de su implicación residual sobre* $\mathscr{A}_1^L$. *Consideremos también los números borrosos discretos sobre* $\mathscr{A}_1^L$ *dados por*

$$A = \{0.7/2, 1/3, 0.9/4\}$$
$$B = \{0.6/1, 0.6/2, 0.7/3, 1/4, 0.8/5\}.$$

Un sencillo cálculo prueba que $MIN(A,B) = \{0.6/1, 0.7/2, 1/3, 0.9/4\}$. *Ahora, sea* $C = MIN(A,B)$ *y calculemos* $\widetilde{I_{\text{mín}}}(A,C)$. *Obtenemos:*

$$\widetilde{I_{\text{mín}}}(A,C) = \{0.6/1, 0.7/2, 0.9/3, 0.9/4, 0.9/5, 0.9/6, 1/7\}\}.$$

De esta manera tenemos que $MIN(A,B) \preceq C$, *pero* $B \not\preceq \widetilde{I_{\text{mín}}}(A,C)$ *y así* $\widetilde{I_{\text{mín}}}$ *no satisface la ecuación (4).*

Entonces, ¿ cómo podemos obtener R-implicaciones sobre $\mathscr{A}_1^L$? A continuación, resolveremos la cuestión planteada.

Primeramente propondremos la siguiente notación. Para cada número borroso discreto $A \in \mathscr{A}_1^L$ sólo un número finito de niveles serán relevantes. Así, dados $A, B \in \mathscr{A}_1^L$, podemos suponer sin pérdida de generalidad que los valores $0 = \alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_m = 1$ representan la unión de los niveles relevantes comunes de A y B, y supondremos esta condición a partir de ahora.

Proposición 2 *Sea T una t-norma discreta sobre L y* I_T *su implicación residual. Para cada pareja de números borrosos discretos* $A, B \in \mathscr{A}_1^L$, *sean* $\alpha_0 = 0 \leq \alpha_1 \leq \cdots \leq \alpha_m = 1$ *los niveles de* A, B *y sean* $A^{\alpha_j} = [a_1^{\alpha_j}, a_2^{\alpha_j}]$ *y* $B^{\alpha_j} = [b_1^{\alpha_j}, b_2^{\alpha_j}]$ *los* α_j*-conjuntos de nivel de A y B respectivamente, para* $0 \leq j \leq m$. *Entonces, existe un único número borroso discreto* $R \in \mathscr{A}_1^L$ *cuyos* α_j*-conjuntos de nivel estarán expresados por los intervalos de la cadena finita L,* R^{α_j},

$$[\min_{j \leq i \leq m} I_T(a_1^{\alpha_j}, b_1^{\alpha_j}) \wedge \min_{0 \leq i \leq m} I_T(a_2^{\alpha_j}, b_2^{\alpha_j}), \min_{0 \leq i \leq j} I_T(a_2^{\alpha_j}, b_2^{\alpha_j})] \tag{5}$$

para cada $j = 0, \cdots, m$.

Proposición 3 *Sea T una t-norma discreta sobre L y* I_T *su implicación residual. Sean* $A, B \in \mathscr{A}_1^L$ *y* $\alpha_0 = 0 \leq \alpha_1 \leq \cdots \leq \alpha_m = 1$ *los niveles de estos números borrosos discretos. Consideremos la aplicación binaria*

$$\begin{aligned} \mathscr{I}_{\mathscr{T}} : \mathscr{A}_1^L \times \mathscr{A}_1^L &\to \mathscr{A}_1^L \\ (A,B) &\to \mathscr{I}_{\mathscr{T}}(A,B) \end{aligned}$$

donde $\mathscr{I}_{\mathscr{T}}(A,B)$ *es el número borroso discreto cuyos* α_j*-conjuntos de nivel son los intervalos de L construidos de acuerdo a la expresión dada en (5) para cada* $j = 0, \cdots, m$. *Entonces,* $\mathscr{I}_{\mathscr{T}}$ *es una función de implicación sobre el retículo acotado* $(\mathscr{A}_1^L, 1_0, 1_n)$.

Teorema 4 *Sea T una t-norma discreta sobre L y* I_T *su implicación residual. Sea* $\mathscr{T}$ *la extensión de T sobre* $\mathscr{A}_1^L$. *Entonces la implicación residual sobre* $\mathscr{A}_1^L$ *derivada de la t-norma* $\mathscr{T}$ *es la función de implicación* $\mathscr{I}_{\mathscr{T}}$ *dada en la proposición 3.*

Nota 4 *A partir del teorema anterior se deduce que la operación supremo en la ecuación (4) puede ser sustituida por el máximo. Es decir, para cualquier t-norma* $\mathscr{T}$ *sobre* $\mathscr{A}_1^L$ *que es la extensión de la t-norma T definida sobre L, se tiene que*

$$\mathscr{I}_{\mathscr{T}}(A,B) = \max\{C \in \mathscr{A}_1^L \mid \mathscr{T}(A,C) \preceq B\}. \tag{6}$$

para todo $A, B \in \mathscr{A}_1^L$.

Ejemplo 2 *Consideremos la t-norma* $T(x,y) = \min(x,y)$ *sobre L y* $I_{\text{mín}}$ *su implicación residual. Queremos usar la fórmula 5 para calcular* $\mathscr{I}_{\mathscr{T}}$ *en este caso. Sean* $A, B \in \mathscr{A}_1^L$ *y consideremos* $A^{\alpha_j} = [a_1^{\alpha_j}, a_2^{\alpha_j}]$ *y* $B^{\alpha_j} = [b_1^{\alpha_j}, b_2^{\alpha_j}]$ *los* α_j*-cortes de A y B respectivamente, donde* $0 \leq j \leq m$. *Para calcular los* α_j*-cortes de* $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j}$ *procederemos por pasos:*

i. Supongamos que

$$a_1^{\alpha_k} \leq b_1^{\alpha_k} \text{ para todo } j \leq k \leq m$$
$$a_2^{\alpha_k} \leq b_2^{\alpha_k} \text{ para todo } 0 \leq k \leq m$$

En este caso, obtenemos $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j} = [n,n]$.

ii. Supongamos que

$$a_1^{\alpha_k} \leq b_1^{\alpha_k} \text{ para todo } j \leq k \leq m$$
$$a_2^{\alpha_k} > b_2^{\alpha_k} \text{ para algún } 0 \leq k \leq j$$

Denotemos por $k_0 = \min\{0 \leq k \leq j \mid a_2^{\alpha_k} > b_2^{\alpha_k}\}$. *En este caso, obtenemos* $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j} = [b_2^{\alpha_{k_0}}, b_2^{\alpha_{k_0}}]$.

iii. Supongamos que

$$a_1^{\alpha_k} \leq b_1^{\alpha_k} \text{ para todo } j \leq k \leq m$$
$$a_2^{\alpha_k} \leq b_2^{\alpha_k} \text{ para todo } 0 \leq k \leq j$$
$$a_2^{\alpha_k} > b_2^{\alpha_k} \text{ para algún } j < k \leq m$$

Denotemos por $k_0 = \min\{j < k \leq m \mid a_2^{\alpha_k} > b_2^{\alpha_k}\}$. *En este caso, resulta* $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j} = [b_2^{\alpha_{k_0}}, n]$.

iv. Supongamos que $a_1^{\alpha_k} > b_1^{\alpha_k}$ *para algún* $j \leq k \leq m$. *Ahora, denotemos por* $k_0 = \min\{j \leq k \leq m \mid a_1^{\alpha_k} > b_1^{\alpha_k}\}$. *En este caso, consideramos dos subcasos:*

a) Si $a_2^{\alpha_k} \leq b_2^{\alpha_k}$ *para todo* $0 \leq k \leq j$ *entonces* $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j} = [b_2^{\alpha_{k_0}}, n]$.

b) Si $a_2^{\alpha_k} > b_2^{\alpha_k}$ *para algún* $0 \leq k \leq j$ *entonces* $\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha_j} = [b_1^{\alpha_{k_0}}, b_2^{\alpha_{k_1}}]$ *donde* $k_1 = \min\{0 \leq k \leq j \mid a_2^{\alpha_k} > b_2^{\alpha_k}\}$.

Destacar que el paso anterior cubre todas las posibilidades y así el α_j*-corte de* $\mathscr{I}_{\mathscr{T}}(A,B)$ *será siempre uno de los intervalos considerados en los pasos anteriores. En particular, notar que:*

1. *Si* $A \preceq B$*, estamos en las condiciones del primer caso para todo* α_j *con* $0 \leq j \leq m$ *y entonces* $\mathscr{I}_{\mathscr{T}}(A,B) = 1_n$.
2. *Si* $A \succ B$*, estamos en el segundo caso del cuarto caso donde* $k_0 = k_1 = j$ *para todo* α_j *con* $0 \leq j \leq m$*. Esto es,* $\mathscr{I}_{\mathscr{T}}(A,B) = B$.
3. *Si A y B no son comparables, alguno de los pasos anteriores será aplicado para calcular* $\mathscr{I}_{\mathscr{T}}(A,B)$.

Nota 5 *Alcalde et al. [1] proporcionan un método constructivo para obtener operadores de implicación borrosos intervalo valorados a partir de un operador de implicación borroso definido sobre* $[0,1]$*. En particular, este método es usado para obtener implicaciones residuadas intervalo valoradas. Así, si* I_T *denota la implicación residual obtenida a partir de la t-norma T definida sobre* $[0,1]$*, la R-implicación intervalo valorada* $\mathbf{I_T}$ *(que será llamada extensión de la implicación residual* I_T *sobre* $I([0,1]) = \{[a,b], a,b \in [0,1], a \leq b\}$*) tiene la siguiente expresión*

$$\mathbf{I_T}([a,b],[c,d]) = [I_T(a,c) \wedge I_T(b,d), I_T(b,d)] \qquad (7)$$

Notar que, si consideramos una cadena finita L en vez del intervalo unidad, es posible obtener una fórmula análoga como la expresada en (7) en el conjunto de intervalos cerrados de L, $I(L)$*, usando en este caso operadores discretos (t-normas y funciones de implicación definidas sobre L) en vez de operadores definidos en el intervalo unidad. También, deseamos destacar que si consideramos el conjunto*

$$\mathbb{B} = \{A \in \mathscr{A}_1^L \mid supp(A) = core(A)\} \subseteq \mathscr{A}_1^L, \qquad (8)$$

los elementos de $\mathbb{B}$ *son aquellos números borrosos discretos cuyos* α*-cortes son iguales para todo* $\alpha \in [0,1]$*. Así, si* A,B *son números borrosos discretos de* $\mathbb{B}$ *cuyos* α*-cortes están dados por los intervalos* $[a,b]$ *y* $[c,d]$ *respectivamente, entonces un simple calculo muestra que*

$$\mathscr{I}_{\mathscr{T}}(A,B)^{\alpha} = [I(a,c) \wedge I(b,d), I(b,d)]$$

para todo $\alpha \in [0,1]$*. Esto es, la fórmula (7) puede ser vista como un caso particular de la función de implicación residual sobre* $\mathscr{A}_1^L$.

En la siguiente proposición estudiaremos algunas de las propiedades que son comunes a cualquier función de implicación residual sobre $\mathscr{A}_1^L$.

Proposición 4 *Para cada* $A,B,C \in \mathscr{A}_1^L$ *la función de implicación* $\mathscr{I}_{\mathscr{T}}$ *satisface las siguientes propiedades:*

P1. $\mathscr{I}_{\mathscr{T}}(1_n, A) = A$*. (Neutralidad de la verdad).*

P2. $\mathscr{I}_{\mathscr{T}}(A,B) = 1_n$ *si y solo si* $A \preceq B$*. (Propiedad de ordenación).*

P3. $\mathscr{I}_{\mathscr{T}}(A,A) = 1_n$*. (Principio de identidad).*

P4. $\mathscr{T}(A,B) \preceq C$ *si y solo si* $\mathscr{I}_{\mathscr{T}}(A,C) \succeq B$*. (Principio de residuación).*

P5. $\mathscr{T}(A, \mathscr{I}_{\mathscr{T}}(A,B)) \preceq B$*. (Modus ponens).*

P6. $\mathscr{I}_{\mathscr{T}}(A,B) \succeq B$.

P7. $\mathscr{I}_{\mathscr{T}}(A, \mathscr{I}_{\mathscr{T}}(B,C)) = \mathscr{I}_{\mathscr{T}}(\mathscr{T}(A,B),C)$*. (Ley de importación para* $\mathscr{T}$*).*

P8. $\mathscr{I}_{\mathscr{T}}(A, \mathscr{I}_{\mathscr{T}}(B,C)) = \mathscr{I}_{\mathscr{T}}(B, \mathscr{I}_{\mathscr{T}}(A,C))$*. (Principio de intercambio).*

Corolario 1 *Sea T una t-norma discreta sobre L y* $\mathscr{T}$ *su extensión sobre* $\mathscr{A}_1^L$*. Sea* $\mathscr{I}_{\mathscr{T}}$ *la implicación residual sobre* $\mathscr{A}_1^L$ *obtenida a partir de* $\mathscr{T}$*. Entonces* $(\mathscr{A}_1^L, 1_0, 1_n, MIN, MAX, \mathscr{T}, \mathscr{I}_{\mathscr{T}})$ *retículo acotado residuado.*

Ejemplo 3 *Consideremos la cadena finita* $L = \{0,1,2,3,4,5,6,7\}$ *y sean* $A,B \in \mathscr{A}_1^L$ *dados por,*

$$A = \{0.7/2, 1/3, 0.9/4\}$$
$$B = \{0.6/1, 0.6/2, 0.7/3, 1/4, 0.8/5\}$$

- *Sea T la t-norma de Łukasiewicz* $T_{\mathbf{L}}(x,y) = \max\{0, x+y-7\}$ *y* $I_T(x,y) = \min\{7, 7+y-x\}$ *su implicación residual. Aplicando (5), obtenemos*

$$\mathscr{I}_{\mathscr{T}}(A,B) = \{0.6/6, 1/7\}$$

- *Sea T la t-norma mínimo y* $I_{\text{mín}}$ *su implicación residual. Aplicando de nuevo (5) obtenemos*

$$\mathscr{I}_{\mathscr{T}}(A,B) = \{0.6/1, 0.6/2, 0.6/3, 0.6/4, 0.6/5, 0.6/6, 1/7\}$$

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por los proyectos MTM2009-10962 y MTM2009-10320, ambos con ayuda de fondos FEDER.

Referencias

[1] C. Alcalde, A. Burusco and R. Fuentes-González, A constructive method for the definition of interval-valued fuzzy implication operators, *Fuzzy Sets and Systems*, 153 (2005) 211–227.

[2] K.T. Atanassov, Intuitionistic fuzzy sets, *Fuzzy Sets and Systems*, 20 (1986) 87–96.

[3] M. Baczyński and B. Jayaram, *Fuzzy Implications*, Studies in Fuzziness and Soft Computing, vol. 231, Springer, Berlin Heidelberg, 2008.

[4] B. Bedregal, G. Beliakov, H. Bustince, T. Calvo, R. Mesiar and D. Paternain, A class of fuzzy multisets with a fixed number of memberships, *Information Sciences*, submitted.

[5] J. Casasnovas and J.V. Riera, Maximum and Minimum of Discrete Fuzzy numbers. In *Frontiers in Artificial Intelligence and Applications: artificial intelligence research and development*, 163, pp. 273–280, IOS Press, 2007.

[6] J. Casasnovas and J.V. Riera, Discrete fuzzy numbers defined on a subset of natural numbers, *Theoretical Advances and Applications of Fuzzy Logic and Soft Computing*, in the series: Advances in Soft Computing 42, pp. 573-582, Springer, 2007.

[7] J. Casasnovas and J.V. Riera, Lattice properties of discrete fuzzy numbers under extended min and max. In *Proceedings IFSA-EUSFLAT*, pages 647–652, July, Lisbon (Portugal), 2009.

[8] J. Casasnovas and J.V. Riera, Extension of discrete t-norms and t-conorms to discrete fuzzy numbers, *Fuzzy Sets and Systems*, 167(1) (2011) 65–81.

[9] J. Casasnovas and J.V. Riera, S-implications in the set of discrete fuzzy numbers. In *Proceedings IEEE-WCCI 2010*, pages 2741–2747, July, Barcelona (Spain), 2010.

[10] J. Casasnovas and J.V. Riera, Weighted means of subjective evaluations. In R. Seising, V. Sanz (editors), *Soft Computing in Humanities and Social Sciences*. Springer-Verlag, Berlin. Forthcoming.

[11] B. De Baets, Model implicators and their characterization. In N. Sreele, editor, proceedings of the *First ICSC International Symposium on Fuzzy logic*, (ICSC 1995), pages A42–A49, Academic Press, 1995.

[12] G. Deschrijver and E.E. Kerre, On the relation between some extensions of fuzzy set theory, *Fuzzy Sets and Systems*, 133 (2003) 227–235.

[13] G.J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic (Theory and applications)*. Prentice Hall, 1995.

[14] A. Kolesarova, G. Mayor and R. Mesiar, Weighted ordinal means, *Information Sciences*, 177 (2007) 3822-3830.

[15] M. Mas, G. Mayor and J. Torrens, t-operators and uninorms on a finite totally ordered set, *International Journal of Intelligent Systems*, 14 (1999) 909–922.

[16] M. Mas, M. Monserrat and J. Torrens, S-implications and R-implications on a finite chain, *Kybernetica*, 40 (2004) 3–20.

[17] M. Mas, M. Monserrat and J. Torrens, On two types of discretes implications, *International Journal of Approximate Resoning*, 40 (2005) 262–279.

[18] G. Mayor and J. Torrens, Triangular norms on discrete settings. In E.P. Klement and R. Mesiar, editors, *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier, Netherlands, 2005, pp. 189–230.

[19] O. Castillo and P. Mellin, Type-2 Fuzzy Logic Theory and Applications, *Studies in Fuzziness and Soft Computing*, 223, Springer- Verlag, Berlin, 2008.

[20] H. Ono, *Substructural logics and residuated lattices–an introduction*, in V.F. Hendricks, J. Malinowski (Eds.), Trends in Logic, "50 Years of Studia Logica", Vol. 20, Kluwer Academic Publishers, Dordrecht, 2003, pp. 177-212.

[21] J. V. Riera and J. Torrens, Fuzzy implications defined on the set of discrete fuzzy numbers, Proceedings EUSFLAT 2011, pp. 259-266.

[22] J. V. Riera and J. Torrens, Aggregation of subjective evaluations based on discrete fuzzy numbers, *Fuzzy Sets and Systems*, accepted.

[23] Y. Shang, X. Yuan and E.S. Lee, The n-dimensional fuzzy sets and Zadeh fuzzy sets based on the finite valued fuzzy sets, *Computers and Mathematics with Applications*, 60 (2010) 442-463.

[24] V. Torra, Hesitant fuzzy sets, *International Journal of Intelligent Systems*, 25 (2010) 529–539.

[25] W. Voxman, Canonical representations of discrete fuzzy numbers, *Fuzzy Sets and Systems*, 54 (2001) 457–466.

[26] G. Wang, C. Wu and C. Zhao, Representation and Operations of discrete fuzzy numbers. *Southeast Asian Bulletin of Mathematics* 28: (2005) 1003–1010.

[27] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-I, *Information Sciences*, 8(3) (1975) 199–249.

DE MORGAN TRIPLES REVISITED

Francesc Esteva, **Lluís Godo**

IIIA - CSIC, 08913 Bellaterra, Spain, {esteva,godo}@iiia.csic.es

Abstract

In this paper we overview basic known results about the varieties generated by De Morgan triples and about the problem to find equations defining the variety generated by a concrete De Morgan triple. We also provide some alternative proofs and some new results, specially for the case of Łukasiewicz De Morgan triples.

1 Introduction

De Morgan triples are algebraic structures over the real unit interval, defined by the minimum, the maximum, a t-norm, an involutive negation and the t-conorm obtained by duality. Namely $([0,1],\leq,*,n,\oplus)$ is a De Morgan triple if $*$ is a t-norm, n is an involutive (strong) negation and $\oplus$ is the t-conorm defined by duality from $*$ and n (i.e. $x\oplus y = n(n(x)*n(y))$). Actually, any De Morgan triple is characterized by a t-norm $*$ and an involutive negation n, therefore for the sake of a simpler notation, we will denote De Morgan triples as pairs $(*,n)$. Throughout this paper, we will denote the *standard negation* by n_s (i.e. $n_s(x) = 1-x$) and the De Morgan triples defined by the minimum ($*_G$), product ($*_\Pi$) and Łukasiewicz ($*_L$) t-norms and the standard negation n_s will be called *standard* (minimum, product or Łukasiewicz) *De Morgan triples*.

The literature about these structures is quite large and also the use of them in different applications (see for example [20, 21, 1, 2, 17, 18]). De Morgan triples and their isomorphisms were firstly studied by Garcia and Valverde in [11] and posteriorly by Gehrke et al. in [12]. From the logical point of view they have been first studied by Gerkhe et al. in [13]. In that paper the authors consider logics associated to De Morgan triples only having an involutive negation and two conjunctions and disjunctions (corresponding to min, max and the connectives associated to a t-norm and its dual t-conorm) and the constants 0 and 1 (thus implication is not involved). In the setting of the so-called Mathematical Fuzzy Logic (considering also the residuated implication as connective) the most direct precedent is the paper by Monteiro [15] and Sankapanavar [16] about Heyting algebras with an involutive negation, later generalized to residuated lattices with an involutive negation in [4]. But the most relevant papers in mathematical fuzzy logic strictus sensus are [8, 10, 6, 7]. The first paper [8] studies the logic obtained by adding an involutive negation to a strict BL logic, and in particular to Product and Gödel logics. The papers [10, 6, 7] make further studies in the same topic in a more general setting.

This paper is a summary of basic results (with some new results specially for Łukasiewicz triples) about the varieties generated by De Morgan triples and the problem to find equations defining the variety generated by a concrete De Morgan triple. The generated varieties are understood in this paper as subvarieties of the variety of enriched De Morgan algebras of the kind $(A,\wedge,\vee,\odot,\neg,0,1)$, where $(A,\wedge,\vee,\neg,0,1)$ is a De Morgan algebra[1] and $(A,\leq,\odot,1)$ is a commutative ordered monoid[2] such that $x\odot 0 = 0$ for all $x\in A$. Therefore, when talking in this paper about axiomatizations of varieties generated by De Morgan triples, they will

[1] That is, $(A,\wedge,\vee,0,1)$ is a bounded distributive lattice and $\neg : A\to A$ is an involutive negation on A.

[2] Here $\leq$ denotes the order defined by the lattice operations $\wedge$ and $\vee$.

already assume the following set of axioms:

(EDM1) axioms of bounded distributive lattices for $\wedge, \vee, 0, 1$

(EDM2) $\neg(\neg x) = x$

(EDM3) $\neg(x \wedge y) = \neg x \vee \neg y$

(EDM4) axioms of commutative ordered monoids for $\odot, 1$, plus $x \odot 0 = 0$.

2 On De Morgan triples defined over the three basic t-norms

In this section, we recall basic results on De Morgan triples defined by Łukasiewicz, Product and Minimum t-norms together with an involutive negation. Basic results on these three cases are:

- The minimum t-norm case is the most easy one since any pair of De Morgan triples defined by the minimum t-norm and an arbitary involutive negation are trivially isomorphic (see [11]). Thus the set of valid equations satisfied by them are the same, independently of the particular De Morgan triple. Of course these De Morgan triples are finitely axiomatizable by just adding to (EDM1) – (EDM4) the following single additional axiom:

(MIN) $x \odot y = x \wedge y$

- The Łukasiewicz case is well known for the standard De Morgan triple $(*_{\text{Ł}}, n_s)$. The variety generated by this triple is the variety of MV-algebras, finitely axiomatizable by the following equations in a language using a negation operation $\neg$, the disjunction operation $\oplus$ that is dual to $\odot$ wrt $\neg$, and the constant 0 (see for example [3]):

(MV1) $x \oplus (y \oplus z) = (x \oplus y) \oplus z$
(MV2) $x \oplus y = y \oplus x$
(MV3) $x \oplus 0 = x$
(MV4) $\neg(\neg x) = x$
(MV5) $x \oplus \neg 0 = \neg 0$
(MV6) $(\neg(\neg x \oplus y)) \oplus y = (\neg(\neg y \oplus x)) \oplus x$

But if we take the Łukasiewicz t-norm and an involutive negation different from the standard, then equation (MV6) is not valid anymore. Take, for example the involutive negation defined by:

$$n(x) = \begin{cases} 1 - 3x, & \text{if } x \leq \frac{1}{4} \\ \frac{1}{3} - \frac{1}{3}x, & \text{otherwise} \end{cases}$$

Taking $x = \frac{1}{4}, y = \frac{1}{6}$, the left hand of (MV6) is $\frac{1}{6}$ while the right hand of (MV6) is $\frac{1}{4}$.

We devote Section 3 below to the study of the varieties generated by Łukasiewicz De Morgan triples.

- The case of product t-norm has been largely studied. Observe first that in [8], in order to show that Product logic with an involutive negation is not standard complete, it is proved that the equation $(\neg x \odot x) \leq (\neg(\neg x \odot x))^3$ is valid on the standard product De Morgan triple $(*_\Pi, n_s)$ but not in a product De Morgan triple $(*_\Pi, n)$ with $n \neq n_s$. Section 4 is devoted to discuss the varieties generated by the De Morgan triples defined by product t-norm and an involutive negation.

3 On the lattice of subvarieties generated by Łukasiewicz De Morgan triples

Observe first that the identity is the unique automorphism of the algebraic structure defined over $[0,1]$ by Łukasiewicz t-norm $*_{\text{Ł}}$. This easily follows from the fact that any such an automorphism has to be the identity over the rationals and, since $*_{\text{Ł}}$ is continuous, this implies that it must be the identity over the whole real interval $[0,1]$. As a consequence we obtain the following result.

Proposition 3.1. *There are as many non-isomorphic Łukasiewicz De Morgan triples as different involutions can be defined on the real unit interval.*

And for the generated varieties we can prove the following result.

Theorem 3.2. *The lattice of subvarieties generated by Łukasiewicz De Morgan triples has infinite height and infinite width.*

Proof. Take a sequence of involutive negations $\{n_k \mid k \in \mathbb{N}, k \geq 2\}$ such that the fix point of n_k is $\frac{k-1}{k}$ and $n_k(\frac{k-2}{k}) = \frac{(k \cdot k)-1}{k \cdot k}$ and consider the corresponding De Morgan triples $T_k = (*_{\text{Ł}}, n_k)$ for each $k \in \mathbb{N}$. Notice that $(\frac{k-1}{k})^r = 0$ (where $x^k = x *_L \overset{k}{\ldots} *_{\text{Ł}} x$) if and only if $r \geq k$, $x \wedge n_k(x) \leq \frac{k-1}{k}$, and $(n_k((x \vee n_k(x))^2))^r \leq \frac{k-1}{k}$ if and only if $r \geq k-1$. From these results it is obvious that :

- For $m \in \mathbb{N}$, the equation

$$(x \wedge \neg x)^m = 0 \qquad (A_m)$$

is satisfied in any Łukasiewicz De Morgan triple T_k such that $m \geq k$ and it is not satisfied in each T_k such that $m < k$.

- For $r \in \mathbb{N}$, the equation
$$(\neg((x \vee \neg x)^2))^r \leq y \vee \neg y \qquad (B_r)$$
is satisfied in any Łukasiewiz De Morgan triple T_k such that $r \geq k-1$ and it is not satisfied in each T_k such that $1 < r < k-1$.

This proves the infinite width of the lattice of subvareities. In order to prove the infinite height, consider the sets (parametrized by $m \in \mathbb{N}$) of Łukasiewicz De Morgan triples $\mathscr{T}_m = \{T_k \mid 2 \leq k-1 \leq m\}$. Obviously $\mathscr{T}_m \subset \mathscr{T}_{m+1}$ and thus the subset relation is also true for their generated vareities. Moreover it is also obvious that equation (A_m) is satisfied by all De Morgan triples of $\mathscr{T}_m$ but not for all De Morgan triples of $\mathscr{T}_{m+1}$ and so the subset relation between the generated varieties is strict. □

Some interesting open problems are:
1) whether the infinite height and width in Theorem 3.2 is countable or not;
2) to study under which conditions two different Łukasiewicz De Morgan triples generate the same variety;
3) the (finite or not) axiomatization of the different varieties generated by Łukasiewicz De Morgan triples, or equivalently, to find a (finite or not) family of equations characterizing each Łukasiewicz De Morgan triple.

4 On the lattice of subvarieties generated by product De Morgan triples

We consider here product De Morgan triples, i.e. triples $(*,n)$ such that $*$ is isomorphich to the standard product t-norm $*_\Pi$. A first result shows the important role played by the set $\mathscr{M}_\Pi$ of the product De Morgan triples $(*_\Pi,n)$ defined by the standard product t-norm $*_\Pi$ and any involutive negation n with $\frac{1}{2}$ as its fix point.

Proposition 4.1. *Any product De Morgan triple $(*,n)$ is isomorphic to a product De Morgan triple from $\mathscr{M}_\Pi$.*

Proof. Let $*$ be a t-norm isomorphic to the product t-norm $*_\Pi$, and let $f : [0,1]_* \to [0,1]_{*_\Pi}$ be such isomorphism. Then $(*,n)$ is isomorphic to the product De Morgan triple $(*_\Pi,\bar{n})$ where the involutive negation $\bar{n}$ is defined as $\bar{n}(x) = f(n(f^{-1}(x)))$. Denote by s the fix point of $\bar{n}$. On the other hand, it is well known that any automorphism of the algebraic structure defined by the product t-norm is of the form $x \to x^a$ for a fixed positive real a, where x^a denotes usual exponential. The automorphism g defined by taking a such that $s^a = \frac{1}{2}$ does the job, indeed $g \circ f$ gives the desired isomorphism since g transforms the involutive negation $\bar{n}$ into a new involutive negation with fix point $\frac{1}{2}$. □

Thus, in order to study the subvarieties generated by product De Morgan triples one only needs to consider as generators the chains belonging to $\mathscr{M}_\Pi$. The main result in this section is stated in the following theorem.

Theorem 4.2. *The lattice of subvarieties generated by product De Morgan triples has infinite height and infinite (uncountable) width.*

The result is really surprising if we take into account that the lattice of subvarieties of product-algebras contains Boolean algebras as the only proper subvariety and thus the addition of an involutive negation gives rise to a continuum of subvarieties.

The proof of the next theorem is based on results in the paper [13]. Actually, in that paper only the infinite (uncountable) width result is proved, while the infinite height result is proved in [6]. Next we follow [13] for the proof of the uncountable width result, and from there we provide a new proof of the infinite height result.

Theorem 4.3. *Let $(*,n)$ and $(*,\eta)$ be two product De Morgan triples. The variety generated by $(*,n)$ is comparable with the variety generated by $(*,\eta)$ if and only if the triples $(*,n)$ and $(*,\eta)$ are isomorphic.*

By the previous result, we can restrict ourselves to chains of $\mathscr{M}_\Pi$. Obviously two chains of $\mathscr{M}_\Pi$ are isomorphic if and only if they are the same chain and what this theorem says is that two different chains from $\mathscr{M}_\Pi$ generate incomparable subvarieties. To prove this statement we need several lemmas.

Lemma 4.4. *If $(*_\Pi,n)$ belongs to the variety generated by $(*_\Pi,\eta)$, then $n(\frac{1}{2^k}) = \eta(\frac{1}{2^k})$ for all $k \in \mathbb{N}$.*

Proof. Consider for any $k,l,m \in \mathbb{N}$, the equations[3]

[3] Remember that $a \leq b$ is equivalent to $a \wedge b = a$.

$$(n((x\vee n(x))^k))^l \leq (y\vee n(y))^m \qquad (1)$$

and

$$(n((x\wedge n(x))^k))^l \geq (y\wedge n(y))^m. \qquad (2)$$

Equation (1) is valid over $(*_\Pi, n)$ if the inequality holds for any $a,b\in[0,1]$, which is equivalent to

$$\max_{a\in[0,1]}(n(a\vee n(a))^k)^l \leq \min_{b\in[0,1]}(b\vee n(b))^m.$$

It is obvious that these maxima and minima are attained at $\frac{1}{2}$, the fix point of the negations. Thus (1) holds in $(*_\Pi, n)$ if and only if

$$n(\tfrac{1}{2^k}) \leq (\tfrac{1}{2})^{\frac{m}{l}}.$$

An analogous reasoning proves that (2) holds in $(*_\Pi, n)$ if and only if

$$n(\tfrac{1}{2^k}) \geq (\tfrac{1}{2})^{\frac{m}{l}}.$$

If $(*_\Pi, n)$ belongs to the variety genated by $(*_\Pi, \eta)$, the inequalities that hold for $(*_\Pi, n)$ must hold for $(*_\Pi, \eta)$ as well, and being the set $\{(\frac{1}{2})^{\frac{m}{l}} \mid l,m\in\mathbb{N}\}$ dense in the real unit interval, we conclude that for each $k\in\mathbb{N}$, $n(\frac{1}{2^k}) = \eta(\frac{1}{2^k})$. □

Now for any involutive negation n with fix point $\frac{1}{2}$, define the set

$$M(n) = \{(n(\frac{1}{2^k}))^l \mid k,l\in\mathbb{N}\}.$$

Lemma 4.5. *For any involutive negation n with fix point $\frac{1}{2}$, $M(n)$ is dense in the real unit interval.*

Proof. The sequence $\{n(\frac{1}{2^k}) \mid k\in\mathbb{N}\}$ is an increasing sequence with limit 1, and thus for any $\varepsilon > 0$ there is k_0 such that $1 - n(\frac{1}{2^{k_0}}) < \varepsilon$. But $1-b<\varepsilon$ implies $b^m - b^{m+1} = b^m(1-b) < (1-b) < \varepsilon$. Thus it follows that for each real in $[0,1]$ there is an element in the sequence $\{(n(\frac{1}{2^k}))^l \mid k,l\in\mathbb{N}\}$ whose difference from it is at most ε. □

Lemma 4.6. *If $(*_\Pi, n)$ belongs to the variety genated by $(*_\Pi, \eta)$, then n and η coincide on $M(n)$.*

Proof. If $(*_\Pi, n)$ belongs to the variety generated by $(*_\Pi, \eta)$, we know from Lemma 4.4, that for all $k,l\in\mathbb{N}$, $(n(\frac{1}{2^k}))^l = (\eta(\frac{1}{2^k}))^l$ and thus the sets$M(n)$ and $M(\eta)$ coincide. Now consider the inequalities:

$$(n((n((x\wedge n(x))^k))^l))^r \leq (y\vee n(y))^m \qquad (3)$$

and

$$(n((n((x\vee n(x))^k))^l))^r \geq (y\wedge n(y))^m \qquad (4)$$

By an argument similar to that in Lemma 4.4 we obtain that (3) holds in $(*_\Pi, n)$ if, and only if,

$$n((n(\tfrac{1}{2^k}))^l) \leq (\tfrac{1}{2})^{\frac{m}{r}},$$

and that (4) holds in $(*_\Pi, n)$ if, and only if,

$$n((n(\tfrac{1}{2^k}))^l) \geq (\tfrac{1}{2})^{\frac{m}{r}}.$$

But the same conditions are valid for η and thus, reasoning again as in Lemma 4.4, we obtain $n(a) = \eta(a)$, for all $a\in M(n) = M(\eta)$. □

We have thus shown that if $(*_\Pi, n)$ belongs to the variety generated by $(*_\Pi, \eta)$, then n and η agree on a dense set and since involutive negations are continuous functions, they coincide in the full real unit interval. This ends the proof of Theorem 4.3.

The two families of equations used in the proofs above have been considered separately. However the first family is a special case of the second. Namely, taking $k=1$ equation (3) becomes (1), and (4) becomes (2) as an easy computation shows. Thus, in fact, we only have one family of equations used to separate the subvarieties. From Theorem 4.3 it follows that there are as many incomparable subvarieties as involutive negations having $\frac{1}{2}$ as fixed point. Therefore there are uncountable many of the latter. Summarizing, we have the following result.

Corollary 4.7. *The set of subvarieties generated by single product De Morgan triples contains an uncountable number of pair-wise incomparable subvarieties. Furthermore these subvarieties are separable by the following family of equations:*

$$(n((n((x\wedge n(x))^k))^l))^r \leq (y\vee n(y))^m.$$

Next we will prove the infinite height part of Theorem 4.2. Fix $k_0, m_0\in\mathbb{N}$ and an strictly increasing sequence of naturals $\{l_i\}_{i\in\mathbb{N}}$. Then define the sequence $\{T_i\}_{i\in\mathbb{N}}$ of subsets of $\mathscr{M}_\Pi$ as

$$T_i = \{(*_\Pi, n)\in\mathscr{M}_\Pi \mid n(\frac{1}{2^{k_0}}) \leq (\frac{1}{2})^{\frac{m_0}{l_i}}\}.$$

Since $\{\frac{m_0}{l_i}\}_{i\in\mathbb{N}}$ is a decreasing sequence with limit 0, for all $i\in\mathbb{N}$, $T_i\subset T_{i+1}$ and the same inclusions hold true for the varieties generated by these families. Finally an easy observation shows that these inclusions are proper, since the equation (1) for

k_0, m_0 and l_i is valid for T_i but not for T_{i+1}. Thus we have an infinite sequence of strict inclusions of subvarieties, and thus the height of the lattice of subvarieties is, at least, countably infinite.

Remark In [6] and in [14], further insights into the subvarieties of product De Morgan triples and strict De Morgan triples[4] can be found. In order to separate the subvarieties, in [6] Cintula et al. use a different family of equations. For each natural n, they define the equation[5]

$$n((n(x^n))^n) = x \qquad (D_n)$$

and prove, using these equations, that the lattice of subvarieties of product De Morgan triples and strict De Morgan triples contain a sublattice isomorphic to the lattice of natural numbers $(\mathbb{N}, \preceq)$ with the order $\preceq$ defined by: $1 \preceq n$ for all $n \in \mathbb{N}$ and $n \preceq m$ if there is a natural k such that $n^k = m$. So defined, it is clear that $(\mathbb{N}, \preceq)$ has infinite width and infinite height.

One interesting question is to know whether or not there are finitely-many characterizing equations for each product De Morgan triple of $\mathscr{M}_\Pi$, or equivalently, whether there is a finite equational basis for the subvariety generated by each product De Morgan triple of $\mathscr{M}_\Pi$. The question makes sense because in both papers [13, 7] the authors give different sets of separating equations, i.e. defining different subvarieties, but in each case they need an infinite number of equations to axiomatize the subvariety generated by each one of the product De Morgan triple.

5 Final remarks

In the setting of Mathematical Fuzzy logic, the expansions of t-norm based logics with an involutive negation was firstly studied by Esteva et al. in [8] and posteriorly by Flaminio and Marchioni in [10], Cintula et al. in [6] and finally by Haniková and Savický in [14]. The basic difference with the work on De Morgan triples is the use of one further operation, the residuated implication $\Rightarrow$, and of its associated negation defined as $x \Rightarrow 0$. The first paper mostly deals with completeness results for logics of continuous t-norms expanded with an involutive negation, the second one generalizes it to the more general setting of logics of left-continuous t-norms, and the third one deals with some simplifications on the axioms and with some subvarieties defined by additional axioms (among them the ones we have commented in the remark of the preceding section). In the last paper, Haniková and Savický give a set of equations for product logic with an involutive negation that separate the varieties generated by real product logic chains (that cannot be used in our setting because they use the implication and its associated negation that we do not have in the De Morgan triples). On the other hand they characterize the the strict continuous t-norms defining De Morgan triples for which the result of Theorem 4.3 still remains valid.

Still in the setting of residuated t-nom based logics, the problem of whether the logic corresponding to the expansion of the standard product chain with an involutive is finitely axiomatizable has been (partially) adressed in the literature. As far as we know, only the case of the logic of the standard product chain and the standard negation has been proved to be finitely axiomatizable (when using the product implication and thus in a more general settiing than the De Morgan triples). The proof of this result is not trivial and is based on the study of the logic ŁΠ combining both product and Łukasiewicz logics [9]. The original definition of ŁΠ is done in a language with four basic connectives (both Łukasiewicz and product conjunctions and implications), and is complete with respect to the standard ŁΠ-chain, the chain defined over $[0,1]$ by Łukasiewicz and product t-norms and their associated residuated implications. However, a very nice result due to Cintula (see [5, 6, 19] for further details and references) proves that ŁΠ is also complete with respect tto the expansion of the standard product chain with the standard negation. Indeed, he observes that, over this structure, the standard Łukasiewicz conjunction and implication operations are definable as follows:

(C) $x *_Ł y = x *_\Pi n_s(x \Rightarrow_\Pi n_s(y))$

(I) $x \Rightarrow_Ł y = n_s(x *_\Pi n_s(x \Rightarrow_\Pi y))$

and, following this idea, it is proved in [5, 19] that ŁΠ can be defined as an axiomatic extension of the product logic with an involutive negation by either defining $\&_Ł$ as in (C), i.e., as $\varphi \,\&\, \sim(\varphi \to \sim \psi)$ and adding the axiom:

- $\varphi \,\&_Ł\, \psi \to_Ł \psi \,\&_Ł\, \varphi$

or defining $\to_Ł$ as in (I), i.e. as $\sim(\varphi \&\, \sim(\varphi \to \psi))$

[4] That is, those triples defined by continuous t-norms $*$ satisfying $x * x > 0$ for each $x > 0$.

[5] We have translated the equation to our notation.

and adding the axiom:

- $(\varphi \to_{Ł} \psi) \to_{Ł} ((\psi \to_{Ł} \chi) \to_{Ł} (\varphi \to_{Ł\psi}))$

This implies that if one takes an involutive negation different from n_s the resulting operation in (C) is not commutative any longer and, analogously, the resulting function in (I) is not transitive any longer (thus they do not coincide with Łukasiewicz t-norm and its residuum). As a consequence, the logic that is complete with respect to the standard product chain with the standard involutive negation is in fact ŁΠ, hence it is finitely axiomatizable. [6]

Acknowledgements

The authors thank Eurocores LoMoREVI Eurocores Project (FP006/FFI2008-03126-E/FILO), the Spanish projects TASSAT (TIN2010-20967-C04-01) and ARINF (TIN2009-14704-C03-03 with FEDER Funds), the AT project CONSOLIDER CSD 2007- 0022), and the Catalan Government project 2009SGR-1433/34.

References

[1] Claudi Alsina, Enric Trillas, and Llorenç Valverde. Some logical connectives for fuzzy set theory. *Journal of Mathematical Analysis and Applications*, 93(1):15–26, 1983.

[2] Wyllis Bandler and Ladislav J. Kohout. Fuzzy power sets and fuzzy implication operators. *Fuzzy Sets and Systems*, 4:183–190, 1980.

[3] R. Cignoli, I.M.L. D'Ottaviano and D. Mundici. *Algebraic Foundations of Many-valued Reasoning*. Kluwer, Dordrecht, 2000.

[4] Roberto Cignoli and Francesc Esteva. Commutative integral bounded residuated lattices with an added involution. *Annals of Pure and Applied Logic*, 161(2):150–160, 2009.

[5] Petr Cintula. An alternative approach to the ŁΠ logic. *Neural Network World*, 11(6):561–572, 2001.

[6] Petr Cintula, Erich Peter Klement, Radko Mesiar, and Mirko Navara. Residuated logics based on strict t-norms with an involutive negation. *Mathematical Logic Quarterly*, 52(3):269–282, 2006.

[7] Petr Cintula, Erich Peter Klement, Radko Mesiar, and Mirko Navara. Fuzzy logics with an additional involutive negation. *Fuzzy Sets and Systems*, 161(3):390–411, 2010.

[6] This does not imply that the corresponding De Morgan triple be finitely axiomatizable since here we are using the product implication.

[8] Francesc Esteva, Lluís Godo, Petr Hájek, and Mirko Navara. Residuated fuzzy logics with an involutive negation. *Archive for Mathematical Logic*, 39(2):103–124, 2000.

[9] Francesc Esteva, Lluís Godo, and Franco Montagna. The ŁΠ and ŁΠ$\frac{1}{2}$ logics: Two complete fuzzy systems joining Łukasiewicz and product logics. *Archive for Mathematical Logic*, 40(1):39–67, 2001.

[10] Tommaso Flaminio and Enrico Marchioni. T-norm based logics with an independent involutive negation. *Fuzzy Sets and Systems*, 157(4):3125–3144, 2006.

[11] Pere Garcia and Llorenç Valverde. Isomorphisms between De Morgan triplets. *Fuzzy Sets and Systems*, 30: 27–36, 1989.

[12] M. Gehrke, C. Walker, and E. Walker. De Morgan Systems on the unit interval. *Inter. Journal of Intelligent systems*, 11: 751–759, 1996.

[13] M. Gehrke, C. Walker, and E. Walker. Fuzzy logics arising from strict de Morgan systems. In S. Rodabaugh and E. P. Klement, editors, *Topological and Algebraic Structures in Fuzzy Sets: A Handbook of Recent Developments in the Mathematics of Fuzzy Sets*, pages 257–276. Kluwer, Dordrecht, 2003.

[14] Zuzana Haniková and Petr Savický. Distinguishing Standard SBL-Algebras with Involutive Negations by Propositional Formulas. *Mathematical Logical Quaterly*, 54(6): 579–596, 2008.

[15] António A. Monteiro. Sur les algèbres de Heyting symétriques. *Portugaliae Mathematica*, 39(1–4):1–239, 1980.

[16] H. P. Sankappanavar. Heyting algebras with a dual lattice endomorphism. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 33:565–573, 1987.

[17] Enric Trillas. Sobre funciones de negacin en la teora de conjuntos difusos (in span- ish). *Stochastica*, III(1):47–60, 1979.

[18] Enric Trillas, Claudi Alsina, and Llorenç Valverde. Do we need max, min and $1 - j$ in fuzzy set theory? In R.R. Yager, editor, *Fuzzy Set and Possibility Theory: Recent Developments*, pages 275–287. Pergamon Press, New York, 1982.

[19] Thomas Vetterlein. Partial algebras for Łukasiewicz logic and its extensions. *Archive for Mathematical Logic*, 44(7):913–933, 2005.

[20] Lotfi A. Zadeh. Fuzzy Sets. *Information and Control*, 8(3):338–353, 1965.

[21] Lotfi A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(1):28–44, 1973.

Caracterización de las implicaciones f y g-generadas de Yager

Sebastià Massanet, **Joan Torrens**

Departament de Ciències Matemàtiques i Informàtica,
Universitat de les Illes Balears
Campus UIB, Crta. Valldemossa Km. 7.5,
07122 Palma de Mallorca, {s.massanet, dmijts0}@uib.es

Resumen

Las implicaciones de Yager, introducidas en 2004, han resultado interesantes en razonamiento aproximado por sus propiedades y aplicaciones en inferencia. Sin embargo, aunque se han estudiado sus propiedades algebraicas y sus ecuaciones de distributividad, aún no existen caracterizaciones de este tipo de implicaciones. En esta comunicación, presentamos caracterizaciones de las implicaciones de Yager basadas en la ley de importación, una ecuación funcional que ha sido extensamente estudiada recientemente. A partir de estos resultados, también se caracterizan algunos casos particulares como son las implicaciones de Reichenbach, Yager y Goguen.

Palabras Clave: Implicación borrosa, Implicaciones de Yager, f-implicación, g-implicación, ley de importación.

1 INTRODUCCIÓN

En el campo de la lógica borrosa, la caracterización y representación de los conectivos lógicos borrosos, como son los conectivos AND o OR, las negaciones y las implicaciones borrosas, es uno de los puntos de interés de la comunidad investigadora. Esta cuestión ha sido una constante a lo largo del tiempo como se puede apreciar por la gran cantidad de trabajos publicados sobre el tema. Por ejemplo, se puede encontrar la caracterización de diferentes familias de t-normas y t-conormas en [18] (y las referencias interiores), caracterizaciones de uninormas se encuentran en [10, 11, 15, 16, 19, 25], negaciones en [12] y [26] e implicaciones borrosas en [1, 4, 5, 7, 8, 13, 20, 24].

Entre todos estos operadores, las funciones de implicación borrosas se han convertido en claves en el control borroso y el razonamiento aproximado, como también en todos los campos donde se aplican estas teorías. Por lo tanto, las implicaciones borrosas se han convertido en el principal campo de investigación de mucha gente, hecho que corroboran el artículo recopilatorio [23] e incluso la publicación de libros exclusivamente dedicados a este tema (ver [7]).

Debido al gran número de aplicaciones que tienen, existen muchos modelos diferentes de implicaciones borrosas para responder al problema concreto que deben de modelar. En [27], se indica que las caracterizaciones a través de propiedades algebraicas son esenciales para comprender el comportamiento de los diferentes modelos de implicaciones borrosas. Así, por ejemplo, las (S,N) (excepto cuando la negación N no es continua) y las R-implicaciones están caracterizadas completamente (ver [5] y [14] respectivamente), mientras que las QL y las D-operaciones sólo lo están en algunos casos particulares ([20]).

Existe un enfoque distinto para obtener implicaciones borrosas basado en usar directamente funciones generadoras aditivas. De esta forma, las implicaciones de Yager f y g-generadas se pueden interpretar como implicaciones generadas a partir de generadores continuos aditivos de t-normas y t-conormas continuas Arquimedianas, respectivamente. Estas implicaciones son interesantes en razonamiento aproximado por sus propiedades y aplicaciones en inferencia ([28]). También, se han estudiado algunas propiedades algebraicas y ecuaciones de distributividad de estas implicaciones en [6] y [9]. Sin embargo, su caracterización es aún un problema abierto (ver Problema 3.3.1 en [7]).

Las caracterizaciones de las implicaciones de Yager que se presentan en este trabajo están basadas en la ley de importación. Esta ley, que deriva de una tautología en lógica clásica, se convierte en la siguiente ecuación funcional en lógica borrosa:

$$I(T(x,y),z) = I(x,I(y,z)) \quad \text{para todo } x,y,z \in [0,1], \quad (1)$$

donde T es una t-norma e I una implicación borrosa. La ley de importación ha sido estudiada recientemente en [17, 21, 22] para varios tipos de implicaciones. Finalmente, como consecuencia de estas caracterizaciones, se derivarán caracterizaciones particulares de algunas implicaciones importantes, como son las implicaciones de Reichenbach, Yager y Goguen, resolviendo algunos problemas abiertos en la Sección 7.6 en [7].

2 PRELIMINARES

Se supondrá conocida la teoría de t-normas y t-conormas (todas las notaciones y resultados necesarios pueden encontrarse en [2], [14] y [18]).

2.1 NEGACIONES BORROSAS

Definición 1. ([26]; Definición 11.3, [18]) Una función decreciente $N : [0,1] \to [0,1]$ se dice que es una *negación borrosa*, si $N(0) = 1$ y $N(1) = 0$. Una negación borrosa N es

(i) *estricta*, si es continua y estrictamente decreciente.

(ii) *fuerte*, si es una involución, i.e., $N(N(x)) = x$ para todo $x \in [0,1]$.

Ejemplo 1. Algunas negaciones importantes que se usarán en este trabajo son la *negación clásica*, $N_C(x) = 1 - x$, y la *negación borrosa mínima*,

$$N_{D_1}(x) = \begin{cases} 1 & \text{si } x = 0, \\ 0 & \text{si } x \in (0,1]. \end{cases}$$

2.2 IMPLICACIONES BORROSAS

Definición 2. (Definición 1.1.1, [7]) Un operador binario $I : [0,1]^2 \to [0,1]$ se dice que es una *una función de implicación*, o una *implicación*, si satisface:

(I1) $I(x,z) \geq I(y,z)$ cuando $x \leq y$, para todo $z \in [0,1]$.

(I2) $I(x,y) \leq I(x,z)$ cuando $y \leq z$, para todo $x \in [0,1]$.

(I3) $I(0,0) = I(1,1) = 1$ e $I(1,0) = 0$.

Observar que de la definición se deriva que $I(0,x) = 1$ e $I(x,1) = 1$ para todo $x \in [0,1]$ mientras que los valores simétricos $I(x,0)$ e $I(1,x)$ no quedan determinados. Algunas propiedades de las implicaciones borrosas que se usarán posteriormente son las siguientes:

- El *principio de neutralidad por la izquierda*,

$$I(1,y) = y, \quad y \in [0,1]. \qquad (\mathbf{NP})$$

- La *ley de importación* con una t-norma T,

$$I(T(x,y),z) = I(x,I(y,z)), \quad x,y,z \in [0,1]. \qquad (\mathbf{LI})$$

- La *propiedad de orden*,

$$x \leq y \Longleftrightarrow I(x,y) = 1, \quad x,y \in [0,1]. \qquad (\mathbf{OP})$$

Definición 3. (Definición 1.14.15, [7]) Sea I una implicación borrosa. La función N_I definida como $N_I(x) = I(x,0)$ para todo $x \in [0,1]$, se dice que es la *negación natural* asociada a I.

2.2.1 IMPLICACIONES DE YAGER

A continuación, se dan las definiciones de las clases de implicaciones de Yager.

Definición 4. ([28]; Definición 3.1.1, [7]) Sea $f : [0,1] \to [0,\infty]$ una función continua estrictamente decreciente con $f(1) = 0$. La función $I : [0,1]^2 \to [0,1]$ definida por

$$I(x,y) = f^{-1}(x \cdot f(y)), \quad x,y \in [0,1],$$

entendiendo que $0 \cdot \infty = 0$, se dice que es una *implicación f-generada*. La función f es un f-generador de la I generada arriba. En tal caso, se denotará I_f en lugar de I.

Definición 5. ([28]; Definición 3.2.1, [7]) Sea $g : [0,1] \to [0,\infty]$ una función continua y estrictamente creciente con $g(0) = 0$. La función $I : [0,1]^2 \to [0,1]$ definida por

$$I(x,y) = g^{(-1)}\left(\frac{1}{x} \cdot g(y)\right), \quad x,y \in [0,1],$$

entendiendo que $\frac{1}{0} = \infty$ y $\infty \cdot 0 = \infty$, se dice que es una *implicación g-generada*, donde la función $g^{(-1)}$ es la pseudo-inversa de g definida como

$$g^{(-1)}(x) = \begin{cases} g^{-1}(x) & \text{si } x \in [0, g(1)], \\ 1 & \text{si } x \in [g(1), \infty]. \end{cases}$$

La función g es un g-generador de la I generada arriba. En tal caso, se denotará I_g en lugar de I.

Ejemplo 2. Los f-generadores y los g-generadores se pueden ver como generadores aditivos de t-normas y t-conormas continuas Arquimedianas, respectivamente. Los siguientes ejemplos corroboran este hecho. El dibujo de estas implicaciones se puede ver en la Figura 1.

1. Si se considera el f-generador $f(x) = -\ln x$, que es un generador continuo aditivo de la t-norma producto T_P, o el g-generador $g(x) = -\frac{1}{\ln x}$, se obtiene la implicación de Yager I_{YG},

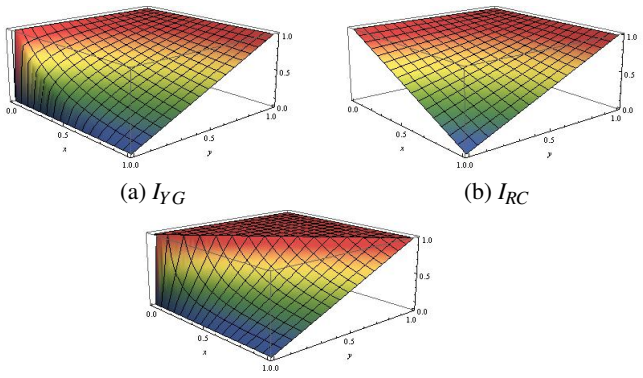
$$I_{YG}(x,y) = \begin{cases} 1 & \text{si } x = 0 \text{ e } y = 0, \\ y^x & \text{si } x > 0 \text{ o } y > 0. \end{cases}$$

2. Con el f-generador $f(x) = 1 - x$, que es el generador continuo aditivo de la t-norma de Łukasiewicz, se obtiene la implicación de Reichenbach,

$$I_{RC}(x,y) = 1 - x + xy.$$

3. Considerando el g-generador $g(x) = x$, que es el generador continuo aditivo de la t-conorma de Łukasiewicz S_{LK}, se obtiene la implicación de Goguen I_{GG}, definida por

$$I_{GG}(x,y) = \begin{cases} 1 & \text{si } x \leq y, \\ \frac{y}{x} & \text{si } x > y. \end{cases}$$



(a) I_{YG} (b) I_{RC}

(c) I_{GG}

Figura 1: Algunas implicaciones de Yager.

3 CARACTERIZACIONES DE LAS IMPLICACIONES f-GENERADAS

El objetivo en esta sección es presentar las caracterizaciones de las implicaciones f-generadas de Yager. Se dará un teorema de caracterización para cada clase de implicaciones f-generadas de Yager dependiendo del valor del f-generador en $x = 0$.

Las implicaciones f-generadas son implicaciones obtenidas mediante el uso de generadores aditivos de t-normas continuas Arquimedianas. Existen dos subclases de implicaciones f-generadas si $f(0) = \infty$ o $f(0) < \infty$. En primer lugar, se obtiene la caracterización de las implicaciones f-generadas que satisfacen que $f(0) < \infty$.

Teorema 1. *Sea $I : [0,1]^2 \to [0,1]$ una función binaria. Entonces las siguientes afirmaciones son equivalentes:*

(i) I es una implicación f-generada con $f(0) < \infty$.

(ii) I satisface (LI) con T_P y N_I es una negación estricta.

Además, en este caso el f-generador es único excepto por una constante multiplicativa positiva y viene dado por $f(x) = N_I^{-1}(x)$.

El caso de las implicaciones f-generadas con $f(0) = \infty$ es bastante más complejo y se resume en el siguiente resultado. Se denotan por $h_k : [0,1] \to [0,1]$ las secciones horizontales de I, dadas por $h_k(x) = I(x,k)$, y por $v_k : [0,1] \to [0,1]$ las secciones verticales de I, dadas por $v_k(x) = I(k,x)$, para todo $x \in [0,1]$ y para todo $k \in [0,1]$.

Teorema 2. *Sea $I : [0,1]^2 \to [0,1]$ una función binaria. Entonces las siguientes afirmaciones son equivalentes:*

(i) I es una implicación f-generada con $f(0) = \infty$.

(ii) I satisface (LI) con T_P, I es continua excepto en $(0,0)$ e $I(x,y) = 1 \Leftrightarrow x = 0$ o $y = 1$.

(iii) I satisface (LI) con T_P, $N_I = N_{D_1}$, I es continua excepto en $(0,0)$ y existe $k_0 \in (0,1)$ tal que

- *h_k es estrictamente decreciente con $h_k(0) = 1$ y $h_k(1) = k$ para todo $k \in (0,k_0]$,*
- *v_k es estrictamente creciente en el intervalo $[0,k_0]$ para todo $k \in (0,1)$.*

(iv) I satisface (LI) con T_P, $N_I = N_{D_1}$ y existe $k \in (0,1)$ tal que

- *h_k es continua y estrictamente decreciente con $h_k(0) = 1$ y $h_k(1) = k$,*
- *$h_k^{-1}(k) : (0,k) \to [0,1]$ que asigna $h_y^{-1}(k)$ a cada $y \in (0,k)$ es una función bien definida, continua y estrictamente creciente tal que $\lim_{y \to 0^+} h_y^{-1}(k) = 0$.*

Además, en este caso el f-generador, que es único excepto por una constante multiplicativa positiva, viene dado por

$$f(x) = \begin{cases} h_k^{-1}(x) & \text{si } k \leq x \leq 1, \\ \frac{1}{h_x^{-1}(k)} & \text{si } 0 < x < k, \\ \infty & \text{si } x = 0, \end{cases} \tag{2}$$

donde k es un valor fijo cualquiera $0 < k < 1$.

En la Tabla 1 se dan ejemplos de funciones $I : [0,1]^2 \to [0,1]$ mostrando que las propiedades en el Teorema 2-(ii) son mútuamente independientes.

En este punto, queremos dar una caracterización para algunas implicaciones bien conocidas e importantes como casos particulares de las caracterizaciones de las implicaciones f-generadas de Yager dadas en esta sección. Algunos de los resultados que se muestran mejoran los existentes ya que requieren de un menor número de condiciones que además son más sencillas, mientras que la caracterización de la implicación de Reichenbach responde al problema abierto puesto en [7], Sección 7.6.

3.1 IMPLICACIÓN DE REICHENBACH

La implicación de Reichenbach es una implicación f-generada con $f(0) < \infty$ generada por el f-generador $f(x) = 1 - x$. Esta implicación es también una S-implicación generada por la t-conorma suma probabilística S_P y la negación clásica N_C. Sin embargo, no se conoce aún caracterización

Tabla 1: La independencia mútua de las propiedades del Teorema 2-(ii)

Función I	**(LI) con** T_P	$I(x,y)=1 \Leftrightarrow x=0$ **o** $y=1$	I **cont.**$\backslash\{(0,0)\}$
$I=T_P$	✓	X	X
$I=I_{KD}$	X	✓	X
$I(x,y)=\begin{cases}\frac{xy}{x^2+y^2} & \text{si } x,y>0,\\ 0 & \text{si } x=y=0.\end{cases}$	X	X	✓
$I=I_{RC}$	✓	✓	X
$I=I_{GG}$	✓	X	✓
$I(x,y)=\begin{cases}y^{x(1-y)} & \text{si } x>0 \text{ o } y>0,\\ 1 & \text{si } x=y=0.\end{cases}$	X	✓	✓

para I_{RC} (ver Sección 7.6 en [7]). Ahora, usando el Teorema 1, se puede derivar el siguiente resultado.

Teorema 3. *Dada una función* $I:[0,1]^2\to[0,1]$ *las siguientes afirmaciones son equivalentes:*

(i) *I es la implicación de Reichenbach* I_{RC}*, esto es,* $I(x,y)=I_{RC}(x,y)=1-x+xy$ *para todo* $x,y\in[0,1]$.

(ii) *I satisface (LI) con* T_P *y* $N_I(x)=1-x$ *para todo* $x\in[0,1]$.

3.2 IMPLICACIÓN DE YAGER

La implicación de Yager es una implicación f-generada con $f(0)=\infty$ con f-generador $f(x)=-\ln x$. Como las implicaciones f-generadas con $f(0)=\infty$ son, de hecho, también implicaciones g-generadas con $g(1)=\infty$, la implicación de Yager es una implicación g-generada con $g(1)=\infty$ con g-generador $g(x)=\frac{-1}{\ln x}$. Esta implicación no es ni una (S,N)-implicación ni una R-implicación.

La caracterización actual de esta implicación se basa también en la ley de importación y en una ecuación distributiva. El resultado, un caso particular del Corolario 2 en [3] o del Teorema 7.5.6 en [7], es el siguiente.

Teorema 4. *Dada una función* $I:[0,1]^2\to[0,1]$ *las siguientes afirmaciones son equivalentes:*

(i) *I es la implicación de Yager,* I_{YG}.

(ii) *I es una implicación borrosa, I es continua excepto en el punto* $(0,0)$*, satisface (LI) con* T_P *y se cumplen las siguientes ecuaciones*

$$\begin{aligned} I(x,yz) &= I(x,y)\cdot I(x,z)\\ I(0.5,y^2) &= y \end{aligned}$$

para todo $x,y,z\in[0,1]$.

Observar que se parte de una función binaria I que necesita ser una implicación borrosa. Por tanto, las condiciones (I1), (I2) e (I3) estan implícitas en la caracterización. En cambio, la caracterización que se presenta no necesita de estos requisitos y se basa en el Teorema 2.

Teorema 5. *Dada una función* $I:[0,1]^2\to[0,1]$ *las siguientes afirmaciones son equivalentes:*

(i) *I es la implicación de Yager,* I_{YG}.

(ii) *I satisface (LI) con* T_P*,* $N_I=N_{D_1}$ *y existe* $k\in(0,1)$ *tal que* $h_k(x)=k^x$ *para todo* $x\in[0,1]$ *e*

$$I\left(\frac{\ln k}{\ln y},y\right)=k, \quad \text{para todo } y\in(0,k). \tag{3}$$

4 CARACTERIZACIONES DE LAS IMPLICACIONES g-GENERADAS

A continuación, el objetivo es caracterizar las implicaciones g-generadas. Estas implicaciones son aquellas obtenidas mediante el uso de generadores aditivos de t-conormas continuas Arquimedianas. También existen dos subclases de implicaciones g-generadas si $g(1)=\infty$ o $g(1)<\infty$. El caso de las implicaciones g-generadas con $g(1)=\infty$ es trivial porque estas implicaciones son de hecho las mismas que las implicaciones f-generadas con $f(0)=\infty$ (ver Proposición 4.4.1 en [7]). Dado un f-generador con $f(0)=\infty$, la siguiente función

$$g(x)=\frac{1}{f(x)}, \quad x\in[0,1],$$

es un g-generador con $g(1)=\infty$. Por lo tanto, el teorema de caracterización se obtiene reescribiendo adecuadamente el Teorema 2 teniendo en cuenta esta relación.

Finalmente, el siguiente teorema caracteriza las implicaciones g-generadas con $g(1)<\infty$.

Teorema 6. *Sea* $I:[0,1]^2\to[0,1]$ *una función binaria. Entonces las siguientes afirmaciones son equivalentes:*

(i) *I es una implicación g-generada con* $g(1)<\infty$.

(ii) I satisface (LI) con T_P y existe una función continua y estrictamente creciente $t : [0,1] \to [0,1]$ con $t(0) = 0$ y $t(1) = 1$ tal que $I(x,y) = 1 \Leftrightarrow y \geq t(x)$.

Además, en este caso el g-generador es único excepto por una constante multiplicativa positiva y viene dado por $g(x) = t^{-1}(x)$.

Una vez caracterizadas todas las implicaciones g-generadas de Yager, se pueden obtener caracterizaciones concretas de algunas implicaciones g-generadas conocidas como casos particulares.

4.1 IMPLICACIÓN DE GOGUEN

La implicación de Goguen es una implicación g-generada con $g(1) < \infty$ con g-generador $g(x) = x$. Es también una R-implicación obtenida a partir de la t-norma producto, pero no es una (S,N)-implicación. De hecho, es la única implicación que pertenece a la intersección de las implicaciones g-generadas con las R-implicaciones obtenidas a partir de una t-norma continua por la izquierda. No existe una caracterización concreta de esta implicación. Sin embargo, hay una caracterización de todas las φ-conjugadas de la implicación de Goguen. La caracterización actual se obtiene en [4] a través de las propiedades de las R-implicaciones.

Teorema 7. (Teorema 19, [4]) *Dada una función $I : [0,1]^2 \to [0,1]$ las siguientes afirmaciones son equivalentes:*

(i) I es una φ-conjugada de la implicación de Goguen I_{GG}, i.e., existe un automorfismo φ, que está únicamente determinado excepto por un exponente constante positivo, tal que I tiene la forma

$$I(x,y) = (I_{GG})_\varphi(x,y) = \varphi^{-1}\left(\min\{1, \frac{\varphi(y)}{\varphi(x)}\}\right),$$

para todo $x,y \in [0,1]$.

(ii) I es continua excepto en el punto $(0,0)$, satisface (I2), (EP), (OP) y $N_I = N_{D_1}$.

La caracterización de la implicación de Goguen que presentamos se basa en el Teorema 6 y en que esta implicación es la única g-generada con $g(1) < \infty$ que satisface (OP).

Teorema 8. *Dada una función $I : [0,1]^2 \to [0,1]$ las siguientes afirmaciones son equivalentes:*

(i) I es la implicación de Goguen, I_{GG}.

(ii) I satisface (OP) y (LI) con T_P.

5 CONCLUSIONES Y TRABAJO FUTURO

En esta comunicación, se han dado caracterizaciones axiomáticas de las implicaciones de Yager. Estos resultados se basan en algunas propiedades algebraicas de la función binaria inicial I, especialmente en la ley de importación con la t-norma producto T_P. Usando estos resultados, hemos caracterizado algunas implicaciones conocidas como las de Goguen, Reichenbach y Yager.

A continuación, sería interesante estudiar la caracterización de las implicaciones φ-conjugadas de las de Yager. También, ahora que se dispone de la caracterización de estas implicaciones, se puede resolver el problema abierto, Problema 4.8.3, en [7], sobre la intersección de las implicaciones de Yager con las QL o D-operaciones. Hemos abordado ya ambas cuestiones con resultados satisfactorios. Finalmente, recordar que existen otros problemas abiertos con respecto a las implicaciones de Yager como la caracterización de su envoltura convexa (Problema 6.5.1 en [7]) y el estudio de la propiedad de T-condicionalidad en esta clase de implicaciones.

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto nacional MTM2009-10320 con el apoyo de fondos FEDER.

Referencias

[1] I. Aguiló, J. Suñer, and J. Torrens. A characterization of residual implications derived from left-continuous uninorms. *Information Sciences*, 180(20):3992–4005, 2010.

[2] C. Alsina, M. Frank, and B. Schweizer. *Associative functions. Triangular norms and copulas*. World Scientific Publishing Company, 2006.

[3] M. Baczyński. On a class of distributive fuzzy implications. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9:229–238, 2001.

[4] M. Baczyński. Residual implications revisited. Notes on the Smets-Magrez Theorem. *Fuzzy Sets and Systems*, 145:267–277, 2004.

[5] M. Baczyński and B. Jayaram. On the characterization of (S,N)-implications. *Fuzzy Sets and Systems*, 158:1713–1727, 2007.

[6] M. Baczyński and B. Jayaram. Yager's classes of fuzzy implications: some properties and intersections. *Kybernetika*, 43:157–182, 2007.

[7] M. Baczyński and B. Jayaram. *Fuzzy Implications*. Springer, Berlin Heidelberg, 2008.

[8] M. Baczyński and B. Jayaram. (U,N)-implications and their characterizations. *Fuzzy Sets and Systems*, 160:2049–2062, 2009.

[9] J. Balasubramaniam. Yager's new class of implications J_f and some classical tautologies. *Information Sciences*, 177:930–946, 2007.

[10] B. De Baets. Idempotent uninorms. *European Journal of Operational Research*, 118(3):631 – 642, 1999.

[11] P. Drygas. On properties of uninorms with underlying t-norm and t-conorm given as ordinal sums. *Fuzzy Sets and Systems*, 161(2):149–157, 2010.

[12] J. C. Fodor. A new look at fuzzy connectives. *Fuzzy Sets and Systems*, 57:141–148, 1993.

[13] J. C. Fodor. Contrapositive symmetry on fuzzy implications. *Fuzzy Sets and Systems*, 69:141–156, 1995.

[14] J. C. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, Dordrecht, 1994.

[15] J. C. Fodor, R. R. Yager, and A. Rybalov. Structure of uninorms. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, 5:411–127, 1997.

[16] S. Hu and Z. Li. The structure of continuous uninorms. *Fuzzy Sets and Systems*, 124:43–52, 2001.

[17] B. Jayaram. On the law of importation $(x \wedge y) \rightarrow z \equiv (x \rightarrow (y \rightarrow z))$ in fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 16:130–144, 2008.

[18] E. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[19] J. Martin, G. Mayor, and J. Torrens. On locally internal monotonic operators. *Fuzzy Sets and Systems*, 137:27–42, 2003.

[20] M. Mas, M. Monserrat, and J. Torrens. QL versus D-implications. *Kybernetika*, 42:351–366, 2006.

[21] M. Mas, M. Monserrat, and J. Torrens. The law of importation for discrete implications. *Information Sciences*, 179:4208–4218, 2009.

[22] M. Mas, M. Monserrat, and J. Torrens. A characterization of (U,N), RU, QL and D-implications derived from uninorms satisfying the law of importation. *Fuzzy Sets and Systems*, 161:1369–1387, 2010.

[23] M. Mas, M. Monserrat, J. Torrens, and E. Trillas. A survey on fuzzy implication functions. *IEEE Transactions on Fuzzy Systems*, 15(6):1107–1121, 2007.

[24] S. Massanet and J. Torrens. The law of importation versus the exchange principle on fuzzy implications. *Fuzzy Sets and Systems*, 168(1):47 – 69, 2011.

[25] D. Ruiz-Aguilera, J. Torrens, B. De Baets, and J. Fodor. Some remarks on the characterization of idempotent uninorms. In E. Hüllermeier, R. Kruse, and F. Hoffmann, editors, *Computational Intelligence for Knowledge-Based Systems Design*, volume 6178 of *Lecture Notes in Computer Science*, pages 425–434. Springer Berlin / Heidelberg, 2010.

[26] E. Trillas. Sobre funciones de negación en la teoría de conjuntos difusos. *Stochastica III/1*, pages 47–60, 1979.

[27] E. Trillas, M. Mas, M. Monserrat, and J. Torrens. On the representation of fuzzy rules. *Int. J. Approx. Reasoning*, 48(2):583–597, 2008.

[28] R. Yager. On some new classes of implication operators and their role in approximate reasoning. *Information Sciences*, 167:193–216, 2004.

Solutions of systems of fuzzy relation equations as concepts of a formal context

Juan Carlos Díaz, Jesús Medina, Rafael Rodríguez

Department of Mathematics. University of Cádiz. Spain, {juancarlos.diaz,jesus.medina,rafael.rodriguez}@uca.es

Abstract

Fuzzy relation equations have been generalized in order to consider any residuated conjunctor. Moreover, the solutions of the systems of these general fuzzy relation equations and the concepts of a fuzzy property-oriented concept lattice have been related. These contributions are showed in a worked out example.

Keywords: Fuzzy relation equations; Galois connection; property-oriented concept lattice.

1 INTRODUCTION

Fuzzy relation equations, introduced by E. Sanchez [28], are associated with the composition of fuzzy relations and have been used to investigate theoretical and applicational aspects of fuzzy set theory [11], e.g., approximate reasoning, time series forecast, decision making, fuzzy control, as an appropriate tool for handling and modeling of non-probabilistic form of uncertainty, etc. Many papers have investigated the capacity to solve (systems of) fuzzy relation equations, e.g., in [1, 10, 12, 24, 26].

This paper presents a generalization of the fuzzy relation equations [28, 23]. This general environment provides more flexibility in order to relate the variables considered in the system.

Moreover, a fuzzy generalization of the property-oriented concept lattices, given by Düntsch and Gediga [13, 14], is introduced. These concept lattices and the generalized relation equations have been related in order to obtain results that assure the existence of solutions in these systems. These definitions and results are illustrated by a worked out example to improve the readability and comprehension of the paper.

The most important consequence is that this relation provides that the properties given, e.g., in [2, 3, 4, 15, 17, 21, 20, 27] can be applied to obtain many properties of these systems. Indeed, it can be considered that the algorithms presented, e.g., in [5, 6, 18] obtain solutions for these systems.

The plan of this paper is the following: Section 2 recalls classical fuzzy relation equations; in Section 3 fuzzy property-oriented concept lattices as well as the basic operators used will be presented; later, in Section 4, a general fuzzy relation equations environment is introduced, which is resolved using fuzzy property-oriented concept lattices theory, in Section 5; all these results are illustrated in a worked out example, in Section 6; the paper ends with some conclusions and prospects for future work.

2 FUZZY RELATION EQUATIONS

Fuzzy relation equations [10, 12, 26] are very important in applicative example, such as in approximate reasoning, time series forecast, decision making and fuzzy control, etc.

The operators used in order to obtain the systems of fuzzy relation equations are the sup-$*$-composition, denoted $\circ$ and introduced in [29], and inf-$\rightarrow$-composition, denoted $\triangleleft$ and introduced in [1]. From now on, *a residuated lattice* $\mathcal{L} = (L, \vee, \wedge, *, \leftarrow, 0, 1)$ will be fixed.

A *system of fuzzy relation equations with* sup-$*$-*composition* (SFRE$*$), is the following system of equations

$$A_i \circ R = B_i, \quad \text{or} \tag{1}$$

$$\bigvee_{u \in U} (A_i(u) * R(u,v)) = B_i(v) \tag{2}$$

for all $i \in \{1,\dots,n\}$, and where $R \in L^{U\times V}$ is an unknown fuzzy relation, $A_1,\dots,A_n \in L^U$ and $B_1,\dots,B_n \in L^V$. Its counterpart is a *system of fuzzy relation equations with* inf-$\rightarrow$*-composition* (SFRE$\rightarrow$), that is,

$$A_j^* \lhd R = D_j \quad \text{or} \tag{3}$$

$$\bigwedge_{v\in V} (A_j^*(v) \rightarrow R(u,v)) = D_j(u) \tag{4}$$

for all $j \in \{1,\dots,m\}$, and considered with respect to unknown fuzzy relation $R \in L^{U\times V}$, and where $A_1^*,\dots,A_m^* \in L^V$ and $D_1,\dots,D_m \in L^U$.

Let $U = \{u_1,\dots,u_m\}$ and $V = \{v_1,\dots,v_n\}$ be two universes, $R \in L^{U\times V}$ an unknown fuzzy relation and $A_1\dots,A_n \in L^U$, $B_1,\dots,B_n \in L^V$ as arbitrarily chosen fuzzy subsets of respective universes.

If an element of V is fixed, $v \in V$, $A_i(u_j) = a_{ij}$ for $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$, $R(u_j,v) = x_j$, $B_i(v) = b_i$, then System (1) can be written as

$$\begin{array}{rcl} a_{11} * x_1 \vee \cdots \vee a_{1m} * x_m &=& b_1 \\ \vdots \quad \vdots & & \vdots \quad \vdots \\ a_{n1} * x_1 \vee \cdots \vee a_{nm} * x_m &=& b_n \end{array} \tag{5}$$

Therefore, for each $v \in V$, we obtain a column of R, thus, solving n SFRE$*$ systems, we will obtain the unknown relation. In order to obtain a SFRE$\rightarrow$ system where the considered universes U, V and the unknown fuzzy relation $R \in L^{U\times V}$ are as in the previous system, an element of U is fixed, $u \in U$, fuzzy subsets $A_1^*,\dots,A_m^* \in L^V$, $D_1,\dots,D_m \in L^U$ are assumed, such that $A_j^*(v_i) = a_{ij}$ for $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$, $R(u,v_i) = y_i$ and $D_j(u) = d_j$. Hence, System (3) can be written as

$$\begin{array}{rcl} a_{11} \rightarrow y_1 \wedge \cdots \wedge a_{n1} \rightarrow y_n &=& d_1 \\ \vdots \quad \vdots & & \vdots \quad \vdots \\ a_{1m} \rightarrow y_1 \wedge \cdots \wedge a_{nm} \rightarrow y_n &=& d_m \end{array} \tag{6}$$

Hence, for each $u \in U$, we obtain a row of R and solving m SFRE$\rightarrow$ systems, the unknown relation will be obtained.

These above systems have the same goal, searching for the unknown relation R, although the mechanism is different.

The aim in [25] was to use the theory of Galois connections in order to characterize whether these systems are solvable and, if so, finding their solution sets.

The next sections will introduce a generalization of fuzzy relational equations in order to consider adjoint triples and the results given in [25] will be generalized. Moreover, the Galois conections will be related to a concept lattice environment.

The specific concept lattice environment that will be used is the fuzzy property-oriented concept lattice, which will be presented in the next section.

3 FUZZY PROPERTY-ORIENTED CONCEPT LATTICES

Adjoint triples are the basic operators in this environment, which are formed by three mappings: a non-commutativity conjunctor and two residuated implications [16], that satisfy the well-known adjoint property.

Definition 1 *Let* $(P_1,\leq_1)$, $(P_2,\leq_2)$, $(P_3,\leq_3)$ *be posets and* $\&\colon P_1\times P_2 \to P_3$, $\swarrow\colon P_3\times P_2 \to P_1$, $\nwarrow\colon P_3\times P_1 \to P_2$ *be mappings, then* $(\&,\swarrow,\nwarrow)$ *is an* adjoint triple *with respect to* P_1,P_2,P_3 *if:*

1. $\&$ *is order-preserving in both arguments.*
2. $\swarrow$ *and* $\nwarrow$ *are order-preserving on the first argument*[1] *and order-reversing on the second argument.*
3. $x \leq_1 z \swarrow y$ *iff* $x \,\&\, y \leq_3 z$ *iff* $y \leq_2 z \nwarrow x$, *where* $x \in P_1$, $y \in P_2$ *and* $z \in P_3$.

Example of adjoint triples are the Gödel, product and Łukasiewicz t-norms together with their residuated implications.

Example 1 *Since the Gödel, product and Łukasiewicz t-norms are commutative, the residuated implications satisfy that* $\swarrow^G = \nwarrow_G$, $\swarrow^P = \nwarrow_P$ *and* $\swarrow^L = \nwarrow_L$. *Therefore, the Gödel, product and Łukasiewicz adjoint triples are defined on* $[0,1]$ *as:*

$$\&_P(x,y) = x\cdot y \qquad z \nwarrow_P x = \min(1, z/x)$$

$$\&_G(x,y) = \min(x,y) \qquad z \nwarrow_G x = \begin{cases} 1 & \text{if } x \leq z \\ z & \text{otherwise} \end{cases}$$

$$\&_L(x,y) = \max(0, x+y-1)$$

$$z \nwarrow_G x = \min(1, 1-x+z)$$

In [22] more general examples of adjoint triples are given.

[1] Note that the antecedent will be evaluated on the right side, while the consequent will be evaluated on the left side, as in logic programming framework.

Example 2 *Let* $[0,1]_m$ *be a regular partition of* $[0,1]$ *into* m *pieces, for example* $[0,1]_2 = \{0, 0.5, 1\}$ *divide the unit interval into two pieces.*

A discretization of the product t-norm is, e.g., the operator $\&_P^*\colon [0,1]_{20} \times [0,1]_8 \to [0,1]_{100}$ *defined, for each* $x \in [0,1]_{20}$ *and* $y \in [0,1]_8$ *as:*

$$x \,\&_P^*\, y = \frac{\lceil 100 \cdot x \cdot y \rceil}{100}$$

where $\lceil _ \rceil$ *is the ceiling function.*

For this operator, the corresponding residuated implication operators $\swarrow_P^*\colon [0,1]_{100} \times [0,1]_8 \to [0,1]_{20}$ *and* $\nwarrow_P^*\colon [0,1]_{100} \times [0,1]_{20} \to [0,1]_8$ *are defined as:*

$$\begin{aligned} b \swarrow_P^* a &= \frac{\lfloor 20 \cdot \min\{1, b/a\} \rfloor}{20} \\ b \nwarrow_P^* c &= \frac{\lfloor 8 \cdot \min\{1, b/c\} \rfloor}{8} \end{aligned}$$

where $\lfloor _ \rfloor$ *is the floor function.*

The triple $(\&_P^*, \swarrow_P^*, \nwarrow_P^*)$ *is an adjoint triple, although the operator* $\&_P^*$ *is neither commutative nor associative.*

The basic structure, which fixes the triplet of lattices and the adjoint triple, is the fuzzy property-oriented frame.

Definition 2 *Given two complete lattices* $(L_1, \preceq_1)$ *and* $(L_2, \preceq_2)$*, a poset* $(P, \leq)$ *and one adjoint triple with respect to* P, L_2, L_1*,* $(\&, \swarrow, \nwarrow)$*, a* fuzzy property-oriented frame *is the tuple*

$$(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&)$$

The definition of context is analogous to the one given in [21].

Definition 3 *Let* $(L_1, L_2, P, \&)$ *be a fuzzy property-oriented frame. A* context *is a tuple* (A, B, R) *such that* A *and* B *are non-empty sets (usually interpreted as attributes and objects, respectively),* R *is a* P*-fuzzy relation* $R\colon A \times B \to P$

From now on, we will fix a fuzzy property-oriented frame and context, $(L_1, L_2, P, \&)$, (A, B, R).

Now we define the following mappings $^{\uparrow_\pi}\colon L_2^B \to L_1^A$ and $^{\downarrow^N}\colon L_1^A \to L_2^B$ as

$$\begin{aligned} g^{\uparrow_\pi}(a) &= \sup\{R(a,b) \,\&\, g(b) \mid b \in B\} \\ f^{\downarrow^N}(b) &= \inf\{f(a) \nwarrow R(a,b) \mid a \in A\} \end{aligned}$$

Clearly, these definitions generalize the classical possibility and necessity operators [14]. Moreover, $(^{\uparrow_\pi}, ^{\downarrow^N})$ is an isotone Galois connection and $^{\uparrow_\pi\downarrow^N}\colon L_2^B \to L_2^B$ is a closure operator and $^{\downarrow^N\uparrow_\pi}\colon L_1^A \to L_1^A$ an interior operator.

A concept, in this environment, is a pair of mappings $\langle g, f\rangle$, with $g \in L^B, f \in L^A$, such that $g^{\uparrow_\pi} = f$ and $f^{\downarrow^N} = g$, which will be called *fuzzy property-oriented formal concept*. In that case, g is called the *extent* and f, the *intent* of the concept. The set of all these concepts will be notated as $\mathcal{F}_{\pi N}$.

Definition 4 *A* fuzzy property-oriented concept lattice *is the set*

$$\mathcal{F}_{\pi N} = \{\langle g, f\rangle \mid g \in L_2^B, f \in L_1^A \text{ and } g^{\uparrow_\pi} = f, f^{\downarrow^N} = g\}$$

in which the ordering is defined by $\langle g_1, f_1\rangle \preceq \langle g_2, f_2\rangle$ *iff* $g_1 \preceq_2 g_2$ *(or equivalently* $f_1 \preceq_1 f_2$*).*

The pair $(\mathcal{F}_{\pi N}, \preceq)$ is a complete lattice [19], which generalizes the concept lattice introduced in [9] to a fuzzy environment.

4 GENERAL FUZZY RELATION EQUATIONS

This section considers adjoint triples as generalization of the sup-$*$-composition, introduced in [29], and inf-$\to$-composition, introduced in [1]. Hence, from now on, a fuzzy property-oriented frame, $(L_1, L_2, P, \&)$ will be fixed.

Let $U = \{u_1, \ldots, u_m\}$ and $V = \{v_1, \ldots, v_n\}$ be two universes, $R \in {L_2}^{U \times V}$ an unknown fuzzy relation, and $K_1 \ldots, K_n \in P^U$, $D_1, \ldots, D_n \in {L_1}^V$ as arbitrarily chosen fuzzy subsets of respective universes.

A *system of fuzzy relation equations with* sup-&-*composition* is the following system of equations

$$\bigvee_{u \in U} (K_i(u) \,\&\, R(u,v)) = D_i(v) \tag{7}$$

for all $i \in \{1, \ldots, n\}$.

If an element v of V is fixed and for each $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$, $K_i(u_j) = k_{ij}$, $R(u_j, v) = x_j$, $D_i(v) = d_i$, then System (7) can be written as

$$\begin{aligned} k_{11} \,\&\, x_1 \vee \cdots \vee k_{1m} \,\&\, x_m &= d_1 \\ \vdots \quad \vdots \qquad\qquad \vdots \quad \vdots & \\ k_{n1} \,\&\, x_1 \vee \cdots \vee k_{nm} \,\&\, x_m &= d_n \end{aligned} \tag{8}$$

Hence, for each $v \in V$, if we solve System (8), then we obtain a "column" of R (i.e. the elements $R(u_j, v)$, with $j \in \{1, \ldots, m\}$). Thus, solving n similar systems, one for each $v \in V$, the unknown relation R is obtained.

Its counterpart is a *system of fuzzy relation equations with* inf-$\nwarrow$-*composition*, that is

$$\bigwedge_{v\in V}(R(u,v)\nwarrow K_j^*(v)) = E_j(u) \tag{9}$$

for all $j\in\{1,\dots,m\}$, that is considered with respect to unknown fuzzy relation $R\in {L_1}^{U\times V}$, and where $K_1^*,\dots,K_m^*\in P^V$ and $E_1,\dots,E_m\in {L_2}^U$.

If an element $u\in U$ is fixed, fuzzy subsets $K_1^*,\dots,K_m^*\in P^V$, $E_1,\dots,E_m\in L_2^U$ are assumed, such that $K_j^*(v_i)=k_{ij}$, $R(u,v_i)=y_i$ and $E_j(u)=e_j$, for each $i\in\{1,\dots,n\}$, $j\in\{1,\dots,m\}$, then System (9) can be written as

$$\begin{array}{rcccc} y_1\nwarrow k_{11}\wedge\cdots\wedge y_n\nwarrow k_{n1} & = & e_1 \\ \vdots \quad\quad \vdots & & \vdots & \vdots \\ y_1\nwarrow k_{1m}\wedge\cdots\wedge y_n\nwarrow k_{nm} & = & e_m \end{array} \tag{10}$$

Hence, for each $u\in U$, we obtain a "row" of R (i.e. the elements $R(u,v_i)$, with $i\in\{1,\dots,n\}$), consequently, solving m similar systems, the unknown relation R is obtained.

Systems (8) and (10) have the same aim, searching for the unknown relation R although the mechanism is different.

5 SOLVING GENERAL FUZZY RELATION EQUATIONS

Now we will see that Systems (8) and (10) can be interpreted as a fuzzy property-oriented concept lattice. And so, the properties given to the isotone Galois connection $(^{\uparrow_\pi},{}^{\downarrow^N})$, as well as to the complete lattice $\mathcal{F}_{\pi N}$ can be used in the resolution of Systems (8) and (10).

Therefore, the environment needed to define Systems (8) and (10) is considered.

Let $C_K\colon L_2^m\to L_1^n$ be the mapping, where $C_K(\bar{x})$ has n components $C_K(\bar{x})_i$, defined as:

$$C_K(\bar{x})_i \;=\; k_{i1}\,\&\,x_1\vee\cdots\vee k_{im}\,\&\,x_m \tag{11}$$

for each $i\in\{1,\dots,n\}$ and $\bar{x}=(x_1,\dots,x_m)\in L_2^m$, and the mapping $I_{K^*}\colon L_1^n\to L_2^m$ has m components $I_{K^*}(\bar{y})_j$, defined as

$$I_{K^*}(\bar{y})_j \;=\; y_1\nwarrow k_{1j}\wedge\cdots\wedge y_n\nwarrow k_{nj} \tag{12}$$

for each $j\in\{1,\dots,m\}$ and $\bar{y}=(y_1,\dots,y_n)\in L_1^n$.

Given a system of fuzzy relation equations, defined above, a fuzzy context (A,B,S) can be considered, where $A=\{a_1,\cdots,a_n\}$ is a set of arbitrary attributes with the same cardinality of V, set $B=\{b_1,\cdots,b_m\}$ is formed by arbitrary objects with the same cardinality of U, and $S\colon A\times B\to P$ is defined as $S(a_i,b_j)=k_{ij}$, for all $i\in\{1,\dots,n\}$, $j\in\{1,\dots,m\}$.

In this framework, we obtain the following result:

Lemma 1 *For each $\bar{x}\in L_2^m$ and $\bar{y}\in L_1^n$, there exist $\mu\in L_2^B$ and $\lambda\in L_1^A$, such that:*

$$\begin{array}{rcl} \mu^{\uparrow_\pi}(a_i) & = & C_K(\bar{x})_i \\ \lambda^{\downarrow^N}(b_j) & = & I_{K^*}(\bar{y})_j \end{array}$$

for all $i\in\{1,\dots,n\}$ and $j\in\{1,\dots,m\}$.

Therefore, the mappings $C_K\colon L_2^m\to L_1^n$ and $I_{K^*}\colon L_1^n\to L_2^m$ are equivalents to the mappings $^{\uparrow_\pi}\colon L_2^B\to L_1^A$ and $^{\downarrow^N}\colon L_1^A\to L_2^B$, respectively. As a consequence, the following proposition is obtained.

Proposition 1 *The maps $C_K\colon L_2^m\to L_1^n, I_{K^*}\colon L_1^n\to L_2^m$, establish an isotone Galois connection. And $I_{K^*}\circ C_K\colon L_2^m\to L_2^m$ is a closure operator and $C_K\circ I_{K^*}\colon L_1^n\to L_1^n$ an interior operator.*

As (C_K,I_{K^*}) is an isotone Galois connection, any result in order to solve one system of fuzzy relation equations has its dual counterpart in the other system, which can be obtained by replacing C_K by I_{K^*}, $\leq$ by $\geq$, $\vee$ by $\wedge$, respectively. The following result explains when these systems can be solved and how a solution can be obtained.

Theorem 1 *Considering the above environment to define Systems (8) and (10), we obtain that System (8) can be solved if and only if $I_{K^*}(\bar{d})$ is a solution. In this case, $I_{K^*}(\bar{d})$ is the largest solution.*

Furthermore, System (10) can be solved if and only if $C_K(\bar{e})$ is a solution, which is the smallest solution.

Hence, given the environment needed to define Systems (8) and (10), there exists a context (A,B,S) such that resolving System (8) is equivalent to obtaining a mapping $\mu\in L_2^B$, which holds

$$\mu^{\uparrow_\pi} \;=\; \lambda_{\bar{d}} \tag{13}$$

where $\lambda_{\bar{d}}\in L_1^A$, defined as $\lambda_{\bar{d}}(a_i)=d_i$, for all $i\in\{1,\dots,n\}$. Analogously, there exists a context (A,B,S) such that System (10) is solvable if and only if there exists a mapping $\lambda\in L_1^A$, which holds

$$\lambda^{\downarrow^N} \;=\; \mu_{\bar{e}} \tag{14}$$

where $\mu_{\bar{e}}\in L_2^B$, defined as $\mu_{\bar{e}}(b_j)=e_j$, for all $j\in\{1,\dots,m\}$.

Therefore, Theorem 1 can be rewritten as follows:

Lemma 2 *Considering the above environment, we obtain that System (8) can be solved if and only if* $\lambda_{\bar{d}}^{\downarrow^N\uparrow_\pi} = \lambda_{\bar{d}}$.

Analogously, System (10) can be solved if and only if $\mu_{\bar{e}}^{\uparrow_\pi\downarrow^N} = \mu_{\bar{e}}$.

The following consequences can be stabilized from Lemma 2 and the definition of fuzzy property-oriented concept lattice $\mathcal{F}_{\pi N}$, assuming the above framework.

Theorem 2 *System (8) can be solved if and only if* $\langle \lambda_{\bar{d}}^{\downarrow^N}, \lambda_{\bar{d}} \rangle$ *is a concept of* $\mathcal{F}_{\pi N}$. *Moreover, System (10) can be solved if and only if* $(\mu_{\bar{e}}, \mu_{\bar{e}}^{\uparrow_\pi})$ *is a concept of* $\mathcal{F}_{\pi N}$.

The main contribution of the relation introduced in this paper is not only the above consequences, but a lot of other properties for Systems (8) and (10) that can be stabilized from the results proved, for example, in [2, 3, 4, 15, 17, 27], applying Theorem 2.

Moreover, in order to obtain the solutions of Systems (8) and (10), the algorithms developed, for example, in [5, 6, 18], can be applied.

6 A WORKED OUT EXAMPLE

Now we will consider an example used in [7, 8]. In these papers the authors assume a fuzzy relational equation and they look for (minimal) solutions. The following example will consider this fuzzy relation equation, together with other equations in order to form a system of fuzzy relational equations, and the results above will be used to prove that the system has a solution and to obtain the maximum solution. In order to obtain the maximum solution, an algorithm to get the concepts of a fuzzy property-oriented concept lattice will be applied.

Let $U = \{u_1, u_2, u_3, u_4\}$ and $V = \{v_1, v_2, v_3, v_4, v_5\}$ be two universes, $R \in [0,1]^{U\times V}$ an unknown fuzzy relation, which is represented by the matrix $R = (R(u_i, v_j))$ and $K_1, K_2, K_3, K_4, K_5 \in [0,1]^U$, $D_1, D_2, D_3, D_4, D_5 \in [0,1]^V$ fuzzy sets in U and V, defined by the following matrices K and D

$$K = \begin{pmatrix} 0.3 & 0.2 & 0.7 & 0.8 \\ 0.5 & 0.4 & 0.4 & 0.9 \\ 0.7 & 0.3 & 0.2 & 0.7 \\ 0.9 & 0.6 & 0.1 & 0.2 \\ 0.8 & 0.5 & 0.6 & 0.4 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.7 & 0.3 & 0.3 & 0.5 & 0.8 \\ 0.4 & 0.4 & 0.4 & 0.5 & 0.9 \\ 0.4 & 0.3 & 0.3 & 0.7 & 0.7 \\ 0.3 & 0.4 & 0.6 & 0.9 & 0.5 \\ 0.6 & 0.4 & 0.5 & 0.8 & 0.6 \end{pmatrix}$$

as $K_i(u_j)) = K(i,j)$ and $D_i(v_j) = D(i,j)$. That is, the mappings K_i and D_i are represented by the rows of K and D, respectively.

Firstly, the goal will be to obtain a matrix R, such that $K \odot R = D$, where $\odot$ is the *sup-&-composition* and the conjunctor & is the minimum, $\&(x,y) = min(x,y)$.

Hence, the fuzzy property-oriented frame considered will be $([0,1],[0,1],[0,1],\&)$.

From the system given, we need to consider the context (A, B, S), where $A = \{a_1, a_2, a_3, a_4, a_5\}$, $B = \{b_1, b_2, b_3, b_4\}$, and $S \in [0,1]^{A\times B}$ defined as $S(a_i, b_j) = k_{ij}$. Note that the number of attributes is equal to the cardinality of V and the number of objects is equal to the cardinality of U.

Let $\mathcal{F}_{\pi N}$ be the fuzzy property-oriented concept lattice associated to the frame and context above. Following the steps given by the theoretical results, we need to consider the fuzzy subsets $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \in [0,1]^A$ associated to the columns of the matrix D, that is, for λ_1 we have $\lambda_1(a_1) = 0.7$, $\lambda_1(a_2) = 0.4$, $\lambda_1(a_3) = 0.4$, $\lambda_1(a_4) = 0.3$, $\lambda_1(a_5) = 0.6$. Therefore, $\lambda_j(a_i) = D(i,j)$, for all $i, j \in \{1, \ldots, 5\}$.

Now, we need to check whether the fuzzy subsets of attributes $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ are the intent of a concept in $\mathcal{F}_{\pi N}$.

For that, first of all, it is necessary to compute the fuzzy subsets of objects $\lambda_1^{\downarrow^N}, \lambda_2^{\downarrow^N}, \lambda_3^{\downarrow^N}, \lambda_4^{\downarrow^N}, \lambda_5^{\downarrow^N}$. Then we apply the mapping to these subsets and, finally, we must check the following equality:

$$\lambda_j^{\downarrow^N\uparrow_\pi} = \lambda_j$$

for all $j \in \{1, \ldots, 5\}$.

If these equalities are verified, then the system is solvable and the maximum solution is given by the matrix R in which each element of the matrix $R(i,j)$ is $\lambda_j^{\downarrow^N}(b_i)$, for all $i \in \{1, \ldots, 4\}$ and $j \in \{1, \ldots, 5\}$.

In order to obtain the concepts of the fuzzy property-oriented concept lattice assumed, we have used an implementation of the necessity and possibility operators in PYTHON programming language.

$$\lambda_1^{\downarrow^N}(b_1) = \inf\{0.7 \nwarrow 0.3, 0.4 \nwarrow 0.5, 0.4 \nwarrow 0.7 \\ 0.3 \nwarrow 0.9, 0.6 \nwarrow 0.8\}$$

Hence, $\lambda_1^{\downarrow^N}(b_1) = 0.3$. The mapping $\lambda_1^{\downarrow^N}$ can be interpreted as a list of values, the truth values for each object and its computation could be represented as the

following list

$$\begin{array}{l}(\inf\{0.7 \nwarrow 0.3, 0.4 \nwarrow 0.5, 0.4 \nwarrow 0.7, 0.3 \nwarrow 0.9, 0.6 \nwarrow 0.8\},\\ \inf\{0.7 \nwarrow 0.2, 0.4 \nwarrow 0.4, 0.4 \nwarrow 0.3, 0.3 \nwarrow 0.6, 0.6 \nwarrow 0.5\},\\ \inf\{0.7 \nwarrow 0.7, 0.4 \nwarrow 0.4, 0.4 \nwarrow 0.2, 0.3 \nwarrow 0.1, 0.6 \nwarrow 0.6\},\\ \inf\{0.7 \nwarrow 0.8, 0.4 \nwarrow 0.9, 0.4 \nwarrow 0.7, 0.3 \nwarrow 0.2, 0.6 \nwarrow 0.4\})\\ = (0.3, 0.3, 1, 0.4)\end{array}$$

Now, $\lambda_1^{\downarrow^N\uparrow_\pi}$ must be obtained in order to check whether the equality $\lambda_1^{\downarrow^N\uparrow_\pi} = \lambda_1$ is satisfied. The algorithm is applied once again and we obtain:

$$\lambda_1^{\downarrow^N\uparrow_\pi} = (0.7, 0.4, 0.4, 0.3, 0.6) = \lambda_1$$

Therefore, the system associated to the first column of D has a solution and this solution (the vector $(0.3, 0.3, 1, 0.4)$) is equal to the one obtained in [8] but now using a fuzzy concept lattice framework and an algorithm developed in this mathematical environment.

Once the verification of the equality for λ_1 is done, we need to compute the rest of subsets of objects. These subsets will be represented by the following lists.

$$\begin{array}{rcl}\lambda_2^{\downarrow^N} & = & (0.3, 0.4, 0.3, 0.3)\\ \lambda_3^{\downarrow^N} & = & (0.3, 1, 0.3, 0.3)\\ \lambda_4^{\downarrow^N} & = & (1, 1, 0.5, 0.5)\\ \lambda_5^{\downarrow^N} & = & (0.5, 0.5, 1, 1)\end{array}$$

Finally, we check that, effectively, the equalities $\lambda_j^{\downarrow^N\uparrow_\pi} = \lambda_j$, with $j \in \{2, \ldots, 5\}$, hold.

Consequently, the equality $K \odot R = D$ can be solved and the maximum solution is given by the matrix:

$$R = \begin{pmatrix} 0.3 & 0.3 & 0.3 & 1 & 0.5\\ 0.3 & 0.4 & 1 & 1 & 0.5\\ 1 & 0.3 & 0.3 & 0.5 & 1\\ 0.4 & 0.3 & 0.3 & 0.5 & 1\end{pmatrix}$$

Thus, we have a new tool in order to prove the existence of solutions of the equations $K \odot R = D$, where R is an unknown matrix. Moreover, fuzzy property-oriented concept lattices theory is used to obtain the maximum solution.

In this example, the "product" (composition) of the matrices $\odot$ is defined from the Gödel conjunctor, as in [7, 8]. However, we must remark that any non-commutative residuated conjunctor can be considered, as the one given in Example 2, which improves the flexibility of the language and it adds more possibilities to applicative examples.

For instance, if the operator $\odot$ is related to the product conjunctor, that is,

$$\&_P(x, y) = x \cdot y \qquad z \nwarrow_P x = \min(1, z/x)$$

which will be denoted as $\odot_P$, we obtain that

$$\lambda_1^{\downarrow^N} = (0.3\hat{3}, 0.5, 1, 0.4\hat{4})$$

and

$$\lambda_1^{\downarrow^N\uparrow_\pi} = (0.7, 0.4, 0.31\hat{1}, 0.3, 0.6) \neq \lambda_1$$

Hence, the system associated to the first column of D does not have a solution, although the difference between $\lambda_1^{\downarrow^N\uparrow_\pi}$ and λ_1 is not very large. We have checked above that, changing the product conjunctor to the Gödel one, we find a solution for this system. Therefore, one possibility to obtain a solution of a system is to change the corresponding conjunctor. In general, this modification of the system can work or not. In the future, we will study conditions and mechanisms to get these operators from data.

Thus, if we know that the equation $K \odot R = D$ has a solution, then we might define a conjunctor ad hoc for which the system has solutions or, if this is not possible, we will define the conjunctor so that each column of D will be as close as possible to an intent of a concept of the corresponding fuzzy property-oriented concept lattice.

7 Conclusions and future work

Considering adjoint triples, the fuzzy relational equations have been generalized. Moreover, several of the results presented in [25] have been obtained directly from the theory developed for the fuzzy property-oriented concept lattice framework. Furthermore, several of the other properties provided, e.g., in [2, 3, 4, 15, 17, 27] can be applied to obtain additional characteristics of these systems or other more general ones.

Besides to obtain the largest solution of a fuzzy relational equation system, another important problem in this theory is to find the minimal solutions of this kind of systems. In the future, we will try to use concept lattice theory to find these special solutions.

Moreover, we will compare the actual methods with the use of fuzzy property-oriented concept lattices theory in order to know whether these systems has a solution and to obtain the maximum solution. Furthermore, we will use the concept lattice theory to find minimal solutions, which is another important problem in fuzzy relational equations theory.

Acknowledgements

Partially supported by Junta de Andalucía grant P09-FQM-5233, and by the EU (FEDER), and the Spanish Science and Education Ministry (MEC) under grant TIN2009-14562-C05-03.

References

[1] W. Bandler and L. Kohout. Semantics of implication operators and fuzzy relational products. *Int. J. Man-Machine Studies*, 12:89–116, 1980.

[2] E. Bartl, R. Bělohlávek, J. Konecny, and V. Vychodil. Isotone galois connections and concept lattices with hedges. In *4th International IEEE Conference "Intelligent Systems"*, pages 15.24–15.28, 2008.

[3] R. Bělohlávek. Lattices of fixed points of fuzzy Galois connections. *Mathematical Logic Quartely*, 47(1):111–116, 2001.

[4] R. Bělohlávek. Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic*, 128:277–298, 2004.

[5] R. Bělohlávek, B. D. Baets, J. Outrata, and V. Vychodil. Lindig's algorithm for concept lattices over graded attributes. *Lecture Notes in Computer Science*, 4617:156–167, 2007.

[6] R. Bělohlávek, B. D. Baets, J. Outrata, and V. Vychodil. Computing the lattice of all fixpoints of a fuzzy closure operator. *IEEE Transactions on Fuzzy Systems*, 18(3):546–557, 2010.

[7] L. Chen and P. Wang. Fuzzy relation equations (ii): The branch-point-solutions and the categorized minimal solutions. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 11:33–40, 2007. 10.1007/s00500-006-0050-1.

[8] L. Chen and P. P. Wang. Fuzzy relation equations (i): the general and specialized solving algorithms. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 6:428–435, 2002. 10.1007/s00500-001-0157-3.

[9] Y. Chen and Y. Yao. A multiview approach for intelligent data analysis based on data operators. *Information Sciences*, 178(1):1–20, 2008.

[10] B. De Baets. Analytical solution methods for fuzzy relation equations. In D. Dubois and H. Prade, editors, *The Handbooks of Fuzzy Sets Series*, volume 1, pages 291–340. Kluwer, Dordrecht, 1999.

[11] A. Di Nola, E. Sanchez, W. Pedrycz, and S. Sessa. *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*. Kluwer Academic Publishers, Norwell, MA, USA, 1989.

[12] A. Di Nola, S. Sessa, W. Pedrycz, and E. Sanchez. *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*. Kluwer, 1989.

[13] I. Düntsch and G. Gediga. Approximation operators in qualitative data analysis. In *Theory and Applications of Relational Structures as Knowledge Instruments*, pages 214–230, 2003.

[14] G. Gediga and I. Düntsch. Modal-style operators in qualitative data analysis. In *Proc. IEEE Int. Conf. on Data Mining*, pages 155–162, 2002.

[15] G. Georgescu and A. Popescu. Non-dual fuzzy connections. *Arch. Math. Log.*, 43(8):1009–1039, 2004.

[16] P. Hájek. *Metamathematics of Fuzzy Logic*. Trends in Logic. Kluwer Academic, 1998.

[17] H. Lai and D. Zhang. Concept lattices of fuzzy contexts: Formal concept analysis vs. rough set theory. *International Journal of Approximate Reasoning*, 50(5):695–707, 2009.

[18] C. Lindig. Fast concept analysis. In G. Stumme, editor, *Working with Conceptual Structures-Contributions to ICCS 2000*, pages 152–161, 2000.

[19] J. Medina. Towards multi-adjoint property-oriented concept lattices. *Lect. Notes in Artificial Intelligence*, 6401:159–166, 2010.

[20] J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.

[21] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets and Systems*, 160(2):130–144, 2009.

[22] J. Medina, M. Ojeda-Aciego, A. Valverde, and P. Vojtáš. Towards biresiduated multi-adjoint logic programming. *Lect. Notes in Artificial Intelligence*, 3040:608–617, 2004.

[23] W. Pedrycz. Fuzzy relational equations with generalized connectives and their applications. *Fuzzy Sets and Systems*, 10(1-3):185 – 201, 1983.

[24] I. Perfilieva. Fuzzy function as an approximate solution to a system of fuzzy relation equations. *Fuzzy Sets and Systems*, 147(3):363–383, 2004.

[25] I. Perfilieva. Fuzzy relation equations in semilinear spaces. In E. Hüllermeier, R. Kruse, and F. Hoffmann, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, volume 80 of *Communications in Computer and Information Science*, pages 545–552. Springer, 2010.

[26] I. Perfilieva and L. Nosková. System of fuzzy relation equations with inf-$\rightarrow$ composition: Complete set of solutions. *Fuzzy Sets and Systems*, 159(17):2256–2271, 2008.

[27] A. M. Radzikowska and E. E. Kerre. A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems*, 126(2):137–155, 2002.

[28] E. Sanchez. Resolution of composite fuzzy relation equations. *Information and Control*, 30(1):38–48, 1976.

[29] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I, II, III. *Information Sciences*, 8–9:199–257, 301–357, 43–80, 1975.

Possibilistic evaluation of fuzzy temporal intervals

José Enrique Pons[1], **Antoon Bronselaer**[2], **Olga Pons**[1], **Guy De Tré**[2]

[1]*Department of Computer Science and Artificial Intelligence,University of Granada, Spain*, {jpons,opc}@decsai.ugr.es
[2]*Department of Telecommunications and Information Processing, Ghent University, Belgium*,
{Antoon.Bronselaer,Guy.De.Tre}@telin.ugent.be

Abstract

The study of fuzzy intervals is of particular interest in temporal database research. In order to optimize the storage of fuzzy temporal intervals, some transformations have been proposed. In this paper we analyze the possibilistic evaluation of the ill-known temporal intervals. We propose a framework to deal with the evaluation of ill-known temporal intervals. It is shown how the reasoning behind the transformations implies a possibilistic information lost.

Keywords: Possibility theory, ill-known values, ill-known sets, intervals, temporal databases.

1 INTRODUCTION

The connection between possibility theory [12],[20],[3] and fuzzy sets has been studied among others by Zadeh [11]. This connection establishes two different natures for a fuzzy set [5]:

- **Conjunctive nature**: The fuzzyfication of a regular set. This interpretation corresponds with the following two semantics: *Degree of preference* and *degree of similarity*.
- **Disjunctive nature**: In this case, the disjunctive nature indicates a description of incomplete knowledge. This interpretation corresponds with the semantics for the *degree of uncertainty*.

The disjunctive nature can be seen as an additional connection between fuzzy set theory and possibility theory. Dubois and Prade [4] have studied the problem of inferring uncertainty about sets from uncertainty about values. They explain how the incomplete knowledge about a set can be represented as a possibility distribution over a universe $\mathscr{P}(U)$, the power set over the universe U. The main drawback of the approach is the large overhead that such a distribution introduces. A mechanism for the propagation of uncertainty about set elements towards uncertainty about sets is introduced in *'Possibilistic evaluation of sets'* [18].

The transformation from two ill-known temporal points to an ill-known set has been discussed in [9]. In this work we will analyze the major pitfalls of these approaches and the benefit of using the mechanism proposed in [18].

The paper is organized as follows. In Section 2, some preliminaries and concepts about the possibilistic evaluation framework and the temporal databases are presented. The proposal for interval evaluation is done in Section 3. In Section 4, the analysis among the different transformations is done. Finally, Section 5 presents the main conclusions and some lines for the future work in this research.

2 PRELIMINARIES

In this section we will introduce some basic concepts. First of all, an introduction to possibility theory is included. Then, the concepts of fuzzy numbers and fuzzy intervals introduced by Dubois and Prade [6] are discussed. Finally, a brief introduction to temporal databases is done.

2.1 Possibilistic variables

A possibilistic variable X over a universe U is defined as a variable taking exactly one value in U, but for which this value is (partially) unknown. The possibility distribution π_X gives the available knowledge about the value that X takes. For each $u \in U$, $\pi_X(u)$ represents the possibility that X takes the value u.

It is important to understand the difference between the following two concepts:

- A *possibilistic variable X* is bounded to take only one value , but this value is not known due to incomplete

knowledge.

- An *ill-known set* [4]: a possibilistic variable defined over the universe $\mathscr{P}(U)$.

Note that while a possibilistic variable refers to one (partially) unknown value, an ill-known set is a crisp set but, for some reason, (partially) unknown.

2.2 Fuzzy numbers and fuzzy intervals

Dubois and Prade [6] proposed the following definition of *fuzzy interval*. A fuzzy interval is a fuzzy set M on the set of real numbers $\mathbb{R}$ such that:

$$\forall (u,v) \in \mathbb{R}^2 : \tag{1}$$
$$\forall w \in [u,v] : \quad \mu_M(w) \geq \min(\mu_M(u), \mu_M(v))$$
$$\exists m \in \mathbb{R} : \quad \mu_M(m) = 1 \tag{2}$$

If m is unique, then M is referred to as a *fuzzy number*, instead of a *fuzzy interval*. In other words, if the core of a fuzzy interval is a singleton, it can be seen as a fuzzy number. In their work, Dubois and Prade propose four different functions (two possibility and two necessity functions) to asses the position of a fuzzy number N relative to a fuzzy number M taken as a reference.

The most convenient representation for the membership function of a fuzzy number is a triangular function. The membership function μ_M for the fuzzy set M has also the properties of convexity and normalization. Three values represent a triangular function:

- D is the singleton value in the core of $\mu_M(x)$.
- $D-a$ is the lower value in the support of $\mu_M(x)$.
- $D+b$ is the upper value in the support of $\mu_M(x)$.

Note that the values a and b are values in the underlying ordered domain. E.g. $(a,b) \in \mathbb{R}^2$. The notation for a triangular function adopted here is $[D,a,b]$.

The most simple representation for the membership function of a fuzzy interval is a trapezoidal function. This membership function μ_T for the fuzzy interval T is convex and normalized. Four values represent a trapezoidal membership function $[\alpha,\ \beta,\ \gamma,\ \delta]$.

2.3 Temporal databases

A temporal database [8] is a database that manages some aspects of time in its schema. A **chronon** is the shortest duration of time supported by the database. The time can be represented either as points or intervals [2]. Fuzzy temporal models [19] are proposed when the time points or intervals are not precisely known. There are proposals for the fuzzyfication of the time point [7] and the fuzzyfication of the time interval [9].
Allen [1] studied the comparison between two crisp time intervals. For fuzzy intervals, several proposals [19], [15],[17] have been done.
Despite of user-defined time (an uninterpreted attribute; supported in the standard SQL2 [14]), there are 3 types of time in a temporal database:

- **Transaction time**: The time when the fact is stored in the database.
- **Valid time**: The time when the fact is true in the modelled reality.
- **Decision time** [16]: The time when an event was decided to happen.

The database model can be then classified into: transaction time [13],[10], Valid time, bi-temporal [21](both valid and transaction time) or tri-temporal [16] (valid, transaction and decision time).

Fuzzy temporal models [19] deal with time points [7] or intervals [9] that are ill-known. Some fuzzy temporal models assume that the time stored in the database is an interval. The temporal interval is represented by two ill-known time points: X an ill-known starting point and Y an ill-known ending point. The interval $[X,Y]$ is not a fuzzy interval but an ill-known interval: it is a crisp interval but it is partially unknown which values are in this interval.

3 INTERVAL EVALUATION BY ILL-KNOWN CONSTRAINTS

The problem of the interval evaluation is explained in [18]: For a crisp interval $I = [a,b]$, we want to know if all points in this interval reside between the boundaries of the ill-known interval $[X,Y]$, where X,Y are ill-known values on the set of real numbers $\mathbb{R}$, and λ is the evaluation function. The following expressions compute the possibility and the necessity:

$$\text{Pos}\,(\lambda([a,b])) = \tag{3}$$
$$\min\left(\sup_{a\geq w} \pi_X(w), \sup_{b\leq w} \pi_Y(w)\right)$$
$$\text{Nec}\,(\lambda([a,b])) = \tag{4}$$
$$\min\left(\inf_{a<w} 1-\pi_X(w), \inf_{b>w} 1-\pi_Y(w)\right).$$

Note that the expressions in (3),(4) cannot be simplified in a more efficient expressions. An ill-known point is given by a triangular membership function. The ill-known values X

and Y are represented by the triangular membership functions $[D_X, a_X, b_X]$ and $[D_Y, a_Y, b_Y]$ respectively. The equation (3) is then derived to:

$$\sup_{a \geq w} \pi_X(w) = \begin{cases} 0 & \text{if } a \leq (D_X - a_X) \\ 1 & \text{if } a \geq D_X \\ \frac{a - D_X + a_X}{a_X} & \text{if } a \in (D_X - a_X, D_X) \end{cases} \quad (5)$$

$$\sup_{b \leq w} \pi_Y(w) = \begin{cases} 0 & \text{if } b \geq (D_Y + b_Y) \\ 1 & \text{if } b \leq D_Y \\ \frac{D_X + b_X - b}{b_X} & \text{if } b \in (D_Y, D_Y + b_Y) \end{cases} \quad (6)$$

The necessity expression of (4) for triangular membership functions is given by:

$$\inf_{a < w} 1 - \pi_X(w) = \begin{cases} 0 & \text{if } a \leq D_X \\ 1 & \text{if } a \geq (D_X + b_X) \\ \frac{a}{b_X} - \frac{D_X}{b_X} & \text{if } a \in (D_X, D_X + b_X) \end{cases} \quad (7)$$

$$\inf_{b > w} 1 - \pi_Y(w) = \begin{cases} 0 & \text{if } b \geq D_Y \\ 1 & \text{if } b \leq (D_Y - a_Y) \\ \frac{D_Y - b}{a_X} & \text{if } b \in (D_Y, D_Y + b_Y) \end{cases} \quad (8)$$

Note that with this approach, both possibility and necessity measures are obtained. When querying a temporal database, these two measures may be used to compute a ranking. E.g. the results for a query may be ordered by a lexicographical order (Pos, Nec). Note also that with only the possibility measure, as provided in the transformations, some interesting information is lost: e.g. If the user wants to know if the interval $[a,b]$ is inside the ill-known interval $[X,Y]$, and the evaluation returns a possibility of 1 and a necessity of 0.5. Then, the user may interpret that the interval $[a,b]$ is inside $[X,Y]$ but it is not between the cores of both X and Y.

Example Consider the ill-known values $X = [3,2,1]$ and $Y = [7,2,3]$. The knowledge about the evaluation of the interval $[a,b]$ w.r.t. X and Y is given by the expressions (3),(4). Figure 1 shows the graphical representation for the ill-known values X and Y. The representation for equations (5),(6),(7),(8) are shown respectively in figures 2,3,4,5.

Consider now that the value for the interval $[a,b]$ is $[2,6]$. The evaluation for the possibility and necessity are:

$$\text{Pos}\,(\lambda([3,6])) = \min(1,1) = 1$$
$$\text{Nec}\,(\lambda([3,6])) = \min(1,0.5) = 0.5$$

The main difference between this proposal and the transformations is that this framework always allows to obtain both possibility and necessity measures for each temporal comparison. In this case, there are the following possibilities:

- $\text{Pos}(\lambda([a,b])) = 0 \rightarrow \text{Nec}(\lambda([a,b])) = 0$: The interval $[a,b]$ is not contained the fuzzy interval $[X,Y]$.
- $\text{Pos}(\lambda([a,b])) > 0$, $\text{Nec}(\lambda([a,b])) = 0$: The interval $[a,b]$ is contained the fuzzy interval $[X,Y]$ but it is not inside the cores of both X and Y.
- $\text{Pos}(\lambda([a,b])) = 1$, $\text{Nec}(\lambda([a,b])) = 1$: The interval $[a,b]$ is contained the fuzzy interval $[X,Y]$ and it is inside the cores of both X and Y.

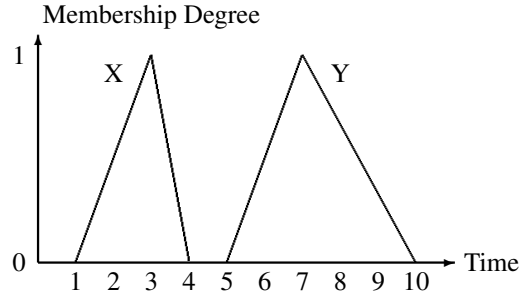


Figure 1: Triangular membership functions for the two ill-known time points X and Y.

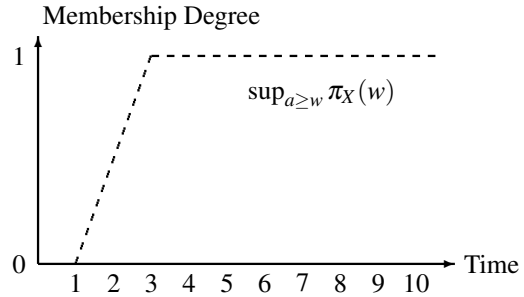


Figure 2: Membership function for $\sup_{a \geq w} \pi_X(w)$.

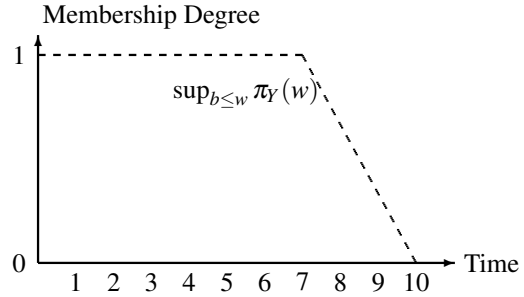


Figure 3: Membership function for $\sup_{b \leq w} \pi_Y(w)$.

4 ANALYSIS OF PROPOSED TRANSFORMATIONS

There are several proposals to transform two fuzzy numbers that represent an ill-known interval into one fuzzy interval. This reasoning is however not consistent with possibility theory. We provide several arguments to prove this,

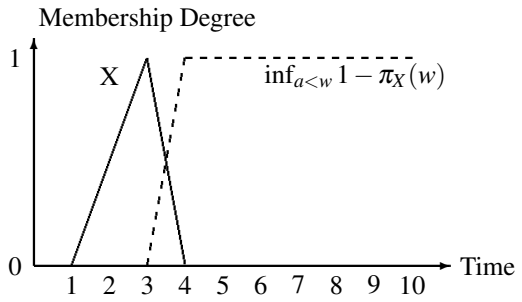


Figure 4: Membership function for $\inf_{a<w} 1-\pi_X(w)$.

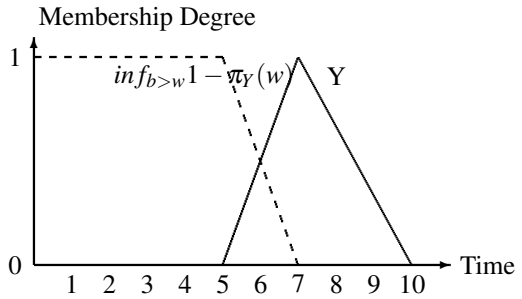


Figure 5: Membership function for $\inf_{b>w} 1-\pi_Y(w)$.

which have been previously reasoned by Dubois and Prade [4]:

1. The fuzzy set given by the two fuzzy numbers is conceptually different from a fuzzy set represented by the fuzzy interval. The fuzzy set given by the two fuzzy numbers belong to $\mathscr{P}(\mathbb{R})$ while the fuzzy interval obtained after the transformation is not a possibility distribution on $\mathscr{P}(\mathbb{R})$ but on $\mathbb{R}$.
2. In some cases, (like in the convex hull transformation) because of the underlying ordered domain, the membership function provided for the fuzzy interval preserves the possibility. However, the necessity is not preserved at all.

In the next subsections we will analyze the two major approaches: The transformation that preserves the imprecision and the convex hull.

4.1 Transformation that preserves the imprecision

This transformation has been proposed in [9]. They assume that the imprecision of a fuzzy set is its area. Therefore the transformation consists of preserving the area of both fuzzy numbers that compose the interval. The transformation is performed by means of geometrical computations. The starting point of the fuzzy temporal interval is given by the fuzzy number $S = [d_s, a_s, b_s]$ respectively, the ending point of the fuzzy interval is given by $E = [d_e, a_e, b_e]$ (see Figure 6).

The transformation defines a fuzzy interval T by means of a trapezoidal membership function. In order to define the

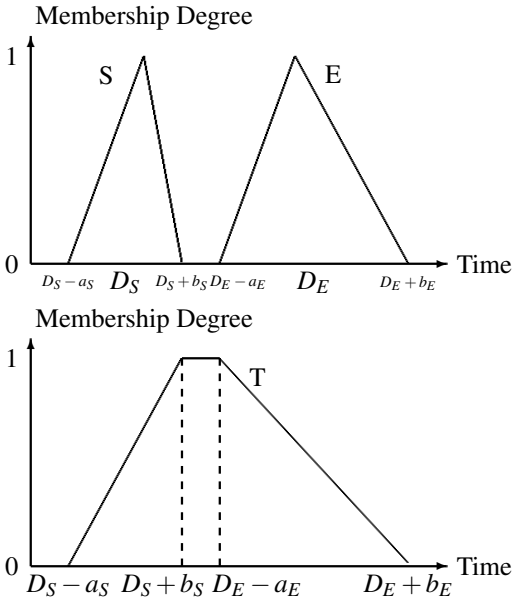


Figure 6: Transformation of two fuzzy numbers X and Y into one fuzzy interval T preserving the amount of imprecision.

membership function for the fuzzy interval, the following two equalities between the areas from the fuzzy numbers and the fuzzy interval are forced: $S1 = S2$ and $S3 = S4$. The calculation for the first equality, based on triangle equivalence is:

$$S1 = S2 \Rightarrow \frac{(d_s + b_s) - (d_s - a_s)}{2} \tag{9}$$

The calculation for the right side ($S3 = S4$) is analogous. This transformation is specially widely used when transforming fuzzy time points into fuzzy intervals. The reasoning is the following: the core of the transformed fuzzy interval corresponds with the time when the fact happened. The left and right slopes of the possibility distribution show the imprecision around the starting and the ending times.

The membership function for this fuzzy interval T is represented by the trapezoid $[\alpha, \beta, \gamma, \delta]$ and is equal to $[d_s - a_s, d_s + b_s, d_e - a_e, d_e + b_e]$:

Note that when the supports of the two fuzzy numbers overlap, this transformation makes no sense.

Example Consider the previous two ill-known points representing a time interval: $X = [3,2,1]$ and $Y = [7,2,3]$. P_1 is the value for the transformation T, which is $[1,4,5,10]$ (see figure 6. The value for $[a,b]$ is $[3,6]$ therefore, the value for P_2 is $[3,3,6,6]$. The knowledge about the interval $[a,b]$ w.r.t. the transformation T using the preserving the imprecision approach is given by the *CONTAINS* operator [9].

$$\text{CONTAINS}(P_1,P_2) = \min \sup_x (1-\mu_{P_2}(x), \mu_{P_1}(x)) =$$
$$\min(\mu_{\sup_x}(3), \mu_{\sup_x}(6) =$$
$$\min(\frac{2}{3}, \frac{4}{5}) = \frac{2}{3} = 0.667$$

Figure 7 shows the two membership degrees for both $(1-\mu_{P_2}(x))$ and $(\mu_{P_1}(x))$.

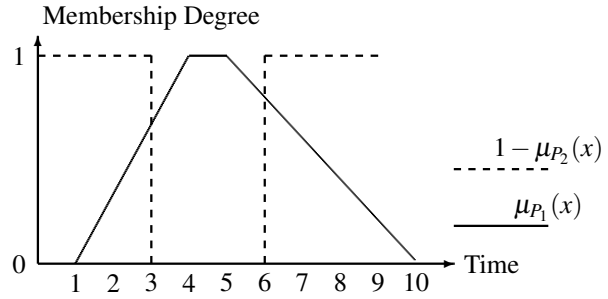


Figure 7: Graphical representation for the membership functions $(1-\mu_{P_2}(x))$ and $(\mu_{P_1}(x))$.

In this example it is shown how this transformation is misleading: it is clear that the temporal interval $[3,6]$ is completely contained by the fuzzy temporal interval represented by $[X,Y]$. Nevertheless, the *CONTAINS* operator applied to this transformation returns a 0.667 membership degree. Therefore, if the user is looking for a temporal interval $[a,b]$ and the database contains a tuple with a fuzzy interval $[X,Y]$, this transformation returns a possibility below 1, thus the user may interpret that the temporal interval $[a,b]$ is not completely inside the fuzzy interval $[X,Y]$. Plus, it is not possible to compute a necessity measure for the *CONTAINS* operator in order to obtain more information about $[a,b]$.

4.2 Transformation based on the convex hull

The transformation based on the convex hull is more intuitive than the previous one. The transformation builds a trapezoidal possibility distribution with the left slope equal to the left slope of the starting fuzzy number (in the graphic, from $d_s - a_s$ to d_s) , the core equal to the interval $[d_s, d_e]$ and the right slope equal to the right slope of the ending fuzzy number (from d_e to $d_e + b_e$). The interpretation of this convex hull is the "most possible" starting point and the "most possible" ending point. Figure 9 shows the resulting transformation.

In this approach, the trapezoidal membership function ($[\alpha, \beta, \gamma, \delta]$) is built in a different way: $[d_s - a_s, d_s, d_e, d_e + b_e]$

Example Consider again the previous two ill-known points representing a time interval: $X = [3,2,1]$ and $Y = [7,2,3]$. P_1 is the value for the transformation T, which is $[1,3,7,10]$ The value for $[a,b]$ is $[3,6]$ therefore, the value for P_2 is $[3,3,6,6]$. The evaluation for the *CONTAINS* operator is the following:

$$\text{CONTAINS}(P_1,P_2) = \min \sup_x (1-\mu_{P_2}(x), \mu_{P_1}(x)) =$$
$$\min \sup_x (1,1) = 1$$

Figure 8 shows the two membership degrees for both $(1-\mu_{P_2}(x))$ and $(\mu_{P_1}(x))$.

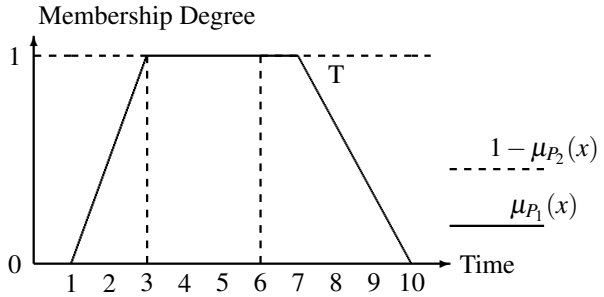


Figure 8: Graphical representation for the membership functions $(1-\mu_{P_2}(x))$ and $(\mu_{P_1}(x))$.

This transformation is also misleading. Even though the possibility obtained is similar to the possibility obtained in our proposal, it is not possible to compute the necessity for the *CONTAINS* operator. Therefore, it is not possible to obtain a complete information for the interval $[a,b]$ w.r.t the fuzzy interval $[X,Y]$.

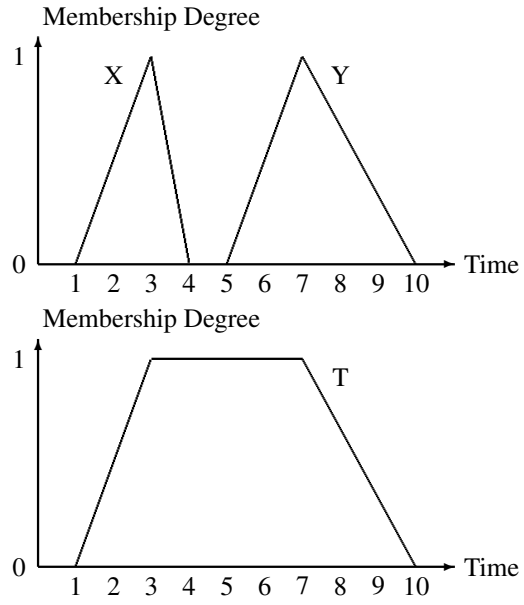


Figure 9: Construction of the convex hull T w.r.t the ill-known values X and Y.

5 CONCLUSIONS AND FUTURE WORK

In this work we have presented a framework to deal with the evaluation for the ill-known interval $[X,Y]$ to contain the crisp interval $[a,b]$. Our proposal keeps the reasoning in the power set $\mathscr{P}(U)$ while the transformations make the reasoning in the set U, therefore the transformations result in a possibilistic information lost. The proposed framework offer more accurate results, as shown in the examples. Also the two measures provided: possibility and necessity can be used for the ranking of results.The framework will be extended with the implementation of the Allen's operators and the comparison between two ill-known intervals.

Acknowledgements

Part of this research is supported by the grant BES-2009-013805 within the research project TIN2008-02066: *Fuzzy Temporal Information treatment in relational DBMS*.

References

[1] ALLEN, J. Maintaining knowledge about temporal intervals. *Communications of the ACM 26* (1983), 832–843.

[2] BOHLEN, M., BUSATTO, R., AND JENSEN, C. Point-versus interval-based temporal data models. In *Data Engineering, 1998. Proceedings., 14th International Conference on* (Feb. 1998), pp. 192 –200.

[3] BRIAN GAINES, AND LADISLAV KOHOUT. Possible automata. In *Proceedings of the International symposium on multiple-valued logic* (Bloomington, USA, 1975), pp. 183–196.

[4] DIDIER DUBOIS, AND HENRI PRADE. Incomplete conjunctive information. *Computers and Mathematics with Applications 15*, 10 (1988), 797–810.

[5] DIDIER DUBOIS, AND HENRI PRADE. The three semantics of fuzzy sets. *Fuzzy Sets and Systems 90*, 2 (1997), 141–150.

[6] DUBOIS, D., AND PRADE, H. Ranking fuzzy Numbers in the Setting of Possibility Theory. *Information Sciences 224* (1983), 183–224.

[7] DUBOIS, D., AND PRADE, H. Processing fuzzy temporal knowledge. *IEEE Transactions on Systems, Man, and Cybernetics 19* (1989), 729–744.

[8] DYRESON, C., AND GRANDI, F. E. A. A consensus glossary of temporal database concepts. *SIGMOD Rec. 23* (1994), 52–64.

[9] GARRIDO, C., MARÍN, N., AND PONS, O. Fuzzy Intervals To Represent Fuzzy Valid Time in a Temporal Relational Database. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 17*, Supp (2009), 173.

[10] JENSEN, C. S., MARK, L., AND ROUSSOPOULOS, N. Incremental implementation model for relational databases with transaction time. *IEEE Trans. on Knowl. and Data Eng. 3* (December 1991), 461–473.

[11] LOTFI ZADEH. Fuzzy sets. *Information and Control 8* (1965), 338–353.

[12] LOTFI ZADEH. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems 1* (1978), 3–28.

[13] L.ROWE, AND M.STONEBREKER". The postgres papers, 6 1987.

[14] MELTON, J., AND SIMON, A. R. *Understanding the new SQL: a complete guide*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

[15] NAGYPÁL, G., AND MOTIK, B. A fuzzy model for representing uncertain, subjective, and vague temporal knowledge in ontologies. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, vol. 2888 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2003, pp. 906–923.

[16] NASCIMENTO, M. A., AND EICH, M. H. Decision time in temporal databases. In *In Proceedings of the Second International Workshop on Temporal Representation and Reasoning* (1995), pp. 157–162.

[17] OHLBACH, H. J. Relations between fuzzy time intervals. *Temporal Representation and Reasoning, International Syposium on 0* (2004), 44–51.

[18] PONS, J. E., BRONSELAER, A., DE TRÉ, G., AND PONS, O. Possibilistic evaluation of sets. Submitted to International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2011.

[19] SCHOCKAERT, S., DE COCK, M., AND KERRE, E. Fuzzifying allen's temporal interval relations. *Fuzzy Systems, IEEE Transactions on 16*, 2 (april 2008), 517 –533.

[20] SHACKLE, G. *Decision, order and time in human affairs*. Cambridge University Press, 1961.

[21] SNODGRASS, R. The temporal query language tquel. In *Proceedings of the 3rd ACM SIGACT-SIGMOD symposium on Principles of database systems* (New York, NY, USA, 1984), PODS '84, ACM, pp. 204–213.

INCORPORACION DE UNA CAPA ELT PARA EL USO DE ETIQUETAS LINGUISTICAS EN UN DATA WAREHOUSE EN PENTAHO

ELT INCORPORATION OF A LAYER FOR USE OF LINGUISTIC LABELS IN A DATA WAREHOUSE IN Pentaho

Angélica Urrutia[1] y Javier Jirón[1]

[1]Universidad Católica del Maule, Departamento de Computación e Informática, Chile
Grupo de Bases de Datos (Tricahue) {jiron@ucm.cl aurrutia@ucm.cl}

RESUMEN

La presente investigación, es una propuesta del uso de etiquetas lingüísticas formuladas con conjuntos difusos asociados a datos de sistemas de gestión de la información. La contribución es la incorporación de una capa ETL, en una arquitectura de Data Warehouse (DW), que permita analizar la información de indicadores de gestión, clásicos o difusos, implementados en una herramienta Open Source (Pentaho).
Para ello, se analizan e implementan las capas de la arquitectura de un Data Warehouse, utilizando dos ETL; el primero para obtener la extracción de los indicadores de gestión desde una base de datos fuentes, y el segundo para incorporar etiquetas lingüísticas a los indicadores de gestión clásicos, para ambos casos, se analizan y comparan resultados. Finalmente, seleccionamos un caso de estudio implementado con una taxonomía para probar la arquitectura propuesta.

PALABRAS CLAVES: Indicadores de gestión con etiquetas lingüísticas, Data Warehouse, ETL.

1. INTRODUCCION

Los sistemas de inteligencia de negocios Business Intelligence (BI), tienen como elemento clave de su arquitectura a los almacenes de datos o Data Warehouse (DW) y algoritmos de minería de datos Data Mining (DM). Para iniciar un proceso BI, los datos son extraídos desde base de datos fuentes utilizando herramientas ETL *(*Extract, Transform and Load), que permite Extraer**,** Transformar y cargar datos creando una estructura multidimensional, que es almacenada en un DW, un ejemplo de ello se presenta en [6].

Un DW es una base de datos histórica, orientada a temas y satisfacer indicadores de gestión considerando un Proceso Analítico en Línea (OLAP), el cual se implementa sobre la base de alguna de las siguientes alternativas: MOLAP (OLAP sobre una base de datos multidimensional), ROLAP (OLAP sobre una base de datos relacional) o HOLAP que combina ambos enfoques [4]. Algunas propuestas en este tipo de plataformas utilizando vistas materializadas y modelos multidimensionales se encuentran en [5] [1].

Nuestro estudio considera la flexibilización de consultas de una implementación BI a partir de la extracción de indicadores de gestión a un DW. Dicha extensión utiliza conjuntos difusos, en forma especial etiquetas lingüísticas aplicadas a algunos datos de análisis [9] [8]. Una representación de los conjuntos difusos es la función trapezoidal **T(x),** la que se muestra en la Figura 1.

a)

$$T(x)=\begin{cases} 0 & \text{Si } (x \le a) \text{ o } (x \ge d) \\ (x-a)/(b-a) & \text{Si } x \in (a, b) \\ 1 & \text{Si } x \in (b, c) \\ (d-x)/(d-c) & \text{Si } x \in (b, d) \end{cases}$$

b)

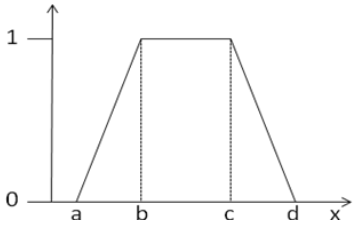


Figura 1: Representación de la función trapezoidal a) matemática b) gráfica.

Los conjuntos difusos son una extensión de los conjuntos clásicos, pues permiten nociones de acercamiento a lo verdadero, falso o imprecisos, relajando la restricción de los conjuntos clásicos de pertenencia o no-pertenencia absoluta al mismo. Un conjunto difuso *A* sobre un universo de discurso *U* es un conjunto de pares dado por: $A = \mu_A(u)/u : u \in a\ U$, $\mu_A(u)$ pertenece a $[0,1] \in R\}$
Donde, μ es llamada función característica o de pertenencia y $\mu_A(u)$ es el grado de Pertenencia del elemento *u* al conjunto difuso A. Este grado de pertenencia oscila entre los extremos 0 y 1 [2], donde:
$\mu_A(u) = 0$, indica que *u* no pertenece en absoluto al conjunto difuso *A*,

μ_A (u) = 1, indica que u pertenece totalmente al conjunto difuso A.
Un ejemplo del uso de la lógica difusa aplicada a la cubos multidimencionales de encuentra en [7], aquí se muestra como se flexibilizan los datos asociados a las dimensiones con etiquetas lingüísticas con algunos tratamientos del su uso en Roll Up.

Las contribuciones esperadas de este trabajo son: implementación de un DW en una plataforma Open Source con dos ETL y comparación de los resultados entregados por datos clásicos y difusos para el análisis de información de indicadores de gestión. Los apartados para tratar este tema son: Arquitectura DW extendida para etiquetas lingüísticas y taxonomía de construcción del Data Warehouse Difuso, los que son presentados a continuación.

2. ARQUITECTURA DW EXTENDIDA PARA ETIQUETAS LINGÜÍSTICAS

El desarrollo tecnológico, es fuente de desafíos y oportunidades para mejorar herramientas que se disponen en el mercado, en el ámbito de la Business Intelligence asociados a un Data Warehouse, sobre todo en una extensión con lógica difusa. Existen diversas propuestas de arquitectura para implementar soluciones DW [4] [3]. La Figura 3, muestra nuestra propuesta que se centra especialmente en el proceso ETL. La Capa 2 obtiene la extracción de indicadores de gestión clásicos, que los almacena en un diseño de estrella (MOLAP), y la Capa 5 selecciona y extienden por medio de difusificación o transformación a etiquetas lingüísticas.
La definición asociada al ETL incorporado a la arquitectura propuesta es la siguiente:

Definición 1: Sea b una colección de objetos extraídos de la capa 1 y W_D es la colección de objetos correspondientes al Data Warehouse de la capa 3, donde $b \epsilon W_D$, y $\mu_E(d_i)$ la función de transformación y carga de la capa 2.

$$\mu_E(d_i) \rightarrow W_D \quad (1)$$

Definición 2: Sea W_D definida para aplicar una función trapezoidal que implica T(X) $\mu_F(W_D)$ la pertenencia al conjunto difuso definida como un set difuso para los pares ordenados:

$$A = \{(x, \mu_F(w_D)) | w_D \epsilon W_D\} \quad (2)$$

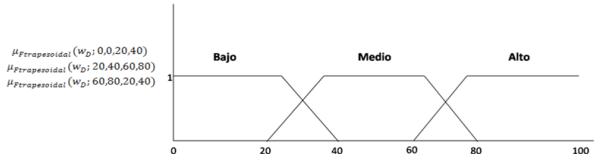


Figura 2: Ejemplo de la representación de etiquetas lingüísticas para la Capa 5.

El detalle de cada una de las capas, se especifica a continuación:

Capa 1: Esta es la capa más baja, en ella se encuentra la base de datos operacional, que se utiliza como bases de datos fuente.
Capa 2: En esta capa se utiliza el ETL para la extracción y transformación de los datos que se diseñan a partir de los indicadores de gestión.
Capa 3: En esta capa se crea el Data Warehouse, usando un diseño estrella u copo de nieve.
Capa 4: En esta capa se crea un cubo multidimensional utilizando herramientas Open Source. En esta capa es posible realizar consultas clásicas al cubo multidimensional.
Capa 5: En esta capa se realiza la extensión de la arquitectura por medio de un ETL, que trasforma datos del cubo a etiquetas lingüísticas con sus respectivas consultas.
Capa 6: En esta capa se utiliza una herramienta Front End para la visualización de los datos.

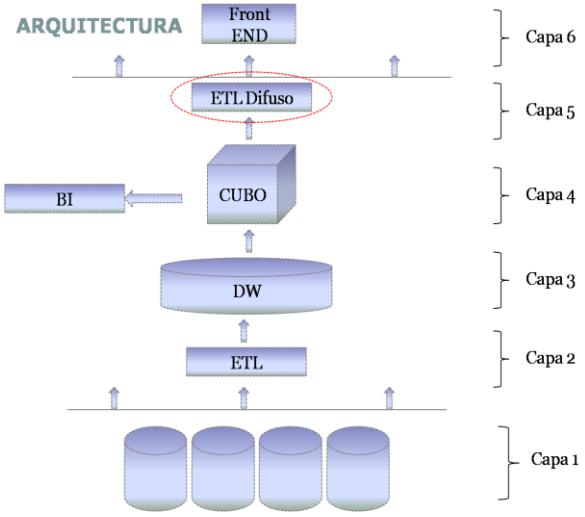


Figura 3: Arquitectura Propuesta para un DW con ETL clásico y difuso.

En la implementación de cada capa, se utilizan herramientas Open Source, tales como, ETL Keller, la generación de cubos con Mondrian, como partes de Pentaho, para una arquitectura Business Intelligence. En su implementación se obtienen resultados idénticos a los obtenidos con herramientas privativas (SQL SERVER 2008).

3. TAXONOMÍA DE CONSTRUCCIÓN DEL DATAWAREHOUSE DIFUSO

La construcción de un Data Warehouse Difuso (DWF) se desarrolla con once actividades. Estas actividades las llamaremos taxonomía y permiten el desarrollo e implementación del DWD, asociados a tres fases fundamentales: la primera especificación de requerimientos (pasos 1-5), la segunda etapa corresponde a la implementación de cubos precisos o cuantitativos en un Data Warehouse (pasos 6-7) y la tercera etapa corresponde a la definición de los componentes imprecisos, con etiquetas lingüísticas que

extenderán algunos datos del data Warehouse a datos cualitativos o datos difusos (pasos 8-11).
Cada paso de la taxonomía se explica en el siguiente apartado.

3.1. APLICACIÓN DE LA PROPUESTA A UN CASO PRÁCTICO.

En este apartado se procede a implementar la propuesta tanto de arquitectura como de la taxonomía. La fuente de datos que se utiliza es la base de datos *AdventureWorks*, que es una base de datos de prueba que incorpora SQL Server 2008, como caso es la venta de bicicletas por internet. A continuación mostramos cada paso de la taxonomía propuesta.

Etapa: Especificación de Requerimiento.

Paso 1: Análisis de requerimientos de información.
Para este caso se propone el análisis de información que solicita un Gerente de Ventas, quien desea conocer la información de las ventas de bicicletas durante cierto período. Las posibles preguntas que se podría plantear el gerente serían las siguientes:
¿Cuánto fue las ventas, totales de bicicletas vendidas en un tiempo determinado y que cliente fue el comprador?
¿Compra total realizada por un determinado cliente?
¿Qué profesión tienen los clientes que compran más bicicletas durante un tiempo determinado?
¿Qué nivel de estudio tienen estos clientes?
Con estas preguntas, se establecen las necesidades de información que requiere el usuario, para poder definir sus objetivos y planificar de manera estratégica, cuáles serán los pasos a seguir por la empresa y cuáles son las medidas que hay que tomar para mejorar el rendimiento del departamento de ventas.

Paso 2: Identificaciones de indicadores y perspectivas
Una vez recopilada las preguntas donde se identifican las necesidades del negocio, se procede a descomponerlas para descubrir los indicadores y las perspectivas. La descomposición que se realizó a las preguntas del usuario, se muestran el la Figura 4.

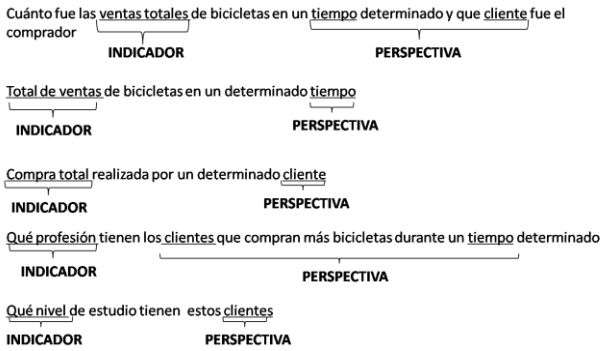


Figura 4: Descomposición de indicadores y perspectivas.

Paso 3: Creación de modelo conceptual
Obtenidos e identificados los indicadores y las perspectivas, se procede a la creación de un modelo conceptual. La Figura 5, muestra dos dimensiones o perspectivas: *Clientes* y *Tiempo*. Los indicadores son: *ventas totales* y *nivel de estudio*. La relación que representa este modelo es *Ventas por Internet*.

Figura 5: Modelo conceptual para el caso de estudio.

Paso 4: Determinar base de datos Fuentes.
La base de datos que se utiliza es AdventureWorks, que es una base de datos de prueba para ventas de bicicletas.

Paso 5: Correspondencias entre modelo conceptual y base de datos fuentes
Una vez conocida la fuente de datos, es necesario establecer la correspondencia entre el modelo conceptual de la Figura 5 y la base de datos fuentes OLTP.
Luego de definir la relación entre el modelo lógico y la base de datos fuente, se consideran los atributos que se encuentran en las tablas de base de datos relacionales, para dicho modelo. En el caso de estudio se toman todos los datos que son de interés para las diferentes perspectivas, por ejemplo:
Cliente: ID_Cliente,ID, Geografia, Titulo, Primer nombre, Segundo nombre, Nombre del estilo, Fecha de nacimiento, Estado civil, Sufijo, Genero, email, Ingresos anuales, Número de hijos, Educación, Profesión, Indicador si es propietario, Número de autos del dueño de casa, Dirección 1, Dirección 2, Teléfono, Primera fecha de compra, Distancia de viaje.
Fecha: ID_Tiempo, Fecha, Nombre del día, Número del día de la semana, Número del día del año, Número de la semana del año, Nombre del mes, Número del mes del año, Año del calendario, Semestre del calendario, Trimestre del calendario, Trimestre fiscal, Año fiscal, Semestre fiscal.

Etapa: Implementación de Cubos Precisos.

Paso 6: Creaciones del Data Warehouse.
El tipo de diseño que se utiliza en el caso de estudio es un diseño de tipo estrella, en donde se generan una tabla de hechos central acompañado de dos tablas dimensiones.
Perspectiva Tiempo:
- Nombre para la dimensión: DimDate
- Clave principal: Datekey.
- Atributos: atributos opcionales.

Perspectiva Cliente:
- Nombre para la dimensión: DimCustomer
- Clave principal: Customer Key.
- Atributos: atributos opcionales.

Tabla de Hecho: La tabla de hechos es *ventas por internet*, la cual contiene el indicador de gestión "*ventas totales*". Sus clave primaria, corresponden a las claves de las perspectivas o dimensiones y su *diseño para este proceso es un diagrama estrella, asociada a las tablas de dimensión* cliente *y* tiempo.

Paso 7: Proceso ETL (extracción, transformación y carga).
Actividad encargada de extraer desde la base de datos fuentes y llenar la tabla de hecho y tablas dimensiones, con todos los datos necesarios.

Etapa: Etiquetas lingüísticas en DW.

Paso 8: Creación de indicador de gestión (Cubo).
El indicador de gestión que se seleccionó, es la suma de ventas por Internet, para lo cual se implementó un cubo, el que se muestra en la Figura 6.

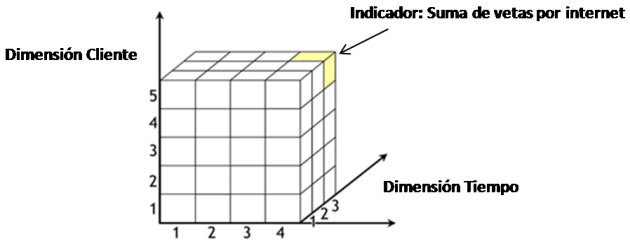


Figura 6: Visualización de indicador de gestión gráficamente.

Para la implementación del indicador de gestión, se analizaron dos herramientas de la suite de Pentaho Open Source.
La primera de ellas fue Cube Designer, la cual nos permite la creación de cubos de manera muy fácil, ya que cuenta con un entorno gráfico que permite un trabajo muy sencillo con las dimensiones y las tablas como se muestra en Figura 7. Este editor genera cubos en formato XML, no es un cubo bien especificado, por lo que, muestra falencias en su creación.
La segunda es el editor para cubos Schema-Workbench, la desventaja es no tener un entorno gráfico intuitivo, por lo que es necesario tener conocimiento sobre la creación de cubos antes de proceder a la creación del mismo.

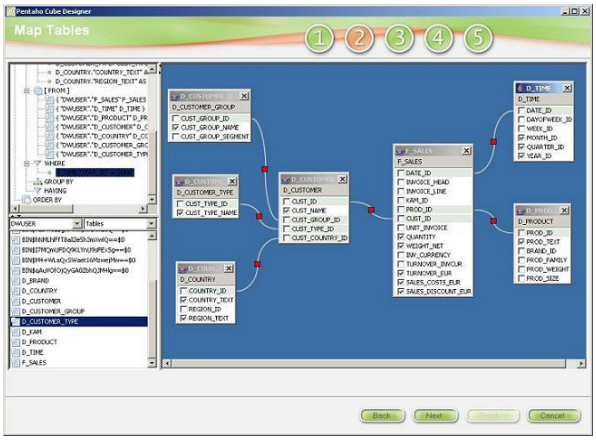


Figura 7: Imagen de Cube Designer que representa un diseño estrella y copo de nieve.

Paso 9: Visualización de indicador de gestión.
En este paso se selecciona y utiliza la herramienta Biserver-ce de la suite Pentaho, que permite visualizar el indicador de gestión, primero es necesario especificar la conexión a la fuentes de datos antes de visualización del cubo, ya que se debe especificar donde se encuentran las tablas de dimensiones y de hechos. Biserver-ce cuenta con administration-console el cual realiza la conexión con las fuentes de datos.

Paso 10: Diseño de calificadores difusos.
En [7], se muestra el uso de ETL solo para la extracción, nuestra propuesta genera un ETL en el sistema OLAP.
La definición de las etiquetas lingüísticas asociadas para ventas realizadas por internet fueron: Ventas Bajas, Medias y Alto. La Figura 8, muestra los intervalos a representar en tres trapecios con un referencial ordenado entre [0 100] dado por:

Bajo: comprende un intervalo [0,0,20,40].
Medio: comprende el intervalo [20,40, 60,80].
Alto: comprende el intervalo [60, 80,100, 100].

Con el siguiente código:

```
var Malo = membership (Suma.getNumber(),u,0,0,20,40);
var Medio = membership (Suma.getNumber(),u,20,40,60,80);
var Bueno = membership (Suma.getNumber(),u,60,80,100,100);
```

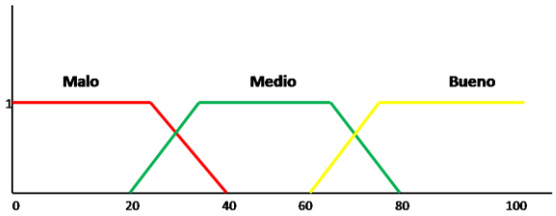


Figura 8: Representación de etiquetas lingüísticas en conjunto difuso.

Una observación importante que se debe realizar es que los valores que son ingresados, a la función, deben ser normalizados para que sus valores se encuentren en el intervalo [0,100]. Para realizar esta transformación, se considera los valores ingresados y se dividen por el mayor valor y luego se multiplica por 100. Luego de definir la función de membresía se debe definir las etiquetas lingüísticas que genera cada conjunto difuso.

```
//Funcion de membresia

var u = Mayo valor;
var salida= "";

function membership (x, u, x1, x2, x3, x4)
{
    x=(x/u)*100

    if(x1<x && x<x2)
        return((x-x1)/(x2-x1));
    else if( x2 <= x && x <= x3)
        return (1);
    else if(x3<x && x<x4)
        return((x4-x)/(x4-x3));
    else return(0);
}
```

La función de membresía para la implementación de los calificadores difusos para extender el cubo, se desarrollo en sentencia MDX, el código generado para este caso se muestra en el párrafo anterior.

Paso 11: Implementación de particiones difusas (ETL Difuso).

En la implementación se selecciona la herramienta Spoon de la suite de Pentaho como ETL, el que tiene un proceso asociado a la transformación, de los datos cuantitativos (datos numéricos) del cubo a datos cualitativos (datos con etiquetas lingüísticas).

La entrada del ETL es el cubo del Paso 8 ya implementado (se utiliza el Lenguaje MDX). Los valores extraídos son guardados en una Tabla. La Figura 9 muestra los pasos de este ETL.

Después de este proceso ETL se procede a transformación de difusificación los datos, para esto será necesario implementar la función de membresía creada en el Paso 10. Esta implementación fue realizada en el lenguaje JAVA SCRIPT y su estructura fue la siguiente sentencia en MDX.

```
//Script here

var u = 9907977.2811;
var salida= "";

function membership (x, u, x1, x2, x3, x4)
{
    x=(x/u)*100

    if(x1<x && x<x2)
        return((x-x1)/(x2-x1));
    else if( x2 <= x && x <= x3)
        return (1);
    else if(x3<x && x<x4)
        return((x4-x)/(x4-x3));
    else return(0);
}
```

Los valores que recibe la función corresponden a:

U: el máximo valor de las ventas por internet.

X: el valor que se quiere fuzzyficar

X1, X2, X3, X4: valores del trapecio a formar. Con esta operación se obtiene el grado de pertenencia que tiene cada elemento a cada una de las etiquetas lingüísticas definidas en el Paso 10. La Figura 10 muestra el proceso siguiente del ETL de difusificación.

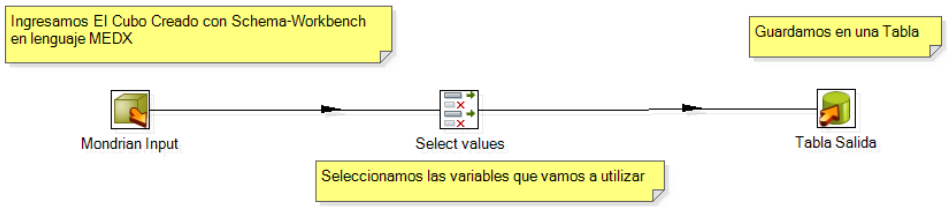


Figura 9: Ingreso del Cubo en lenguaje MDX al ETL

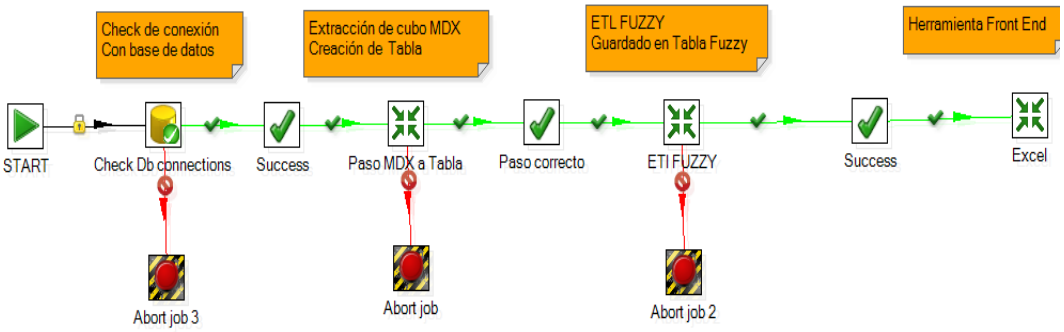


Figura 10: Visualización del ETL difuso.

Paso 12: Obtención de cubo difuso

Luego de tener el cubo difusificado como resultado del ETL, se procede a utilizar una Herramientas Front End, la seleccionada para este caso Open Source, fue OpenOffice, la que genera un archivo con formato XML.

Al finalizar este paso se puede apreciar la gran diferencia que existe al momento de realizar un análisis sobre datos que son cualitativos y cuantitativos, ya que el analista tienen información más generalizada para identificar si una venta en buena, mala o regular (datos difusos) y a la vez más detallas (datos clásicos).

Para la verificación de los resultados se pudo comparar con algunas herramientas propietarias como SQL Server y los datos obtenidos son exactamente los mismo, por lo que se demuestra que utilizar herramientas propietarias o herramienta privativa se los mismos resultados en un BI. Las herramientas BI Free requieren de mayor experiencia del uso de las mismas.

5. CONCLUSIONES Y TRABAJOS FUTUROS

En la actualidad dentro de las organizaciones el poder de la información es de vital importancia para la toma de decisiones, para esto sean utilizado nuevas

tecnologías que permitan un manejo más eficiente y eficaz de la información.

Las organizaciones necesitan tener conocimiento sobre sus indicadores de gestión y como éstos se comportan en el tiempo, ya que así pueden realizar una planificación estratégica en donde identificar los mejores mercados para sus productos o sus poténciales clientes.

La introducción de lógica difusa utilizada aquí, extendiendo la arquitectura clásica Business Intelligence, permitió la creación de indicadores de gestión no sólo cuantitativos, sino que cualitativos asociados a etiquetas lingüísticas, como por ejemplo: mostrar las ventas que son "buenas" asociadas a un conjunto de datos.

La transformación de datos cuantitativos a datos cualitativos debe realizarse siempre en conjunto con un experto, que luego de analizar sus indicadores de gestión, determina que ventas son "buenas", "malas" o "regulares". Con la extensión de la arquitectura Business Intelligence se logra que los indicadores de gestión, además de indicar su valor cuantitativo que se asocia a un valor cualitativo, por lo que el analista identifica con mayor claridad los niveles de venta, aportando de manera significativa al proceso de análisis de indicadores de gestión.

En la construcción de las etiqueta lingüísticas se debe considerar algunas experiencias del experto o usuario final para definirlas y estarán de acorde al contexto.

Se apreció que la implementación de las etiquetas lingüísticas debe partir por un buen diseño del conjunto difuso, que represente el contexto en el que se quiere aplicar. Una mala elección una etiquetas lingüísticas traerá como consecuencia que la representación de los valores cualitativos no tiene relación con los valores cuantitativos. El buen diseño se obtendrá a partir de las especificaciones de requerimiento analizadas en conjunto con el usuario.

La razón por la que se trabajó con lógica difusa dentro de los procesos Business Intelligence, es que la necesidad de manejar información estratégica de la organización de manera más natural y fácil de comprender para los niveles gerenciales, abstraer los complejos resultados a una forma natural al raciocinio humano.

Trabajos futuros. La utilización de otras herramientas de la lógica difusa en los DW, DM y soluciones Front End.

Referencias

[1] Carrera Sabina., Varas Marcela., Urrutia Angelica. (2010) **"**Transformación De Esquemas Multidimensionales Difusos Desde El Nivel Conceptual Al Nivel Lógico", Ingeniare. Revista Chilena. Ingeniería. Vol. 18 N º 2, pp. 165-175.

[2] Galindo, J., Urrutia, A., Piattini, M. (2006), *Fuzzy Databases: Modeling, Design and Implementation*. Idea Group Publishing Hershey, USA.

[3] Inmon, W.H, "Building The Data Warehouse". (2002). Tercera Edición, John Wiley & Sons, 2002.

[4] Kimball Ralph, "The Data Warehouse Toolkit. (1996), Primera Edición, John Wiley & Sons.

[5] Kumar Pavan, Krishua Radha, Kumar Supriya. (2005), "Fuzzy OLAP Cube for Qualitative Analisis", IEEE, Vol 15, pp 290-295.

[6] Lobo Villar Julio. (2010). "Transformación de una datawarehouse de un modelo MDA en un diseño lógico de cubo en XML, de la herramienta Pentaho Businness Inteligence" Trabajo de Titulación, Universidad Católica del Maule.

[7] Molina, C., Rodríguez-Ariza, L., Sánchez, D., Vila, M.A. (2006). "A new fuzzy multidimensional model". IEEE Transactions on Fuzzy Systems 14 (6), pp. 897-912.

[8] Rodríguez Rosseline, Tineo Leonid**,** Urrutia Angélica. (2010). "*Fuzzy Qualified OLAP CUBE*", Revista Tekhné, Indexada en el LATINDEX CATÁLOGO. Número 13, pp 12-30. Venezuela.

[9] Urrutia Angélica, Egaña Diego. (2007). "*Extensión de un Data Warehouse usando etiquetas lingüísticas y operadores CUBE*", Revista de Tecnología Informativa, Volumen y número 16. http://revistas.uis.edu.co/index.php/revistagti/article/view/1242. [8]

Implementación de una máquina HTM

Adolfo R. de Soto, **Conrado Andreu Capdevila**

Escuela de Ingenierías Industrial e Informática
Universidad de León,
adolfo.rdesoto@unileon.es, conrado.andreu@gmail.com

Resumen

Se presenta una implementación de un máquina que persigue modelizar algunos aspectos básicos de la biología del cerebro. Tres son los elementos fundamentales sobre los que se asienta el modelo: representaciones sparse, jerarquía y predicción temporal. El trabajo se basa en la hipótesis de que modelos computaciones pueden ayudar a comprender la complejidad de la inteligencia y de los mecanismos de cognición humanos, así como a la obtención de máquinas de exploración de datos de propósito general. Se busca especialmente un modelo que pueda explicar los procesos de abstracción realizados por la mente humana, considerados básicos a la hora de explicar los mecanismos cognitivos.

Palabras Clave: HTM, Abstracción.

1 INTRODUCCIÓN

Los modelos computacionales basados en la biología del cerebro son tan antiguos como la propia Inteligencia Artificial y se remontan a la definición de neurona artificial dada por McCulloch y Pitts [9]. Y aunque se dista mucho de comprender los procesos del cerebro que dan lugar a la mente humana, el aumento en el conocimiento de la fisiología del cerebro y la evolución técnica permiten plantear modelos computacionales con base biológica cada vez más sofisticados. Desde el punto de vista de los autores, la construcción de estos modelos deben perseguir un doble objetivo: ayudar a comprender los procesos del conocimiento humano y posibilitar la construcción de máquinas de conocimiento de propósito general. Especialmente estas máquinas deben ayudar en la exploración inteligente de datos desde un acercamiento lo más generalista posible, sin buscar soluciones construidas ad-hoc para el problema considerado.

En este trabajo se presenta una implementación de un modelo computacional basado en los trabajos de Jeff Hawkins [8, 11]. En sus trabajos Hawkins da especial relevancia a la disposición jerárquica, por niveles, que se presenta casi de forma generalizada en el conjunto de neuronas del cortex (especialmente visible en la zona V1 dedicada a la visión) y a la capacidad de predicción del cerebro. La disposición jerárquica permitiría al cerebro abordar problemas de una gran complejidad gracias a la "reducción" en los patrones reconocidos que se produciría en los distintos niveles jerárquicos. Los autores consideran que esta jerarquía presentaría como fenómeno emergente la capacidad de abstracción de la mente humana.

Es común definir inteligencia en el campo de la Inteligencia Artificial como la capacidad que permite a un agente actuar adecuadamente para lograr sus objetivos. Sin embargo en muchos actos humanos no está claro cuál es el objetivo y sin embargo actuamos racionalmente; por ejemplo contestar y mantener una conversación por teléfono no tiene por qué tener un objetivo concreto. Aunque también es cierto que en todo caso el cerebro construye un modelo de la situación que nos permite su comprensión. Por ello es plausible creer que la predicción es el objetivo global del cerebro. Anticipar situaciones y corroborarlas mediante pronósticos que se confirman y por tanto se refuerzan, o se niegan y por tanto se abandonan.

2 JERARQUÍA EN EL CORTEX

El modelo clásico del procesamiento visual en el cortex es un modelo jerárquico [14, 10]. Sin embargo y de acuerdo con [6] hay que ir hacia un paradigma de red distribuida para la memoria, donde la unidad de red de memoria o "cognit" jugaría un papel central. Según el autor "Los cognits están jerárquicamente organizados en términos de abstracción semántica y complejidad. Los cognits se superponen e interconectan intensivamente, incluso entre los niveles jerárquicos". Ese nuevo modelo llevaría a una organización "heterojerárquica" en la que la jerarquía aparece pero es superada con una organización basada en relaciones intra e

inter-niveles de la jerárquía. Sin embargo también según el autor la jerarquía sería necesaria puesto que permitiría representaciones de más altas y complejas memorias y conocimiento, desde la información sensorial en la parte inferior hasta las memorias conceptuales y semánticas en la superior. De acuerdo con este modelo, las distintas memorias así como los items de conocimiento consistirían en redes en el cortex organizadas de forma jerárquica y distribuida.

3 CODIFICACIÓN DISPERSA EN EL CORTEX

En general se entiende por codificación dispersa aquella codificación en la que la mayor parte de las unidades de codificación están inactivas la mayor parte del tiempo. Varios estudios (ver por ejemplo [7]) sugieren la posibilidad de que una codificación dispersa sea la adaptación que el cortex ha adoptado para ofrecer una estrategia para almacenar y construir asociaciones a partir del aparato sensorial. Las limitaciones en la energía disponible para los procesos metabólicos que presentan los sistemas neuronales biológicos es una importante razón que justifica la aparición de codificaciones dispersas. Además en varios sistemas neuronales biológicos han sido identificados comportamientos totalmente compatibles con codificaciones dispersas especialmente en el comportamiento de las neuronas V1 en el cerebro humano [12]. Otras ventajas de la codificación dispersa serían el incremento en la capacidad de almacenamiento de las memorias asociativas o facilitar la lectura de datos complejos, citados en [13].

4 PREDICCION EN EL CORTEX

Varios modelos teóricos establecen que el cerebro no es meramente reactivo a los estímulos sino también "proaactivo" o "predictivo". El cerebro humano está continuamente ocupado generando predicciones que aproximan el futuro que le interesa. Esas predicciones facilitan la percepción y la cognición [2, 3, 5]. En [1] la hipótesis de la predicción del cerebro en las áreas visuales es corroborada al comprobar que la activación neuronal es menor para aquellos estímulos predecibles frente a los impredecibles.

5 IMPLEMENTACIÓN

Una red HTM está compuesta de una o más regiones dispuestas en niveles jerárquicos. Cada nivel puede contener varias regiones y la salida de éstas será la entrada para las regiones del nivel superior. En el primer nivel se encuentra el procesamiento de la entrada a la red.

En la figura 1 se muestra una representación de una región. Se compone de un conjunto de células dispuestas en capas. Las células en la misma columna vertical comparten conexiones *bottom-up* con la entrada, de forma que resulta útil hablar de una capa más de elementos "columna" que va a alimentar a las capas de células. Aquí se ha representado como una capa más bajo las células.

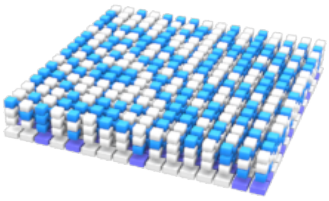


Figura 1: Composición de una región HTM.

Las columnas que están conectadas con la entrada implementan el llamado *spatial pooler*. Es la parte que se encarga de calcular una representación dispersa de lo que la región percibe en cada momento. Como se verá más adelante, el proceso consiste en calcular un estado de activación para cada columna. La matriz final de activaciones es la representación dispersa. Los patrones procesados aquí son exclusivamente espaciales.

El espacio de entrada propiamente dicho se presenta a la región de forma unificada como una disposición de bits, elementos que pueden estar activos o no y que ocupan una posición determinada. Esos bits no codifican necesariamente imágenes, pero en los ejemplos se utilizará ese tipo de entrada para facilitar la explicación.

Las células trabajan contra la representación dispersa para computar la parte del *temporal pooler*. Aquí se incorpora la dimensión temporal, de forma que las células establecen conexiones entre sí para aprender patrones temporales (secuencias de patrones espaciales).

La salida final de la región está compuesta por el agregado de los estados de activación y de predicción de las células. De nuevo se trata de una disposición de bits que es el tipo de entrada que puede alimentarse a una región de un nivel superior en la jerarquía.

El funcionamiento de cada región está gobernado por varios parámetros, que se discutirán en más detalle en los siguientes apartados. En la implementación realizada en este trabajo se ha incorporado la obtención de medidas de diferentes indicadores, como por ejemplo "proporción de células activas en predicción", "número total de sinápsis" o "número de actualizaciones de segmentos implementadas". Los algoritmos, tal y como están planteados en la actualidad, no incluyen elementos de control explícitos que modifiquen los parámetros en ejecución. Como posible trabajo futuro, pensamos que sería interesante estudiar un control *fuzzy* de la máquina HTM.

A continuación se describe de forma resumida el funcionamiento de una región y se ilustran los puntos más importantes con ejemplos. En primer lugar se trata el *spatial pooler*

que calcula la representación dispersa de la entrada, y a continuación el *temporal pooler* que utiliza esa representación para aprender patrones temporales. En [11] se puede consultar una descripción en detalle.

5.1 Spatial Pooler.

El *spatial pooler* calcula una representación dispersa de la entrada en el espacio de columnas. Cada columna mantiene un conjunto de sinápsis *bottom-up* enlazadas con varios bits de la entrada. Cada una de ellas tiene un valor de fuerza que va a determinar si se considera conectada o no. Al presentarse una entrada particular, la combinación del estado de los bits enlazados y de las fuerzas de las sinápsis produce un valor de solapamiento para cada columna. No todas ellas van a alcanzar un estado de activación y a pasar así a formar parte de la representación dispersa final, sino que antes deben superar un proceso de inhibición. Sólo aquellas con un valor de solapamiento alto comparado con los de sus vecinas resultarán ganadoras.

5.1.1 Parámetros.

En la construcción de la región se establecen las conexiones iniciales para cada columna, con una selección aleatoria de bits de entrada dentro de una zona centrada en la proyección de la columna sobre el espacio de entrada. La fuerza de cada conexión depende de la distancia entre esa proyección y el bit enlazado, con un valor máximo en el centro. Este proceso de inicialización está gobernado por varios parámetros como el número de sinápsis por columna, el tamaño de la zona de la entrada a considerar y la forma de la función de fuerza.

Con el objetivo de que las columnas aprendan a representar patrones no triviales, su solapamiento con la entrada debe alcanzar al menos un valor mínimo para ser considerado en el proceso de inhibición. Se trata de otro parámetro que actúa como un umbral de corte; si el solapamiento es menor que el umbral, se considera nulo.

También se desea que la región aproveche el número de columnas disponibles, de forma que todas aprendan a representar algún patrón. Para ello se controla la evolución de la actividad de cada columna respecto a la de sus vecinas y si no es suficientemente alta, se incrementa la agresividad de esa columna en el proceso de inhibición. Este control tiene asociados parámetros que determinan cuándo un nivel de actividad es bajo y cuánto se modifica el comportamiento de las columnas.

5.1.2 Implementación.

El algoritmo consta de tres fases: Solapamiento, actividad y aprendizaje. Cada una de ellas consiste en un cómputo que puede ser distribuído en paralelismo sobre el espacio de columnas, ya que ninguna sobreescribe sus datos de entrada. Por ejemplo, en la primera fase cada columna consulta varios elementos del espacio de entrada y sus sinápsis para generar su valor de solapamiento. En la siguiente fase, otra columna necesitará consultar valores de solapamiento de sus vecinas, pero en la implementación se incluye una barrera de sincronización después de cada fase, de forma que los resultados intermedios estén disponibles.

Las estructuras de datos utilizadas son en su gran mayoría arrrays multidimensionales de tipos básicos. Este diseño es adecuado para aplicar el paradigma propuesto por OpenMP [4], de forma que las fases de cómputo suelen consistir en un lazo “for” paralelo sobre el espacio de columnas.

Toda la infraestructura de entrada/salida se apoya en el formato de datos HDF5, tanto para conjuntos de datos de entrada como para almacenar el estado de la máquina en un momento determinado.

5.1.3 Ejemplo.

Dada una región de 64×64 columnas y una entrada estática con un patrón vertical, en la figura 2 se muestra el estado de activación de las columnas (la representación dispersa calculada) y la fuerza de las conexiones para una columna activa y para otra inactiva. Este estado se alcanza tras presentar la misma entrada a la región durante varias iteraciones, para que las sinápsis reflejen el proceso de aprendizaje.

En el caso de una columna activa, el aprendizaje provoca que las conexiones que no están contribuyendo a la actividad sean penalizadas. La columna está especializándose en la detección del patron espacial representado por las conexiones fuertes. En el caso de una columna inactiva, las conexiones permanencen en un estado similar al inicial, donde se puede ver la forma de la función de inicialización que decrece con la distancia del bit de la entrada respecto a la proyección de la columna.

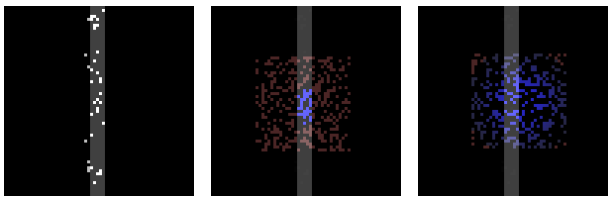


Figura 2: Representación dispersa y aprendizaje

En la figura 3 se muestra, a la izquierda, la entrada percibida en un momento dado por la región y la representación dispersa; a la derecha se represent la unión de los patrones reconocidos por las columnas en actividad.

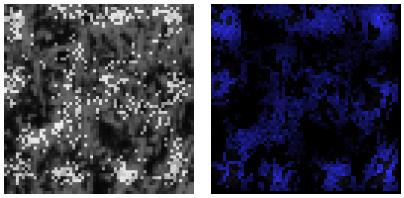


Figura 3: Patrones espaciales percibidos

5.2 Temporal Pooler

El *temporal pooler* trabaja en el espacio de células y utiliza la representación dispersa como entrada. Las células se consideran agrupadas en las columnas correspondientes del *spatial pooler*, de forma que la actividad de una columna en la representación dispersa va a tener influencia en todas las células de esa columna.

El hecho de disponer de varias capas de células (varias células por columna) permite representar el mismo patrón espacial en diferente contexto temporal. Por ejemplo, un patrón *A* será representado por diferentes células si aparece como *BAB* que si aparece como *NAN*.

Además de un estado de actividad, cada célula tiene asociados otros dos estados, uno de aprendizaje y otro de predicción. Las células establecen sinápsis entre sí, pero de forma más compleja que en el caso de las columnas. En primer lugar, las sinápsis se agrupan en segmentos, cero o más por célula, que se activan o no dependiendo del número de sus sinápsis activas. Además, las operaciones de aprendizaje sobre segmentos como la creación o modificación de sus sinápsis se *programan*, pero no se implementan hasta que se cumplan ciertas condiciones.

El objetivo del *temporal pooler* es construir sinápsis que almacenen los patrones temporales presentes en la entrada y que permitan realizar predicciones de la entrada que está por llegar. El funcionamiento se divide en tres fases:

1. En la primera se compara la representación dispersa de la entrada con la última predicción realizada por la región.

   - Si no coinciden la entrada se considera inesperada y todas las células de las columnas involucradas pasan a estar activas; ya que no hay una representación en contexto válida, cualquiera es posible. En cada columna se *fuerza* la selección de una célula de aprendizaje y se crean segmentos para relacionar el patrón resultante (la representación del patrón actual) con el previo.
   - Cuando la predicción sí se cumple se activan únicamente algunas células por columna para obtener la representación del patrón espacial en el contexto de la entrada previa. Las células elegidas son aquellas que contienen segmentos que se activan ante la entrada previa; es decir, no es la primera vez que ambos patrones aparecen en secuencia. Si además los segmentos involucrados se activan con el estado de aprendizaje previo, las células son seleccionadas para el aprendizaje.

2. En la segunda fase se calcula la predicción. Cualquier célula que tenga al menos un segmento que reaccione ante el patrón actual pasa a estar en estado predictivo. Es decir, las células saben que con anterioridad el patrón provocó su activación y esperan que así vuelva a ser.

   Los segmentos involucrados son programados además para especializar las sinápsis respecto al patrón actual, lo que refuerza la relación de predicción entre los patrones. Además, con el objetivo de predecir entradas más posteriores, se selecciona el segmento más activo respecto al patrón previo y se programa su actualización.

3. Para terminar, las células que están en estado de aprendizaje implementan de forma positiva los cambios programados para sus segmentos. Las células que no están en aprendizaje y que además han pasado de la predicción a la no predicción implementan los cambios programados, pero de forma negativa. Estos cambios pueden suponer crear nuevos segmentos o modificar la fuerza de las conexiones, reforzándolas o debilitándolas según la tendencia positiva o negativa de la implementación.

5.2.1 Parámetros

Por una parte, varios parámetros controlan la creación de sinápsis: Número de segmentos por célula, número de nuevas sinápsis en cada actualización, número de sinápsis por segmento y fuerza inicial de las conexiones.

Otros parámetros influyen en el disparo de los segmentos, son umbrales de corte para el nivel de actividad de un segmento que determina si se disparará o si será considerado en el aprendizaje.

Por último, los cambios programados en los segmentos incluyen refuerzo ó debilitación de las conexiones, con incrementos de fuerza dados por otros parámetros.

5.2.2 Implementación.

De igual forma que en *spatial pooler*, las fases de cómputo se realizan de forma distribuída, bien sobre el espacio de columnas o bien sobre el de celdas. Las barreras de sincronización propias de OpenMP garantizan que los resultados intermedios están disponibles.

5.2.3 Ejemplo

Sea una entrada compuesta por cuatro patrones A, B, C y D mostrados de forma cíclica: $ABCDAB\ldots$. La red constará de una única región con una única capa de células. Se verá cómo en cada repetición del ciclo la predicción va llegando cada vez más lejos en el tiempo.

1. En el primer ciclo todos los patrones son inesperados. Una célula por columna activa será seleccionada para representar cada patrón y establecerá un segmento que lo relacione con el patrón inmediatamente anterior. De esta forma, al terminar el ciclo la región relaciona AB, BC y CD mediante segmentos de secuencia.

   En la figura 4 se muestra el estado del sistema en el paso previo (A) y en el actual (B). A la izquierda aparece la entrada y su representación dispersa, y a continuación el estado de las células en actividad (rojo), aprendizaje (azul) y predicción (verde; en el primer ciclo no existe). Los patrones que aparecen en la entrada se han añadido al fondo en las representaciones de células, como referencia. Por simplicidad la región consta de una única capa de células.

   Al observar el patrón B que resulta inesperado, se activan todas las células de las columnas en la representación dispersa. Las seleccionadas como aprendizaje establecen segmentos de secuencia enlazados al aprendizaje previo, lo que efectivamente establece que el patrón B viene después de A.

Figura 4: Patrones A y B del primer ciclo.

2. A lo largo del segundo ciclo ya se produce predicción, pues los nuevos segmentos de secuencia van siendo disparados por los patrones. Las células predictivas programan modificaciones en sus segmentos para reforzar los involucrados en la predicción y para establecer nuevas relaciones contra el patrón de actividad previo.

   Por ejemplo, en la figura 5 se muestra el estado de las células para las entradas A (previo) y B (actual) del segundo ciclo. Al observar el patrón B se cumple la predicción, con lo que se activan determinadas células en cada columna, en lugar de todas. Al tratarse de una única capa, este hecho no puede apreciarse aquí.

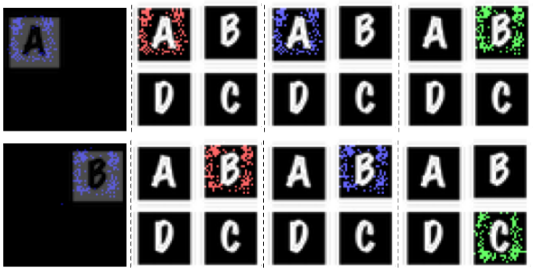


Figura 5: Patrones A y B del segundo ciclo.

   Cuando se presenta B se confirma la predicción que estaban haciendo las células (la dada por el segmento AB con el patrón de actividad previo, A). La entrada ya no es inesperada y sólo las células con esos segmentos pasan a estar activas; se trata de la representación de B en el contexto AB. Esos mismos segmentos enlazan con células que estaban en estado de aprendizaje, por lo que las células pasan además a ser de aprendizaje.

   En la segunda fase del *temporal pooler*, la nueva predicción se produce al disparar la actividad actual B los segmentos BC. Las células predictivas representan C en el contexto BC y programan actualizaciones que refuerzan su respuesta ante el patrón actual B y la construcción de segmentos que las relacionen con el previo A. Estas programaciones quedan pendientes de que se confirme su validez. En este ejemplo serán confirmadas inmediatamente, cuando aparezca C.

   El resultado al término del segundo ciclo es que la región no solo conoce AB, BC, CD y DA mediante segmentos de secuencia, sino que además ha establecido la relación entre $D \cdot B$, $A \cdot C$ y $B \cdot C$ con segmentos no de secuencia.

3. En el tercer ciclo la predicción alcanza no solo el momento inmediatamente posterior, sino un paso más. Así, por ejemplo, al presentarse el patrón B la región predice C y D. Las células predictivas acumulan modificaciones en sus segmentos, entre ellas las que construyen relaciones con el patrón previo. Aquí se añadirán segmentos $A \cdot C$ y $A \cdot \cdot D$.

   En la figura 6 aparecen A (previo) y B (actual) del tercer ciclo. Aquí la predicción ya alcanza un paso más en el tiempo.

6 Conclusiones.

En este trabajo se ha presentado una implementación en C++ de los algoritmos descritos en [11]. En primer lugar se busca disponer de una implementación sobre la que poder experimentar, en un modelo bio-inspirado, con cuestiones como jerarquía, abstracción, predicción temporal o representaciones dispersas. Por otra parte, se trata de un proble-

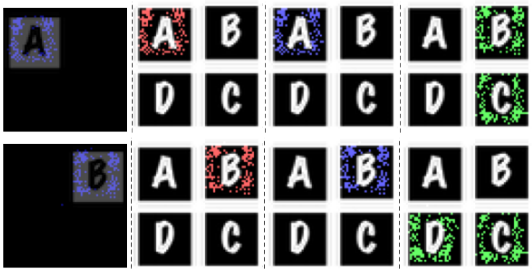


Figura 6: Patrones *A* y *B* del tercer ciclo.

ma ideal para explorar aspectos técnicos relacionados con la explotación del paralelismo inherente a ese modelo.

6.1 Trabajos futuros.

El tipo de entradas con las que se ha experimentado, hasta ahora, están basadas en percepción visual. Sin embargo cualquier secuencia de patrones de bits debería ser válida, se trata de plantear una máquina de propósito general. Estudiar el comportamiento de los algoritmos con entradas heterogéneas, obtenidas de fuentes de datos como tablas o series temporales, parece una buena oportunidad para realmente determinar la capacidad del modelo de afrontar problemas diferentes.

La implementación actual aprovecha el paralelismo dentro de ambos *poolers*, mediante técnicas de memoria compartida (OpenMP). En un futuro inmediato se estudiarán otras formas de paralelismo, como por ejemplo el uso de paso de mensajes para conectar varias regiones.

Por último, el gran número de parámetros que gobiernan el funcionamiento de una región sugiere que podría ser interesante instaurar algún tipo de control *fuzzy* sobre su funcionamiento.

Referencias

[1] A. Alink, C.M. Schwiedrzik, A. Kohler, W. Singer, and L. Muckli. Stimulus predictability reduces responses in primary visual cortex. *J Neurosci.*, 30(8):2960–2966, 2010.

[2] Moshe Bar. The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7):280 – 289, 2007.

[3] Moshe Bar. The proactive brain: memory for predictions. *Philosophical Transactions of The Royal Society B*, 364(1521), 2009.

[4] Barbara Chapman, Gabriele Jost, and Ruud van der Pas. *Using OpenMP: Portable Shared Memory Parallel Programming (Scientific and Engineering Computation)*. The MIT Press, 2007.

[5] Karl Friston and Stefan Kiebel. Predictive coding under the free-energy principle. *Philosophical Transactions of The Royal Society B*, 364:1211–1221, 2009.

[6] Joaquin M. Fuster. Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience*, 21(11):2047–2072, 2009.

[7] Daniel J. Graham and David J. Field. Sparse coding in the neocortex. *Evolution*, III:181–187, 2007.

[8] Jeff Hawkins and Sandra Blakeslee. *On Intelligence*. Times Books, 2004.

[9] Warren McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5:115–133, 1943. 10.1007/BF02478259.

[10] B McNaughton. Cortical hierarchies, sleep, and the extraction of knowledge from memory. *Artificial Intelligence*, 174(2):205–214, 2010.

[11] Numenta. Hierarchical temporal memory.

[12] B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Reseach*, 37:3311–3325, 1997.

[13] Bruno Olshausen and David Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–487, 2004.

[14] Maximilian Riesenhuber and Tomaso Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2:1019–1025, 1999.

IMPRECISIÓN DE LOS VOTANTES EN LOS SISTEMAS DE VOTACIÓN. UNA PROPUESTA BASADA EN MÚLTIPLES TÉRMINOS LINGÜÍSTICOS

Edurne Falcó Díaz de Cerio[1], **José Luis García Lapresta**[2], **Llorenç Roselló Saurí**[3]

[1]*Grupos de Investigación PRESAD (Universidad de Valladolid) y SEED (Universidad Pública de Navarra)*, edurne@eco.uva.es
[2]*Grupo de Investigación PRESAD, Universidad de Valladolid*, lapresta@eco.uva.es
[3]*ESADE - Universitat Ramon Llull*, llorenc.rosello@esade.edu

Resumen

En este trabajo se propone un sistema de votación en el que los votantes juzgan a los candidatos mediante expresiones lingüísticas. Partiendo de una escala finita de términos lingüísticos, los votantes pueden asignar un único término cuando están seguros de sus opiniones (en la escala a utilizar). En caso de duda, los agentes tendrán la oportunidad de asignar varios términos lingüísticos consecutivos. Para determinar el orden de los candidatos que mejor represente las opiniones individuales se desarrolla una metodología en la que, en primer lugar, se halla una calificación colectiva para cada candidato, que puede ser un término lingüístico o un intervalo de términos; finalmente, para dirimir los empates se establece un proceso secuencial de desempate.

Palabras Clave: Toma de decisiones colectivas, Sistemas de votación, Majority Judgment, Términos lingüísticos, Expresiones lingüísticas.

1. INTRODUCCIÓN

La Teoría de la Elección Social muestra que no existe un sistema de votación que sea completamente aceptable. El conocido Teorema de Imposibilidad de Arrow [1] prueba con certeza matemática que ningún procedimiento de agregación de preferencias cumple, de manera simultánea, varias propiedades deseables[1].

El resultado pesimista de Arrow marca los límites a la agregación de preferencias pero no impide la búsqueda de sistemas de votación que cumplan propiedades deseables, a costa de no satisfacer otras. Una de las vías de escape al Teorema de Imposibilidad de Arrow consiste en permitir a los votantes mostrar sus opiniones de manera no puramente ordinal, mediante preferencias o valoraciones numéricas o lingüísticas.

A continuación se hace un breve repaso del sistema de votación con valoraciones lingüísticas más conocido, así como algunas extensiones introducidas recientemente, para después pasar a resumir la propuesta contenida en este trabajo.

1.1. Majority Judgment

Balinski y Laraki [2, 4] han propuesto recientemente un sistema de votación llamado *Majority Judgment* (MJ) que permite a los votantes asignar términos lingüísticos a los candidatos como *Excelente, Muy bueno, Bueno, ...*, en lugar de ordenar dichos candidatos. Entre todos los términos lingüísticos dados por los votantes, MJ escoge la mediana como asignación colectiva. Balinski y Laraki proponen también un proceso de desempate que compara el número de términos que se encuentran por encima de la mediana y por debajo de ella.

Dichos autores llevaron a cabo, además, un análisis experimental de MJ durante las elecciones presidenciales de Francia de 2007 [3] en el distrito de Orsay. En ese trabajo, los autores muestran algunas propiedades interesantes de MJ y defienden que el sistema de votación es de fácil implementación, además de evitar la habitual y costosa segunda ronda de las elecciones francesas presidenciales (si ningún candidato obtiene más de la mitad de los votos en la primera votación, en la segunda ronda se vota sobre los dos candidatos más votados en la primera).

Se han atribuido a MJ ventajas y propiedades deseables frente al marco clásico de Arrow y su agregación de preferencias. Entre estas ventajas se encuentra cómo los votantes muestran sus preferencias de manera más fidedigna y apropiada que en los sistemas de votación convencionales. Además los autores afirman que el método cumple otras propiedades como son el anonimato, la neutralidad, la independencia de alternativas irrelevantes, etc.

[1]Cualquier regla de votación que genere un orden social débil para cada perfil de ordenes débil y satisfaga independencia de alternativas irrelevantes y unanimidad es necesariamente dictatorial siempre y cuando haya al menos tres alternativas y tres votantes.

Aparte de MJ, pueden encontrarse en la literatura otros sistemas de votación donde los votantes emiten sus preferencias mediante términos lingüísticos. Así en García-Lapresta [8] se diseña y estudia una clase general de sistemas de votación que generaliza la mayoría simple mediante la utilización de preferencias lingüísticas. Análogamente, en García-Lapresta *et al.* [11, 9] se diseñan y estudian sistemas de votación que generalizan la regla de Borda.

1.2. Extensiones de Majority Judgment

Conviene señalar que algunos autores han denunciado la aparición de inconsistencias y paradojas en MJ (véanse Felsenthal y Machover [7], Smith [20] y García-Lapresta y Martínez-Panero [10], entre otros).

Con objeto de aminorar alguno de los inconvenientes que se producen en la utilización de MJ en pequeños comités, en García-Lapresta y Martínez-Panero [10] se desarrolla una propuesta en la que la información lingüística se agrega mediante operadores OWA centrados (véase Yager [26]) y el uso de 2-tuplas lingüísticas (véase Herrera y Martínez [13]).

Por otra parte, en Falcó y García-Lapresta [5] se propone una extensión de MJ basada en la utilización de distancias entre términos lingüísticos. Estas distancias se calculan a través de métricas de la familia parametrizada de Minkowski, lo cual permite dar un tratamiento flexible al problema planteado. La extensión que se lleva a cabo en [5] elige como valoración colectiva un término lingüístico que minimiza la distancia a los términos asignados por los agentes (ésta sería la mediana si se utiliza la métrica de Manhattan, caso particular de la familia de Minkowski). Además se propone un método para escoger un único término en el caso en el que haya varios que cumplan el requisito mencionado. En el trabajo se propone también un refinamiento del proceso de desempate de MJ mediante el uso de distancias, de forma que no sólo se cuente el número de valoraciones individuales que hay por encima y por debajo de la valoración colectiva obtenida por cada candidato (como ocurre en MJ), sino que se tenga en cuenta cuáles son.

En Falcó y García-Lapresta [6] se lleva a cabo un primer análisis de las propiedades del operador de agregación que genera la valoración colectiva de cada candidato a partir de las valoraciones individuales, así como las de la regla de decisión asociada al proceso introducido en Falcó y García-Lapresta [5].

1.3. Tratamiento de la imprecisión en la valoración de los candidatos

De acuerdo con Zimmer [28], por lo general, las personas prefieren manejar la incertidumbre con expresiones verbales a hacerlo con números. Por otra parte, las opiniones no suelen ser precisas y, por tanto, no tiene sentido representarlas con precisión. Además, la mayor parte de las personas se siente más cómoda con el uso de palabras que con números (véase Wallsten *et al.* [24]). Fruto de este tipo de reflexiones y evidencias empíricas ha surgido el programa *computing with words*, en donde los agentes muestran sus opiniones a través de términos lingüísticos (véanse Kacprzyk y Zadrożny [14] y Zadeh [27], entre otros).

Si bien el uso de información lingüística acerca el diseño de sistemas de votación a la imprecisión con la que se enfrentan los votantes al juzgar los candidatos, en ocasiones los votantes pueden dudar sobre qué términos lingüísticos usar. Por esto resulta de interés que los votantes puedan emitir sus opiniones de forma más imprecisa, asignando varios términos lingüísticos consecutivos a los candidatos. En este sentido resultan de interés los trabajos de Tang y Zheng [21], Ma *et al.* [15] y, más recientemente, Rodríguez *et al.* [16].

Nuestra propuesta para tratar la imprecisión en la valoración de candidatos se basa en una adaptación de los *espacios de órdenes de magnitud absolutas* propuestos por Travé-Massuyès y Dague [22] y Travé-Massuyès y Piera [23], y más concretamente en las extensiones introducidas en Roselló *et al.* [17, 19].

1.4. Nuestra propuesta

En este trabajo se establece un sistema de votación en el que los votantes pueden expresar sus valoraciones sobre los candidatos mediante un término lingüístico de una escala finita ordenada previamente fijada y, en caso de duda, tienen la posibilidad de asignar una expresión lingüística formada por varios términos consecutivos.

El procedimiento presentado consta de dos fases que tienen como objetivo establecer una ordenación de los candidatos. En la primera fase se asigna una valoración colectiva a cada candidato, que puede ser un término lingüístico o bien una expresión lingüística definida por un intervalo de términos lingüísticos. Esta valoración colectiva está diseñada de forma que se minimice la distancia a las valoraciones individuales de acuerdo a la distancia geodésica en el grafo asociado al conjunto de expresiones lingüísticas. Dado que puede haber varias expresiones lingüísticas que minimicen la distancia a las valoraciones individuales, se establece un procedimiento que permite asignar de forma unívoca una expresión lingüística a cada candidato, la cual será su valoración colectiva.

En la segunda fase se establece un proceso de ordenación de los candidatos en función de las valoraciones colectivas obtenidas. También se introduce un procedimiento de desempate entre los candidatos que estén situados en la misma posición dentro del orden establecido, dando prioridad a los candidatos cuya valoración colectiva venga refrendada por un mayor consenso. Dicho consenso vendrá calculado me-

diante las distancias entre las todas las valoraciones individuales obtenidas por el candidato en cuestión. En el caso de que aún persistiera el empate, se ha previsto un sistema secuencial de desempate que permita ordenar satisfactoriamente los candidatos.

2. PRELIMINARES

De manera análoga al modelo general de MJ, se considera un conjunto de agentes o votantes $V = \{1,\ldots,,m\}$, con $m \geq 2$, que muestra su opinión sobre un conjunto de alternativas o candidatos $X = \{x_1,\ldots,x_n\}$, con $n \geq 2$, a través de una escala finita estrictamente ordenada de términos lingüísticos $L = \{l_1,\ldots,l_g\}$, con $l_1 < \cdots < l_g$ y $g \geq 2$. La *granularidad* de L es el cardinal de L, $\#L = g$. Se supone que la escala es equilibrada y que la distancia entre cada par de términos consecutivos es la misma.

2.1. El conjunto de expresiones lingüísticas

Teniendo en cuenta los espacios de órdenes de magnitud absolutas de Travé-Massuyès y Piera [23], el *conjunto de expresiones lingüísticas* se define como

$$\mathbb{L} = \bigcup\{[l_i,l_j] \mid l_i,l_j \in L,\ 1 \leq i \leq j \leq g\},$$

donde $[l_i,l_j] = \{l_k \mid i \leq k \leq j\}$ y el intervalo $[l_1,l_g]$ se denota con el símbolo ?.

Dado que $[l_i,l_i] = \{l_i\}$, esta expresión lingüística puede identificarse con el término lingüístico l_i, por lo que puede afirmarse que el conjunto de expresiones lingüísticas contiene el conjunto de términos lingüísticos; con abuso de notación, $L \subset \mathbb{L}$.

Ejemplo 1: Sea $L = \{l_1,l_2,l_3,l_4,l_5\}$, cuyos significados vienen dados en el Cuadro 1

Cuadro 1: Significado de las términos lingüísticos.

l_1	l_2	l_3	l_4	l_5
Muy malo	Malo	Regular	Bueno	Muy bueno

Cada expresión lingüística de $\mathbb{L}$ tiene un significado que se deriva de los de los términos lingüísticos de L. Así, $[l_2,l_4]$ significa *entre Malo y Bueno*, $[l_4,l_5]$ significa *entre Bueno y Muy bueno* o *al menos Bueno*, etc. Para una interpretación de las expresiones lingüísticas basada en una gramática libre de contexto, véase Rodríguez *et al.* [16].

El conjunto de expresiones lingüísticas se puede visualizar gráficamente de forma que el nivel inferior corresponda a los términos lingüísticos originales, el segundo nivel a expresiones lingüísticas compuesta por dos términos lingüísticos consecutivos, y así sucesivamente hasta llegar el nivel superior donde sólo hay un elemento, $? = [l_1,l_g]$. En consecuencia, la imprecisión aumenta de abajo a arriba. La Figura 1 ilustra el caso de $g = 5$.

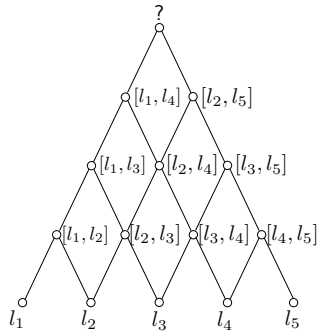


Figura 1: El conjunto de expresiones lingüísticas $\mathbb{L}$ para $g = 5$

Cuando un votante tiene una opinión claramente definida sobre un candidato, tiene la posibilidad de usar un término lingüístico l_i con $1 \leq i \leq g$, mientras que si el votante tiene dudas y prefiere ser menos concreto, puede utilizar una expresión lingüística $[l_i,l_j]$ con $i < j$ (un tratamiento más extenso de este aspecto puede verse en Roselló *et al.* [17, 18, 19]).

2.2. Distancia en el conjunto de expresiones lingüísticas

Para definir una distancia en el conjunto de expresiones lingüísticas $\mathbb{L}$, teniendo en cuenta Roselló *et al.* [19], se considera su grafo asociado $G_{\mathbb{L}}$. Los vértices de $G_{\mathbb{L}}$ son los elementos de $\mathbb{L}$ y las aristas $\mathscr{E} \sim \mathscr{F}$, donde $\mathscr{E} = [l_i,l_j]$ y $\mathscr{F} = [l_i,l_{j+1}]$ o $\mathscr{E} = [l_i,l_j]$ y $\mathscr{F} = [l_{i+1},l_j]$.

Se define la distancia entre dos expresiones lingüísticas $\mathscr{E},\mathscr{F} \in \mathbb{L}$ como la distancia geodésica[2] entre los vértices asociados a $G_{\mathbb{L}}$ y se denota por $d(\mathscr{E},\mathscr{F})$.

La Figura 2 muestra cómo se calcula la distancia entre dos elementos de $\mathbb{L}$ para $g = 5$: $d([l_1,l_3],l_4) = 4$.

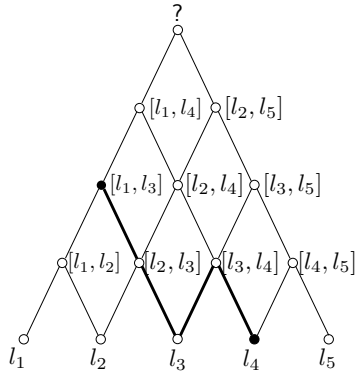


Figura 2: La distancia en $\mathbb{L}$ para $g = 5$

[2]La distancia geodésica entre dos vértices de un grafo es el número mínimo de vértices que debe recorrerse para unirlos, es decir, la longitud del camino más corto que los une (puede haber varios).

2.3. Ordenación del conjunto de expresiones lingüísticas

A continuación se define un orden total $\succ$ en el conjunto de expresiones lingüísticas $\mathbb{L}$ (véase la Figura 3 para el caso de $g = 5$). En primer lugar se mide la distancia de cada expresión lingüística al caso "ideal", que se corresponde con l_g por ser la mejor valoración que se puede obtener, prefiriéndose aquéllas expresiones lingüísticas que estén a menor distancia de l_g. Si las dos expresiones lingüísticas poseen la misma distancia a la valoración "ideal", se preferirá aquélla que tenga una mayor precisión, es decir, la que se encuentre en una posición más baja en el grafo asociado $G_{\mathbb{L}}$ (véase la Figura 1 para el caso de $g = 5$).

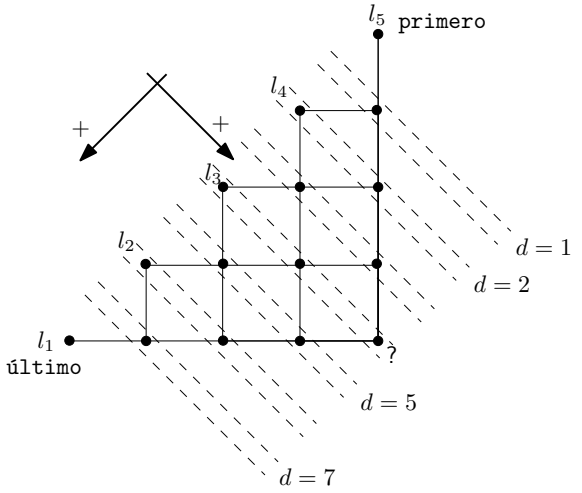


Figura 3: Orden en $\mathbb{L}$ para $g = 5$

Más formalmente, dadas dos expresiones lingüísticas $\mathscr{E}, \mathscr{F} \in \mathbb{L}$, $\mathscr{E} \succ \mathscr{F}$ si se cumple una de las dos condiciones siguientes:

1. $d(\mathscr{E}, l_g) < d(\mathscr{F}, l_g)$
2. $d(\mathscr{E}, l_g) = d(\mathscr{F}, l_g)$ y $\#\mathscr{E} < \#\mathscr{F}$,

donde $\#\mathscr{E}$ es el número de términos que componen $\mathscr{E}$.

Ejemplo 2: Para $g = 5$ la ordenación de todas las posibles valoraciones lingüísticas en $\mathbb{L}$ es:

$$l_5 \succ [l_4, l_5] \succ l_4 \succ [l_3, l_5] \succ [l_3, l_4] \succ [l_2, l_5] \succ$$
$$\succ l_3 \succ [l_2, l_4] \succ ? \succ [l_2, l_3] \succ [l_1, l_4] \succ l_2 \succ$$
$$\succ [l_1, l_3] \succ [l_1, l_2] \succ l_1.$$

3. PROPUESTA DE SISTEMA DE VOTACIÓN

La propuesta que se presenta en este trabajo consta de dos fases diferenciadas que se tratan en cada uno de los subapartados que se presentan a continuación.

3.1. Valoración colectiva de los candidatos

El primer paso a seguir consiste en determinar una *valoración colectiva* para cada candidato. Dicha valoración será una expresión lingüística que sea representativa de las valoraciones individuales obtenidas por el candidato.

Sea $v_j(x_i) \in \mathbb{L}$ la valoración asignada por el votante $j \in V$ al candidato $x_i \in X$. Para cada $x_i \in X$, se define $E(x_i)$ como el conjunto de expresiones lingüísticas que minimizan la suma de distancias a las valoraciones individuales:

$$E(x_i) = \left\{ \mathscr{E} \in \mathbb{L} \mid \forall \mathscr{F} \in \mathbb{L} \sum_{j=1}^{m} d(\mathscr{E}, v_j(x_i)) \leq \right.$$
$$\left. \leq \sum_{j=1}^{m} d(\mathscr{F}, v_j(x_i)) \right\}.$$

Pueden presentarse dos casos:

1. Si $E(x_i)$ tiene un solo elemento, éste constituirá la valoración colectiva de x_i y vendrá denotada por $\mathscr{E}^*(x_i)$.
2. Si $E(x_i)$ tiene más de un elemento, se escogerá como valoración colectiva, $\mathscr{E}^*(x_i)$, la expresión lingüística de $E(x_i)$ con mayor precisión, es decir, la que tenga menor número de términos. Si hubiera más de una, se elegiría la expresión lingüística más cercana a l_g.

3.2. Ordenación de los candidatos

Una vez conocida la valoración colectiva de los candidatos, éstos se ordenarán de acuerdo al orden total establecido en 2.3. Dado que es fácil encontrarse con varios candidatos que compartan la misma valoración colectiva, se hace necesario disponer de un procedimiento de desempate. El que aquí se propone está basado en el grado de consenso de los votantes sobre los candidatos. Entre dos candidatos que compartan la misma valoración colectiva, se preferirá el candidato cuya valoración colectiva tenga un mayor consenso entre los votantes.

El *grado de consenso* entre todos los votantes sobre el candidato $x_i \in X$ se define como

$$\mathsf{cons}(x_i) = 1 - \frac{\sum_{j,k \in V,\, j<k} d(v_j(x_i), v_k(x_i))}{2(g-1)\binom{m}{2}},$$

donde $2(g-1) = d(l_1, l_g)$ es la máxima distancia entre elementos de $\mathbb{L}$.

Este grado de consenso está inspirado en el coeficiente de Gini [12] y cumple $0 \leq \mathsf{cons}(x_i) \leq 1$ y $\mathsf{cons}(x_i) = 1 \Leftrightarrow v_1(x_i) = \cdots = v_m(x_i)$, para cualquier $i \in \{1, \ldots, n\}$.

Si, llegados a este punto, varios candidatos siguieran empatados, se preferiría aquel candidato que tuviera un mayor

número de votos en la valoración ideal l_g. Si aún así persistiera el empate, se tendrían en cuenta los votos conseguidos en la siguiente mejor expresión lingüística, $[l_{g-1}, l_g]$, y así sucesivamente, de acuerdo con el orden total definido en 2.3.

Más formalmente, para establecer la ordenación de los candidatos se tendrá en cuenta el siguiente orden lexicográfico $\rhd$ en X. Dados $x_i, x_j \in X$, $x_i \rhd x_j$ (x_i *vence a* x_j) si se cumple una de las condiciones siguientes:

1. $\mathscr{E}^*(x_i) \succ \mathscr{E}^*(x_j)$
2. $\mathscr{E}^*(x_i) = \mathscr{E}^*(x_j)$ y $\mathsf{cons}(x_i) > \mathsf{cons}(x_j)$
3. $\mathscr{E}^*(x_i) = \mathscr{E}^*(x_j)$, $\mathsf{cons}(x_i) = \mathsf{cons}(x_j)$ y $\#\{k \in V \mid v_k(x_i) = l_g\} > \#\{k \in V \mid v_k(x_j) = l_g\}$
4. $\mathscr{E}^*(x_i) = \mathscr{E}^*(x_j)$, $\mathsf{cons}(x_i) = \mathsf{cons}(x_j)$, $\#\{k \in V \mid v_k(x_i) = l_g\} = \#\{k \in V \mid v_k(x_j) = l_g\}$ y $\#\{k \in V \mid v_k(x_i) = [l_{g-1}, l_g]\} > > \#\{k \in V \mid v_k(x_j) = [l_{g-1}, l_g]\}$
5. ...

A continuación se incluye un ejemplo ilustrativo.

Ejemplo 3: Supongamos $V = \{1, \ldots, 4\}$, $L = \{l_1, \ldots, l_5\}$, $X = \{x_1, x_2\}$ y las valoraciones incluidas en el Cuadro 2.

Cuadro 2: Valoraciones múltiples.

	$v_1(x_i)$	$v_2(x_i)$	$v_3(x_i)$	$v_4(x_i)$
x_1	l_2	$[l_2, l_4]$	$[l_3, l_4]$	$[l_4, l_5]$
x_2	l_1	l_3	$[l_3, l_4]$	l_4

Tras los cálculos pertinentes, se obtiene

$$\mathscr{E}^*(x_1) = [l_3, l_4] \succ l_3 = \mathscr{E}^*(x_2),$$

por lo que x_1 vence a x_2.

Supongamos ahora que los votantes sólo pueden otorgar un término a cada alternativa y que se decantan por asignar las valoraciones incluidas en el Cuadro 3.

Cuadro 3: Valoraciones únicas.

	$v_1(x_i)$	$v_2(x_i)$	$v_3(x_i)$	$v_4(x_i)$
x_1	l_2	l_3	l_3	l_4
x_2	l_1	l_3	l_4	l_4

Ahora se obtendría

$$\mathscr{E}^*(x_2) = l_4 \succ l_3 = \mathscr{E}^*(x_1),$$

por lo que x_2 vencería a x_1.

En consecuencia, la información aportada por los votantes ha sido relevante en la ordenación de los candidatos.

4. Conclusiones

En este trabajo se ha presentado un sistema de votación que permite a los votantes expresar sus opiniones mediante expresiones lingüísticas conformadas por un término lingüístico o varios consecutivos. Con ello se pretende aproximar la toma de decisiones en el ámbito político a la imprecisión con la que los ciudadanos se enfrentan a la hora de juzgar los candidatos en litigio.

El proceso descrito en este trabajo ordena los candidatos en función de una valoración colectiva que minimiza la distancia a las valoraciones individuales. Además se ha presentado un proceso secuencial de desempate basado en el consenso de los individuos y en el número de votos obtenidos con las mejores expresiones lingüísticas.

Aunque ningún sistema de votación está libre de paradojas, el hecho de permitir a los votantes manifestar sus dudas sobre la valoración de los candidatos es de por sí un avance. Permitir a los votantes expresar la imprecisión con la que se enfrentan a la valoración de los candidatos puede generar resultados distintos a los que se obtendrían si se vieran forzados a asignar una valoración inequívoca (véase el Ejemplo 3), pero menos fiel a su forma de pensar.

Al ser éste un trabajo preliminar, quedan por analizar las propiedades del sistema de votación propuesto, así como la comparación con el método *Majority Judgment*, el cual pretende mejorar.

Agradecimientos

Los autores agradecen los comentarios y sugerencias de Jorge Alcalde-Unzu, Ilan Fischer y Francisco Herrera, así como la financiación aportada por el Ministerio de Ciencia e Innovación (proyectos ECO2009-07332, ECO2009-12836, ECO2008-03204-E/ECON, SENSORIAL TIN2010-20966-C02-01 y TIN2010-20966-C02-02) y FEDER.

Referencias

[1] K.J. Arrow: *Social Choice and Individual Values*. 2nd ed., John Wiley and Sons, Nueva York, 1963.

[2] M. Balinski, R. Laraki: A theory of measuring, electing and ranking. *Proceedings of the National Academy of Sciences of the United States of America* 104, pp. 8720-8725, 2007.

[3] M. Balinski, R. Laraki: Election by Majority Judgement: Experimental evidence. *Ecole Polytechnique – Centre National de la Recherche Scientifique* 28, 2007.

[4] M. Balinski, R. Laraki: *Majority Judgment. Measuring, Ranking, and Electing*. The MIT Press, Cambridge MA, 2011.

[5] E. Falcó, J.L. García-Lapresta: A distance-based extension of the Majority Judgement voting system. *Acta Universitatis Matthiae Belii, series Mathematics* 8, pp. 17-27, 2011.

[6] E. Falcó, J.L. García-Lapresta: Aggregating individual assessments in a finite scale. *Proceedings of the World Conference on Soft Computing*, San Francisco, 2011.

[7] D.S. Felsenthal, M. Machover: The Majority Judgement voting procedure: a critical evaluation. *Homo Oeconomicus* 25, pp. 319-334, 2008.

[8] J.L. García-Lapresta: A general class of simple majority decision rules based on linguistic opinions. *Information Sciences* 176, pp. 352-365, 2006.

[9] J.L. García-Lapresta, B. Llamazares, M. Martínez-Panero: A Social Choice analysis of the Borda rule in a general linguistic framework. *International Journal of Computational Intelligent Systems* 3, pp. 501-513, 2010

[10] J.L. García-Lapresta, M. Martínez-Panero: Linguistic-based voting through centered OWA operators. *Fuzzy Optimization and Decision Making* 8, pp. 381-393, 2009.

[11] J.L. García-Lapresta, M. Martínez-Panero, L.C. Meneses: Defining the Borda count in a linguistic decision making context. *Information Sciences* 179, pp. 2309-2316, 2009.

[12] C. Gini: *Variabilità e Mutabilità*. Bologna, Tipografia di Paolo Cuppini, 1912.

[13] F. Herrera, L. Martínez: A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems* 8, pp. 746-752, 2000.

[14] J. Kacprzyk, S. Zadrożny: Computing with words in decision making: through individual and collective linguistic choice rules. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 9 (Suppl.), pp. 89-102, 2001.

[15] J. Ma, D. Ruan, Y. Xu, G. Zhang: A fuzzy-set approach to treat determinacy and consistency of linguistic terms in multi-criteria decision making. *International Journal of Approximate Reasoning* 44 (2), pp. 165-181, 2007.

[16] R. Rodríguez, L. Martínez, F. Herrera: Hesitant fuzzy linguistic terms sets for decision making. *IEEE Transactions on Fuzzy Systems*, en prensa, doi: 10.1109/TFUZ.2011.2170076.

[17] L. Roselló, F. Prats, N. Agell, M. Sánchez: Measuring consensus in group decisions by means of qualitative reasoning. *International Journal of Approximate Reasoning* 51, pp. 441-452, 2010.

[18] L. Roselló, F. Prats, N. Agell, M. Sánchez: A qualitative reasoning approach to measure consensus. En: E. Herrera-Viedma *et al.* (Eds.), *Consensual Processes*, STUDFUZZ, vol. 267, Springer, Heidelberg, pp. 235-261, 2011.

[19] L. Roselló, F. Prats, N. Agell, M. Sánchez, F.A. Mazaira: Using consensus and between generalized multiattribute linguistic assessments for group decision-making. *Information Fusion*, en prensa, doi:10.1016/j.inffus.2011.09.001.

[20] W.D. Smith: On Balinski and Laraki's Majority Judgement median-based range-like voting scheme. `http://rangevoting.org/MedianVrange.html`, 2007.

[21] Y. Tang, J. Zheng: Linguistic modelling based on semantic similarity relation among linguistic labels. *Fuzzy Sets and Systems* 157, pp. 1662-1673, 2006.

[22] L. Travé-Massuyès, P. Dague (Eds.): *Modèles et Raisonnements Qualitatifs*. Hermes Science, París, 2003.

[23] L. Travé-Massuyès, N. Piera: The orders of magnitude models as qualitative algebras. *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Detroit, 1989.

[24] T.S. Wallsten, D.V. Budescu, A. Rapoport, R. Zwick, B. Forsyth: Measuring the vague meanings of probability terms. *Journal of Experimental Psychology* 115, pp. 348-365, 1986.

[25] J.H. Wang, J.Y. Hao: A new version of 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems* 14, pp. 435-445, 2006.

[26] R.R. Yager: Centered OWA operators. *Soft Computing* 11, pp. 631-639, 2007.

[27] L.A. Zadeh: Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems* 4, pp. 103-111, 1996.

[28] A.C. Zimmer: Verbal vs. numerical processing of subjective probabilities. En: R.W. Scholz (ed.), *Decision Making under Uncertainty*. North-Holland, Amsterdam, pp. 159-182, 1983.

AN EVALUATION MODEL FOR COMPANY ENVIRONMENTAL PRACTICES BASED ON INTERACTING CRITERIA

Rocío de Andrés Calle[1], **Teresa González-Arteaga**[2], **Luis Martínez**[3]

[1]*University of Salamanca*, rocioac@usal.es
[2]*University of Valladolid*, teresag@eio.uva.es
[3]*University of Jaén*, martin@ujaen.es

Abstract

Evaluation of the environmental performance is becoming more and more important for companies. In this work we present a multi-criteria decision model for evaluating company's environmental practices. This proposal involves a number of collectives of reviewers related to the company's activity directly or indirectly. In addition, it is considered quantitative and qualitative information with different scales according to the appraisers' knowledge. Due to the relationships among criteria we propose using the Choquet integral as an aggregation operator. The ultimate aim is to compute partial and global indicators that can be used to make decisions regarding the environmental issues.

Keywords: Multi-criteria decision making, Environmental performance, Non-homogeneous information, Choquet integral.

1 INTRODUCTION

Environmental issues are rapidly emerging as an important factor for business and management to consider. The ability of organizations to manage their environmental performance is emerging as a strategic issue for firms. Many studies have been reported about the relation between environmental performance and environmental management systems (see [14]) or the relation with business performance (see [3]) with a variety of results depending on diverse factors.

Environmentally conscious practices refer to programs to improve processes and products environmental performance in diverse forms like design for environment, recycling, waste management, life cycle analysis, green supply chain management, environmental certification (ISO 14000), environmental management systems (EMS), etc. They have evolved with influences from reactive and proactive policies taken by organizations (see [11] and [3]).

Even though evaluation of *environmental performance* (EP) is becoming more and more important recently, there is not a total agreement on what environmental performance is or how to measure it. As a consequence, every organization proposes a different way to define it and to do it (see [14],[3]).

It is recognized the importance to include multiple viewpoints of the concerned parties (see [3]) to increase internal efficiency and external legitimacy. However, many companies use formal methods based just on the opinion of one or various reviewers, in some cases from top environmental managers (see [11] and [3]). Then, the results of the evaluation process could not represent correctly the company environmental situation and they can be biased because the evaluation relies on reviewers that are not the only relevant people to evaluate company's environmental practices.

Several multi-criteria decision making approaches have been used as a basis to deal with environmental issues in various studies in the literature. Some examples of such approaches are Analytical Hierarchy Process (AHP) [10], Analytical Network Process (ANP) [15], grey relational analysis [13], Fuzzy TOPSIS (Technique for Order Performance by Similarity to Ideal Solution) [1], and others specifically developed models.

Moreover, despite there are different types of criteria, quantitative and qualitative, most of the current methods provide only a quantitative precise modelling for their assessments. This fact can produce a lack of precision in the assessments provided by the reviewers for qualitative criteria due to the difficulty of expressing uncertain knowledge in a precise way.

In order to overcome the drawbacks associated with the previous established models to evaluate environmental practices, in this paper we propose a method where different sets of reviewers (internal and external, experts and

non-experts) have to evaluate company's environmental practices attending to different criteria (quantitative and qualitative). Taking into account that the different groups of reviewers may have a different degree of knowledge about the company's environmental practices and the different criteria nature, we then propose a flexible heterogeneous framework in which appraisers may express their assessments by means of numerical or linguistic information according to the criteria nature. In addition, we propose the use of the Fuzzy Linguistic Approach to model qualitative information by means of linguistic variables (see [8]). This approach has been successfully used for this purpose in other evaluation fields and others topics (see [12]).

This paper is organized as follows. Section 2 is devoted to introduce the proposed MCDM model. Section 3 details the rating process that includes the aggregation phase. The paper ends pointing out some concluding remarks.

2 AN ENVIRONMENTAL EVALUATION MODEL

In this section, we present our proposal for evaluating company environmental practices based on a classic decision analysis scheme whose accommodation to our problem is showed in Figure 1.

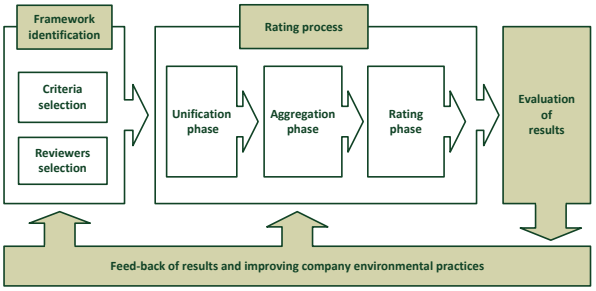


Figure 1: Decision scheme for evaluating company environmental practices

a) Framework identification. In this step of the evaluation process, the criteria and the reviewers of the company's environmental practices are identified.

• *Selection of criteria*. Environmental performance indicators (EPIs) are used to depict the vast quantity of environmental information of a firm in a comprehensive and concise manner. Following the standard ISO 14031, in this paper we consider two general categories of EPIs: *management performance indicators* (MPI) and *operational performance indicators* (OPI). The former provide information about management efforts in relation with the policy, people, practices, procedures, decisions and actions at all levels of the organization. The latter provide information about environmental performance relative to organization's physical facilities and equipment (material and energy flows). Following these categories of indicators we can distinguish two types of criteria associated to them: management criteria and operational criteria.

• *Selection of reviewers*. According to the standard itself [9], in this paper we propose an evaluation method where the assessments are carried out from diverse collectives related to the activity and not only from top environmental managers. Moreover, the proposal method distinguishes between experts and non-experts reviewers.

b) Rating process. To obtain a global environmental performance indicator of each facility site, all information provided by the appraisers must be aggregated. Due to the fact that the framework deals with heterogeneous information, this process is carried out in three phases:

(1) *Normalization and unification information phase*. As the different collectives provide information in different domains we cannot operate directly with it in order to accomplish computations required for the evaluation model. We conduct the original information into a common domain in the way we detail below.

(2) *Aggregation phase*. In this phase the information is aggregated considering different groups of reviewers and different criteria. We use a multi-step aggregation methodology (see Figure 2).

(3) *Rating phase*. The aim of this phase is ranking evaluated facility sites following the goals established by the Department of Environment.

c) Evaluation of the results. One of the most important aspect of an environmental practices evaluation process is to provide feedback about the environmental performance to facility sites. Based on the outcomes, the company can identify and prioritize opportunities for improving the environmental practices.

2.1 EVALUATION FRAMEWORK

In order to show how a company could carry out an environmental practices evaluation process, let us suppose a company which wants to evaluate the environmental practices of its facility sites. It is supposed the company has a set of facility sites $X = \{x_1, ..., x_n\}$ to be evaluated by the following collectives:

• **Internal reviewers**. This reviewers' collective, A, is composed by:

- A set of company's internal experts: $A^E = \{a_1^E, ..., a_m^E\}$.

- A set of company's internal non-experts such as managers, staff, employees, etc.: $A^{NE} = \{a_1^{NE}, ..., a_r^{NE}\}$.

• **External reviewers**. This reviewers' collective, B, consists of:

- A set of company's external experts such as auditors: $B^E = \{b_1^E, ..., b_s^E\}$.

- A set of company's external non-experts evaluators which is made up of two different reviewers' collectives:

1. A set of other "stakeholders" (shareholders, suppliers, government regulators, local communities, intermediate customers, large retailers, final consumers): $B^{NE-G} = \{b_1^{NE-G}, ..., b_t^{NE-G}\}$.
2. A set of social constituents (community groups, trade associations, labor unions, environmental groups): $B^{NE-S} = \{b_1^{NE-S}, ..., b_u^{NE-S}\}$.

Moreover, reviewers evaluate company environmental practices attending to two different types of criteria. Both of them may contain criteria of qualitative or quantitative nature. Due to this fact, we consider a general set of criteria C which includes two subsets of criteria, the quantitative C_1 and the qualitative C_2. The criteria structure is then the following:

- Management performance criteria $C^M = C_1^M + C_2^M$

$$C^M = \{c_k^M \mid c_k^M \in C_1^M \text{ or } c_k^M \in C_2^M,\ k = 1, \ldots, p\}$$

- Operational performance criteria $C^O = C_1^O + C_2^O$

$$C^O = \{c_k^O \mid c_k^O \in C_1^O \text{ or } c_k^O \in C_2^O,\ k = 1, \ldots, q\}.$$

Notice that the same criterion can not be evaluated using different expression domains, in consequence $C_1^M \cap C_2^M = \emptyset$ and $C_1^O \cap C_2^O = \emptyset$.

The assessments provided by the members of the different collectives of reviewers $a_{i,j}^E \in A^E$, $a_{i,j}^{NE} \in A^{NE}$, $b_{i,j}^E \in B^E$, $b_{i,j}^{NE-G} \in B^{NE-G}$ and $b_{i,j}^{NE-S} \in B^{NE-S}$ on the facility site x_j according to the criterion c_k^- are denoted by $a_{i,j,k}^E$, $a_{i,j,k}^{NE}$, $b_{i,j,k}^E$, $b_{i,j,k}^{NE-G}$ and $b_{i,j,k}^{NE-S}$, respectively. Therefore, there are at most $(m+r+s+t+u)(p+q)$ assessments for each locality site since each collective of reviewers could only assess the facility site attending to a number of them following the guidelines established by the Department of Environment.

In this paper we consider an evaluation model which mixes quantitative and qualitative criteria. As a consequence, reviewers can express their opinions for each criterion in different domains, depending on the criteria nature and attending to their knowledge about the evaluated criterion. In this way:

- $a_{i,j,k}^E \in \mathbb{R}^+ \bigvee S_{A^E}^k$ for $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$.
- $a_{i,j,k}^{NE} \in \mathbb{R}^+ \bigvee S_{A^{NE}}^k$ for $i \in \{1, \ldots, r\}$, $j \in \{1, \ldots, n\}$.
- $b_{i,j,k}^E \in \mathbb{R}^+ \bigvee S_{B^E}^k$ for $i \in \{1, \ldots, s\}$, $j \in \{1, \ldots, n\}$.
- $b_{i,j,k}^{NE-G} \in \mathbb{R}^+ \bigvee S_{B^{NE-G}}^k$ for $i \in \{1, \ldots, t\}$, $j \in \{1, \ldots, n\}$.
- $b_{i,j,k}^{NE-S} \in \mathbb{R}^+ \bigvee S_{B^{NE-S}}^k$ for $i \in \{1, \ldots, u\}$, $j \in \{1, \ldots, n\}$.

We can note that any appropriate linguistic term set S_-^k is characterized by its cardinality or *granularity*, $|S_-^k|$.

3 RATING PROCESS

Our proposal considers that reviewers can express their opinions about facility sites in different domains, numerical and linguistic according to the criterion to be evaluated and to the reviewer's knowledge and experience. Therefore and before carrying out the aggregation process, it is necessary to standardize or normalize all gathered information into a unique domain, the interval $[0, g]$, where g is the granularity of the linguistic common domain.

3.1 NORMALIZATION AND UNIFICATION PHASE

Since criteria used to evaluate environmental practices can be quantitative or qualitative, it is necessary to achieve the normalization process by means of two different ways depending on the criterion nature.

On the one side, if gathered information is about a quantitative criterion, we have to distinguish between benefit or cost criterion to accomplish accurately the normalization process. This classification of the quantitative criteria is made based on whether their value is positively or negatively related with the environmental impact incurred. Namely, the same criterion cannot be benefit and cost simultaneously

Let $y_{i,j,k}$ be the assessment of the reviewer j, referred to the facility site x_i over the quantitative criterion k. The normalized previous assessment is defined as $\widetilde{y}_{i,j,k}$ and is carried out as:

$$\widetilde{y}_{i,j,k} = \begin{cases} g\left(\frac{y_{i,j,k}}{y_{k\,max}}\right), & \text{if } k \text{ is a benefit criterion,} \\ g\left(1 - \frac{y_{i,j,k}}{y_{k\,max}}\right), & \text{if } k \text{ is a cost criterion,} \end{cases}$$

where $y_{k\,max}$ is the maximum assessment expressed for all reviewers over all facility sites attending to the k-th criterion, or

$$\widetilde{y}_{i,j,k} = H(y_{i,j,k})$$

On the another side, if the compiled information is with regards to a qualitative criterion, we assume that $S = \{s_0, \ldots, s_h\}$ is a linguistic label set and each label $s_j \in S$ and it should be transformed into a value in the interval $[0, g]$. Moreover, and due to the fact that each collective of reviewers can use different linguistic term sets to evaluate qualitative criteria for each facility sites first of all, we need to unified such sets into a unique linguistic term set commonly called *Basic Linguistic Term Set* (see [6]), denoted by $\overline{S} = \{\overline{s}_0, \ldots, \overline{s}_g\}$ and which granularity $|\overline{S}|$ is g. In order

to achieve that, following [6], we use the next transformation functions

$$H_S : S \xrightarrow{T_{S\overline{S}}} \mathscr{F}(\overline{S}) \xrightarrow{\chi} \langle \overline{S} \rangle \xrightarrow{\Delta_{\overline{S}}^{-1}} [0,g]$$

where $T_{S\overline{S}}$, χ and $\Delta_{\overline{S}}^{-1}$ are defined below.
The idea is to assign a degree of membership to every linguistic label in S for each linguistic label being transformed. The degree of membership is computed by finding the interaction of two linguistic labels belonging to S and $\overline{S}$. After that, the new fuzzy set is transformed into a linguistic 2-tuple.

Definition 1 Let $S = \{s_0, s_1, \ldots, s_h\}$ and $\overline{S} = \{\overline{s}_0, \overline{s}_1, \ldots, \overline{s}_g\}$ be two linguistic term sets, with $h \leq g$. The *linguistic transformation function* $T_{S\overline{S}} : S \longrightarrow \mathscr{F}(\overline{S})$ is defined by:

$$T_{S\overline{S}}(s_j) = \{(\overline{s}_0, \gamma_0), (\overline{s}_1, \gamma_1), \ldots, (\overline{s}_g, \gamma_g)\}$$

$$\gamma_i = \max_y \min\{\mu_{s_j}(y), \mu_{\overline{s}_i}(y)\},\ i = 0, 1, \ldots, g$$

where $\mathscr{F}(\overline{S})$ is the set of fuzzy sets on $\overline{S}$, and μ_{s_j} and $\mu_{\overline{s}_i}$ are the membership functions of the linguistic labels $s_j \in S$ and $\overline{s}_i \in \overline{S}$, respectively.

Definition 2 Let $S = \{s_0, \ldots, s_g\}$ be a set of linguistic terms. The *2-tuple set associated with S* is defined as $\langle S \rangle = S \times [-0.5, 0.5)$. We define the function $\Delta_S : [0,g] \longrightarrow \langle S \rangle$ given by,

$$\Delta_S(\beta) = (s_i, \alpha), \text{ with } \begin{cases} i = \text{round}(\beta), \\ \alpha = \beta - i, \end{cases}$$

where *round* assigns to β the integer number $i \in \{0, 1, \ldots, g\}$ closest to β.

We note that Δ_S is bijective [7] and $\Delta_S^{-1} : \langle S \rangle \longrightarrow [0,g]$ is defined by $\Delta_S^{-1}(s_i, \alpha) = i + \alpha$. In this way, the 2-tuples of $\langle S \rangle$ will be identified with the numerical values in the interval $[0,g]$.

Definition 3 Given the linguistic term set $\overline{S} = \{\overline{s}_0, \overline{s}_1, \ldots, \overline{s}_g\}$, the function $\chi : \mathscr{F}(\overline{S}) \longrightarrow \overline{S} \times [-0.5, 0.5)$, is defined by

$$\chi(\{(\overline{s}_0, \gamma_0), (\overline{s}_1, \gamma_1), \ldots, (s_g, \gamma_g)\}) = \Delta_{\overline{S}} \left(\frac{\sum_{j=0}^{g} j\,\gamma_j}{\sum_{j=0}^{g} \gamma_j} \right).$$

In the end, the transformed assessments for each collective are $\widetilde{a}_{i,j,k}^{E}$ for internal experts reviewers, $\widetilde{a}_{i,j,k}^{NE}$ for internal non-experts reviewers, $\widetilde{b}_{i,j,k}^{E}$ for external experts reviewers and $\widetilde{b}_{i,j,k}^{NE-G}$ and $\widetilde{b}_{i,j,k}^{NE-S}$ for external non-experts reviewers.

We can note that all the information provided by the different collectives has already unified into the interval $[0,g]$. Therefore, we can operate to obtain aggregations of the appraisal results.

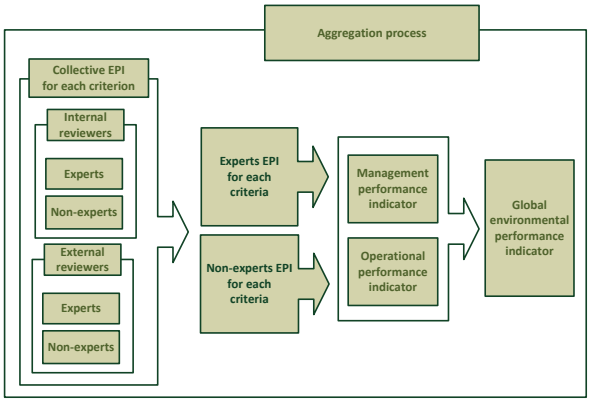


Figure 2: Steps of the aggregation phase

3.2 AGGREGATION PHASE

To obtain a global environmental performance indicator for each facility site, we carry out an aggregation method which has several steps (see Figure 2). To do so, individual assessments will be aggregated by using aggregation operators. They are chosen according to the needs of each aggregation step and considering their properties.

In the specialized environmental evaluation literature, the relationship among EPIs is brought to the attention (see for instance [15]) although it is not explicitly defined. In this paper we are going to manage such a relationship in the aggregation process by means of Choquet integral (see [4] and [5]). This aggregation operator allows us, in comparison to other multi-criteria analysis techniques, to take into account interactions among EPIs or criteria by mean of expressing the significance degree of a combination of them. Also, the global importance of each criterion is considered.

Definition 4 Let $C = \{c_1, \ldots, c_n\}$ be a finite universe. A *fuzzy measure* or *capacity* is a set function $\mu : \mathscr{P}(C) \longrightarrow [0,1]$ which satisfies:

1. $\mu(\emptyset) = 0$ and $\mu(C) = 1$,
2. $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$,

where $\mathscr{P}(C)$ is the set of all subsets of C.

Definition 5 Let $f : C \longrightarrow \mathbb{R}^+$ be a function with $y_k = f(c_k)$ for $k \in \{1, \ldots, n\}$ and $y_0 = 0$. The *discrete Choquet integral* of f with respect to a fuzzy measure μ is given by

$$G_\mu(y_1, \ldots, y_n) = \sum_{k=1}^{n} \left(y_{(k)} - y_{(k-1)}\right)\ \mu\left(A_{(k)}\right),$$

where $\left(y_{(1)}, \ldots, y_{(n)}\right)$ is a non-decreasing permutation of $(y_1, \ldots, y_n)$ and $A_{(k)} = \left\{y_{(k)}, \ldots, y_{(n)}\right\}$.

Remark 1 As it has been proved in several references such as [4], and [2], among others, the Choquet integral fulfills very significant mathematical properties like *idempotence*, *continuity*, *non-decreasing monotonicity*, *stability under the same positive linear transformation*, *decomposability* and *compensativeness*.

There are many different methods to determine an appropriate fuzzy measure (see [5] and [2]) for each application. It is remarkbly to point out that the fuzzy measure plays a key role in the Choquet integral comparable with the choice of weights in the weighted mean. Moreover, the weighted mean and the OWA aggregation operator are special cases of the Choquet integral for specific fuzzy measures which are connected with the weighting vector (see [2]). We propose to use the OWA operator in some of the steps in our aggregation process. In those cases the weighting vectors can be computed using the well-known non-decreasing quantifiers proposed by Yager [16].

A further detailed description of each stage of the aggregation process is presented. It is important to note that each aggregation procedure can use a different fuzzy measure that implies a different Choquet integral. The detailed choice of fuzzy measures are going to be a matter for future works.

1. *Computing environmental performance indicators for each reviewers' collective and each criterion.*

   The reviewers' assessments are aggregated for each criterion and each collective by means of a OWA operator, G^-_- (because OWA operators are anonymous). Then, for each collective and for every criterion c_k^-, the process is conducted in the following manner.

   - For **internal reviewers** (experts and non-experts, respectively):

   $$\begin{aligned} I_k^{A^E}(x_j) &= G_k^{A^E}(\widetilde{a}^E_{1,j,k},\ldots,\widetilde{a}^E_{m,j,k}), \\ I_k^{A^{NE}}(x_j) &= G_k^{A^{NE}}(\widetilde{a}^{NE}_{1,j,k},\ldots,\widetilde{a}^{NE}_{r,j,k}). \end{aligned}$$

   - For **external reviewers** (experts and non-experts, respectively):

   $$\begin{aligned} I_k^{B^E}(x_j) &= G_k^{B^E}(\widetilde{b}^E_{1,j,k},\ldots,\widetilde{b}^E_{s,j,k}), \\ I_k^{B^{NE}}(x_j) &= G_k^{B^{NE}}\left(I_k^{B^{NE-G}}(x_j), I_k^{B^{NE-S}}(x_j)\right), \end{aligned}$$

   where $I_k^{B^{NE-G}}(x_j)$ is the environmental performance indicator for stakeholder reviewers:

   $$I_k^{B^{NE-G}}(x_j) = G_k^{B^{NE-G}}(\widetilde{b}^{NE-G}_{1,j,k},\ldots,\widetilde{b}^{NE-G}_{t,j,k}),$$

   and $I_k^{B^{NE-S}}(x_j)$ is the environmental performance indicator for social constituents reviewers:

   $$I_k^{B^{NE-S}}(x_j) = G_k^{B^{NE-S}}(\widetilde{b}^{NE-S}_{1,j,k},\ldots,\widetilde{b}^{NE-S}_{u,j,k}).$$

2. *Computing environmental performance indicators for experts/non-experts reviewers and each criterion.*

   The previous environmental performance indicators for the x_j facility site: $I_k^{A^E}(x_j)$, $I_k^{A^{NE}}(x_j)$, $I_k^{B^E}(x_j)$ and $I_k^{B^{NE}}(x_j)$ are aggregated for each criterion taking into account if the reviewers are experts or not. The previous indicators belonging to the experts reviewers are then aggregated by means of an OWA operator for each criterion c_k^-. Let $G_k^E : [0,g]^2 \longrightarrow [0,g]$

   $$I_k^E(x_j) = G_k^E\left(I_k^{A^E}(x_j), I_k^{B^E}(x_j)\right).$$

   Analogously to the experts reviewers, an environmental performance indicator is computed for each criterion c_k^- by aggregating the opinions of all non-experts reviewers. Let $G_k^{NE} : [0,g]^2 \longrightarrow [0,g]$

   $$I_k^{NE}(x_j) = G_k^{NE}\left(I_k^{A^{NE}}(x_j), I_k^{B^{NE}}(x_j)\right).$$

3. *Computing management and operational performance indicators.*

   They are obtained by means of aggregating the previous experts and non-experts indicators taking into account the two types of criteria: management and operational. In this step of the aggregation process, we use Choquet integral as an aggregation operator because it allows us, in comparison to other multicriteria analysis techniques, to take into account interactions among EPIs or criteria through expressing the significance degree of a combination of EPIs or criteria.

   • *Management performance indicator.* In order to calculate this indicator we aggregate the experts and non-experts indicators for management criteria, $c^M = \{c_1^M,\ldots,c_p^M\}$ by means of a Choquet integral. Let $G^M : [0,g]^p \longrightarrow [0,g]$

   $$I^M(x_j) = G^M(I_1^E(x_j),\ldots,I_p^E(x_j), I_1^{NE}(x_j)\ldots,I_p^{NE}(x_j)).$$

   • *Operational performance indicator.* Analogously to management performance indicator, an operational performance indicator is computed for operational criteria, $c^O = \{c_1^O,\ldots,c_q^O\}$ by aggregating the experts and no-experts indicators for such criteria by means of a Choquet integral. Let $G^O : [0,g]^q \longrightarrow [0,g]$

   $$I^O(x_j) = G^O(I_1^E(x_j),\ldots,I_q^E(x_j), I_1^{NE}(x_j),\ldots,I_q^{NE}(x_j)).$$

4. *Computing a global environmental performance indicator.*

   It is obtained by aggregating the management and the operational performance indicators, by means of a Choquet integral. Let $G : [0,g]^2 \longrightarrow [0,g]$

   $$I(x_j) = G\left(I^M(x_j), I^O(x_j)\right).$$

All the indicators obtained in each step of the aggregation process, $I_k^{A^E}(x_j)$, $I_k^{A^{\overline{NE}}}(x_j)$, $I_k^{B^E}(x_j)$, $I_k^{B^{\overline{NE}-G}}(x_j)$, $I_k^{B^{\overline{NE}-S}}(x_j)$, $I_k^E(x_j)$, $I_k^{\overline{NE}}(x_j)$, $I^M(x_j)$, $I^O(x_j)$ and I are used for evaluating company environmental practices.

3.3 RATING PHASE

In the exploitation phase, the management team shall classify and order facility sites according to the environmental performance indicators obtained in the previous phase.

Theses values in the interval [0, g] are transformed into lingistic terms in the BLTS using the function Δ_S (definition 2). The sorting and ranking of facility sites are carried out according to the ordinary lexicographic order presented in [7].

4 CONCLUDING REMARKS

In this contribution we have presented a MCDM model for evaluating the company's environmental practices. This model takes into account that appraisers express objective and subjective perceptions and might present different degrees of knowledge about evaluated facility sites. The presented model not only obtain a global environmental performance indicator for each facility site, but also it obtains intermediate environmental performance indicators according to the opinions of each set of reviewers and criterion, and a management and operational performance indicators. Besides, a deeper study about making adjustment of the Choquet integral as an aggregation operator for environmental evaluations seems to be necessary, the particular good properties of these fuzzy integrals look promising. It is worth emphasizing that the proposed model is quite flexible and it allows to the management team customizes how to aggregate the individual opinions.

Acknowledgements

This paper has been partially supported by the research projects: TIN2009-08286, Spanish Ministerio de Educación y Ciencia (Project ECO2009-07332), and ERDF.

References

[1] A. Awasthi, S.S. Chauhan, and S.K. Goyal. A fuzzy multicriteria approach for evaluatin environmental performance of suppliers. *International Journal of Production Economics*, 126:370–378, 2010.

[2] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Springer, Berlin, 2007.

[3] N. Darnall, I. Henriques, and P. Sadorsky. Do environmental management systems improve business performance in an international setting. *Journal of International Management*, 14:364–376, 2008.

[4] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1996.

[5] M. Grabisch and C. Labreuche. A decade of application of the choquet and sugeno integrals in multicriteria decision aid. *Annals of Operations Research*, 175:247–286, 2010.

[6] F. Herrera, E. Herrera-Viedma, and L. Martínez. A fusion approach for managing multi-granularity linguistic term sets in decision making. *Fuzzy Sets and Systems*, 114:43–58, 2000.

[7] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8:746–752, 2000.

[8] F. Herrera, L. Martínez, and P. J. Sánchez. Managing non-homogeneous information in group decision making. *European Journal of Operational Research*, 166:115–132, 2005.

[9] ISO. *Environmental management systems - requirements with guidance for use (ISO 14001: 2004)*. 2004.

[10] A.H.I. Lee, H. Kang, C. Hsu, and H. Hung. A green supplier selection model for high-tech industry. *Expert Systems with Applications*, 36:7917–7927, 2009.

[11] B. Lin, C.A. Jones, and C. Hsieh. Environmental practices and assessment: a process perspective. pages 71–79, 2001.

[12] L. Martínez. Sensory evaluation based on linguistic decision analysis. *International Journal of Aproximated Reasoning*, 44:148–164, 2007.

[13] M.L.Tseng and A.S.F. Chiu. Evaluating firm's green supply chain management in linguistic preferences. *Journal of Cleaner Production*, in press:1–10, 2010.

[14] D. Nawrocka and T. Parker. Finding the connection: environmental management systems and environmental performance. *Journal of Cleaner Production*, 17:601–607, 2009.

[15] J. Sarkis. A strategic decision framework for green supply chain management. *Journal of Cleaner Production*, 11:397–409, 2003.

[16] R. R. Yager. On ordered weighted averaging operators in multicriteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics*, 18:183–190, 1988.

MODELLING INFLUENCE AMONG INDIVIDUALS IN GROUP DECISION MAKING PROBLEMS

I.J. Pérez[1], **F.J. Cabrerizo**[2], **S. Alonso**[3], **E. Herrera-Viedma**[1]

[1]*Dept. of Computer Science and Artificial Intelligence, University of Granada*, {ijperez,viedma}@decsai.ugr.es
[2]*Dept. of Software Engineering and Computer Systems, Distance Learning University of Spain (UNED)*, cabrerizo@issi.uned.es
[3]*Dept. of Software Engineering, University of Granada*, zerjioi@ugr.es

Abstract

With the recent spreading of social networks and digital communities, a new virtual environment has been born where users can share opinions or resources and make decisions. But involving many people in decision making doesn't guarantee success. In addition, in all the societies, there are people that try to exert pressure in order to persuade others that can easily be influenced. In these situations, classical group decision making models fail. Thus, there is still a necessity of developing tools to help users to reach good group decisions as if they where in a real face to face meeting. In such a way, understanding decision making psychology can lead us to reach quick and successful conclusions. Therefore, we propose to analyze the most usual weapons of influence among people and model them with the help of the moderator in order to improve the new virtual decision scenarios making them more precise and realistic.

Keywords: Group decision making, consensus reaching process, weapons of influence, persuasion.

1 INTRODUCTION

Group Decision Making (GDM) arises from many real world situations. Thus, the study of decision making is necessary and important not only in Decision Theory but also in areas such as Management Science, Operations Research, Politics, Social Psychology, Artificial Intelligence, Soft Computing, and so on. In these situations, there is a problem that can be solved in different ways and a group of experts trying to achieve a common solution.

In an ideal world, GDM problems would be solved just regarding the experts self-interest, that is, desire to maximize benefits and minimize cost when driving their decisions. But in the real world GDM problems, experts are suffering continuous attempts by their partners to influence their own opinions. In such a way, experts interaction involves a person trying to change another person opinions and behavior. According to professor Cialdini [2], although there are thousands of different tactics that people employ to influence others, the majority fall within six basic categories:

1. *Social Proof*: People will do things that they see other people are doing.
2. *Authority*: People will tend to obey authority figures, even if they are asked to perform objectionable acts.
3. *Liking*: People are easily persuaded by other people that they like.
4. *Scarcity*: Infrequent items or resources will generate demand.
5. *Consistency*: If people commit, verbally or in writing, to an idea or goal, they are more likely to honor that commitment.
6. *Reciprocation*: People tend to return a favor.

Each of these categories is governed by a fundamental psychological principle that directs human behavior and gives the tactics power of persuasion [2]. These principles of persuasion or weapons of influence can be used as a support for the GDM process as they address the use communication in order to change attitudes, beliefs or the behavior of others in a voluntary manner avoiding the use of coercion.

In this contribution, we propose a preliminary study to overcome such problem by analyzing each one of these weapons of influence on the GDM framework. In such a way, the moderator, who is the only person that interacts with the decision makers, has to manage the influence that each member of the group can exert on the remaining ones, and take it into account when advising members to change their preferences.

In order to do this, the paper is set out as follows. Some general considerations about GDM are presented in Section 2. Section 3 presents the use and modelling of some weapons of influence in GDM problems. Finally, Section 4 draws our conclusions.

2 PRELIMINARIES

In this section, we introduce some concepts and approaches about classical GDM models and the consensus reaching process.

2.1 Group Decision Making

A decision making process, consisting in deriving the best option from a feasible set, is present in just about every conceivable human task. It is obvious that the comparison of different actions according to their desirability in decision problems, in many cases, it cannot be done by using a single criterion or an unique person. Thus, we interpret the decision process in the framework of GDM.

There have been several efforts in the specialized literature to create different models to correctly address and solve GDM situations. Some of them make use of fuzzy theory as it is a good tool to model and deal with vague or imprecise opinions[6, 9, 10, 17, 21].

In a classical GDM situation there is a problem to solve, a solution set of possible alternatives, $X = \{x_1, x_2, \ldots, x_n\}$, $(n \geq 2)$ and a group of two or more experts, $E = \{e_1, e_2, \ldots, e_m\}$, $(m \geq 2)$ characterized by their own ideas, attitudes, motivations and knowledge, who express their opinions about this set of alternatives to achieve a common solution [13, 14, 15]. To do this, each expert has to express his preferences on the set of alternatives by means of a fuzzy preference relation, that is defined as $P^k \subset X \text{x} X$, with a membership function, $\mu_{P^k} : X \text{x} X \rightarrow [0,1]$, where $\mu_{P^k}(x_i, x_j) = p_{ij}^k$ denotes the preference degree of the alternative x_i over x_j for the expert e_k.

- $p_{ij}^k > 1/2$ indicates that x_i is preferred to x_j.
- $p_{ij}^k < 1/2$ indicates that x_j is preferred to x_i.
- $p_{ij}^k = 1/2$ indicates indifference between x_i and x_j.

When cardinality of X is small, the preference relation may be conveniently represented by the n x n matrix $P^k = (p_{ij}^k)$.

Usual resolution methods for GDM problems are composed by two different processes [6] (see Figure 1):

1. *Consensus process:* Clearly, in any decision process, it is preferable that the experts reach a high degree of consensus on the solution set of alternatives. Thus, this process refers to how to obtain the maximum degree of consensus or agreement among the experts on the solution alternatives.

2. *Selection process:* This process consists in how to obtain the solution set of alternatives from the opinions on the alternatives given by the experts.

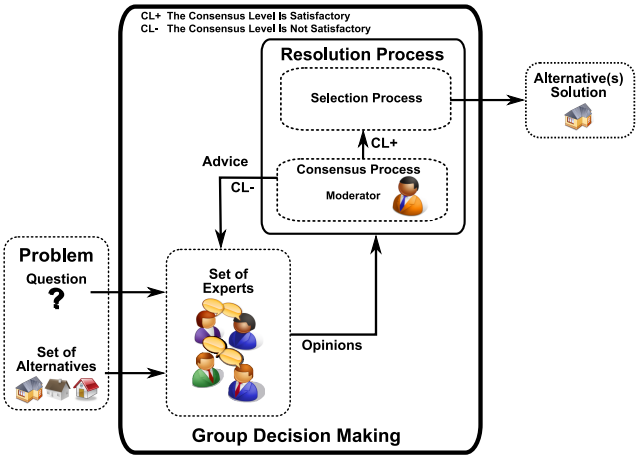


Figure 1: *Resolution process of a GDM*

2.2 Consensus Reaching Process

In group decision making, there are several methods that can be applied. These methods can be classified along a spectrum, from directive to participatory decision making (see Figure 2). The methods that are closer to the directive range, mean that the decision is made by a limited, small number of decision makers in the group. For example, individual dominance method, where one person in the group has the authority or power to make the final decision, or minority influence method, that usually takes the form of decisions delegated from larger groups and made by sub-committees [1]. On the other hand, the methods that are lower on the spectrum, towards the participatory range, mean that the decision is made by all the parties involved. For example, majority rules method usually involves the group voting on the alternatives and the alternative receiving the most votes, wins, or consensus method where the consensual agreement is achieved through group discussion of the alternatives, where every group member can agree on an option and commit to the outcome.

In this contribution we assume the widely studied consensus method [6, 8, 12], going one step further in the way that experts discuss on the alternatives and the relationships and influence among them.

A consensus reaching process in a GDM problem is an iterative process composed by several discussion rounds, in which experts are expected to modify their preferences according to the advice given by the moderator.

Directive	Individual Dominance
↕	Minority Influence
	Majority Rules
Participatory	Consensus

Figure 2: *GDM methods*

Usually, to achieve consensus among the experts, it is necessary to provide the whole group of experts with some advice (feedback information) on how far the group is from consensus, what are the most controversial issues (alternatives), whose preferences are in the highest disagreement with the rest of the group, how their change would influence the consensus degree, and so on.

In such a way, the moderator plays a key role in this process. Normally, the moderator is a person who does not participate in the discussion but knows the preferences of each expert and the level of agreement during the consensus process. He is in charge of supervising and driving the consensus process toward success, i.e., to achieve the maximum possible agreement and reduce the number of experts outside of the consensus in each new consensus round.

Usually, the moderator carries out three main tasks: (i) to compute the consensus measures, (ii) to check the level of agreement and (iii) to produce some advice for those experts that should change their minds.(See Figure 3)

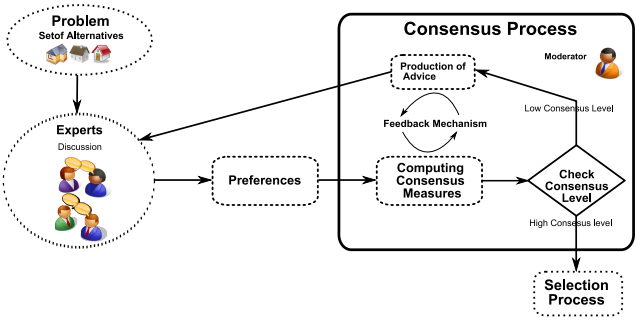


Figure 3: Classical consensus reaching process

Is in this third task where resides the power of the aforementioned weapons of influence. Thus, in the next section we analyze how a GDM moderator can use them in the best way in order to model the real world GDM scenarios where people is persuaded or influenced by others.

3 INFLUENCE AMONG INDIVIDUALS IN GDM PROBLEMS

In any GDM problem, it is desirable that experts reach a high consensus degree on the solution alternative(s). But, when experts' opinions are different among them, it is necessary that some of the experts change their mind in order to achieve the consensus, and is the moderator who has to persuade them to do it.

Usually, persuading people is a difficult task. Our proposal to model this situation is based on the psychology theory that models the human behavior with some built-in automatic responses to stimuli called "fixed-action patterns" activated by a "trigger feature"[2]. In such a way, the moderator should discover the most appropriate trigger feature in each case in order to drive the consensus reaching process in a convergent way.

In this section, we describe the six basic principles of influence among people and their use in GDM problems.

3.1 Social Proof

The principle of social proof states that one important means that people use to decide what to do in a situation is to look at what others are doing or have done there. If many individuals have decided in favor of a particular idea, we are more likely to follow them, because we perceive the idea to be more correct, more valid. Powerful imitation effects have been found among both children and adults and in such diverse activities as purchase decisions, charity donations or phobia remission. This principle can be used to stimulate a person's compliance with a request by informing the person that many other individuals have been complying with it.

Social proof, as weapon of influence, is more influential under two conditions:

1. Uncertainty: When people are unsure, they are more likely to attend to the actions of others and accept them as correct.
2. Similarity: People are more inclined to follow the lead of similar others.

The main advantage of this behavior is that, as there are no many controversial opinions, the consensus is achieved quickly. On the other hand, this principle has no sense if the experts have their own ideas or motivations. In such a case, they will need other reasons, instead of copy the social opinion, to change their minds.

Even though it seems an irrational behavior, this weapon of influence has been successfully used in several GDM models as the only way to compute advice or recommendations to the experts [1, 6, 8, 12, 16].

In these approaches, a collective opinion is computed by aggregating all the individual preferences. Then, those experts whose proximity to the collective opinion is not high enough, are requested to change their minds in order to close their preferences to the collective one.

However, there are more weapons of influence that could help the consensus reaching process under specific circumstances in a more appropriate and realistic way.

3.2 Authority

Sometimes, some individuals with valuable experience, expertise degree or scientific credentials may be seen as authority figures and people will tend to obey them, even if they are asked to perform objectionable acts. This human behavior comes from our education as members of a society where the obedience to legitimate authorities constitutes a correct conduct.

When this obedience is focused to follow individuals with high level of knowledge or wisdom, there is nothing wrong. The problem comes when we are subjected by people with so much power that sometimes abuse of their authority.

The existence of such authority figures has been taken into account in GDM models by giving to their opinions more importance [7, 11, 19, 20]. Even, when advising experts, computing a customized amount of advice recommendations according to their own knowledge level [16]. However, it has never been modeled as a weapon of influence.

To do so, the moderator could act just showing to the experts the opinion of each member considered as an authority, or aggregating all the authorities' opinions and showing to the experts how far are their opinions to the authorities one.

3.3 Liking

Research has pointed out that people tend to rely more on recommendations from people they like than on online recommender systems which generate recommendations based on anonymous people similar to them [18].

There are several examples of research areas where the issue of liking or trust is becoming increasingly important, for example, the semantic web, social networking, virtual communities and, of course recommender systems. In such cases, a kind of trust network is computed in order to model the trust degree between each individual and the remaining ones.

Modeling liking as a weapon of influence in GDM problems could be done by following the same strategy that we explained in the previous subsection for the authority individuals. However, in the previous case, the authority individuals were common to the whole group, but in this situation the moderator needs a trust network [3] in order to recognize the members of the group that are liked by each expert.

3.4 Scarcity

According to the scarcity principle, people assign more value to opportunities when they are less available. This principle holds for two main reasons:

1. Because things that are difficult to attain are typically more valuable, the availability of an item or experience can serve as a shortcut cue to its quality.
2. As things become less accessible, we loose freedoms. According to the psychological reactance theory [5], people respond to the loss of freedoms by wanting to have them (along with the goods and services connected to them) more than before.

This principle of persuasion is most likely to hold true under two optimizing conditions.

1. Scarce items are heightened in value when they are newly scarce. That is, we value those things that have become recently restricted more than those that were restricted all along.
2. We are most attracted to scarce resources when we compete with other for them.

In this case, the only way to use scarcity as a weapon of influence by a moderator in GDM problems, is to present to the experts all the updated information about the alternatives at the beginning of each consensus round. In such a way, every expert always know if an alternative is more valuable than before and they can change their opinion if necessary.

3.5 Consistency

Psychologists have recognized a desire in most people to be and look consistent with their words, beliefs, attitudes and decisions [4]. This tendency for consistency is fed from three sources:

1. Good personal consistency is highly valued by society.
2. Aside from its effect on public image, generally consistent conduct provides a beneficial approach to daily life.
3. A consistent orientation affords a valuable shortcut through the complexity of the modern way of life.

Therefore, this principle can be considered as a weapon of self influence. Once we make a choice, we will encounter personal pressures to behave consistently with that commitment. Those pressures will cause us to respond in ways

that justify our earlier decision. We simply convince ourselves that we have made the right choice and, no doubt, feel better about our decision [4].

By being consistent with earlier decisions, one reduces the need to process all the relevant information in future similar situations. Instead, one merely needs to recall the earlier decision and to respond consistently with it. Commitment decisions, even erroneous ones, have a tendency to be self perpetuating. That is, people often add new reasons and justifications to support the wisdom of commitments they have already made. As a consequence, some commitments remain in effect long after the conditions that spurred them have changed.

Accordingly, the moderator has nothing to do here to influence expert using this principle. However, it is interesting to know which experts change their preferences frequently and which ones have a high level of consistency. This information can be used by the moderator to predict if the use of other weapons of influence on each expert will be successful or it just will be a waste of time.

3.6 Reciprocation

One of the most widespread and basic norms of human culture is embodied in the rule for reciprocation. This rule requires that one person tries to repay what another person has provided. By obligating the recipient of an act to repayment in the future, the rule for reciprocation allows one individual to give something to another with confidence that is not being lost. This sense of future obligation within the rule makes possible the development of different kinds of continuing relationships, transactions, and exchanges that are beneficial to the society.

Moreover, people who do not comply are assigned with labels, such as ungrateful or deadbeat. This happens in reason that it is unpleasant to society in general to be faced with people who take favors and make no efforts to repay them.

One of the reasons for reciprocity is so effective is that this rule has the power to produce a yes as a response to a request that, in the lack of a debt, would surely have been denied. In such a way, this principle can be used by the moderator to remind experts, in a soft way, that they could comply with the change of their opinions in order to close them to the preferences of some expert that previously did the same in the contrary direction.

To model this powerful weapon of influence in the GDM scenarios is not an easy task. To do so, it is necessary to define a framework where the same group of experts has to make several different decisions. Moreover, the moderator has to manage a memory or favor network to remember the different favors among experts. Once the first decision is reached, the moderator can use the network trying to influence experts who received some favors in the previous decisions.

4 CONCLUSIONS

In this contribution, we have done a preliminary study about some of the ways to persuade experts to change their opinions in order to make the consensus process a convergent process. We assume that the moderator has to drive the process by using, if necessary, some weapons of influence to persuade some experts to change their mind as if they were in a real face to face meeting. To do so, we have proposed the introduction of some psychology concepts (or principles or persuasion) in the consensus process, and we have discussed not only the advantages and drawbacks of each of them, but also the way in which they have been or could be used by a virtual moderator in GDM situations.

Acknowledgements

This paper has been developed with the financing of FEDER funds in FUZZYLING Project TIN2007-61079, FUZZYLING-II Project TIN2010-17876, PETRI Project PET2007-0460, Andalusian Excellence Projects TIC-05299 and TIC.5991, and project of Ministry of Public Works 90/07.

References

[1] S. Alonso, I,J. Perez, F.J. Cabrerizo and E. Herrera-Viedma. A Fuzzy Group Decision Making Model for Large Groups of Individuals. In: *Proceedings of 2009 IEEE International Conference on Fuzzy Systems (Fuzz-IEEE09)*, Jeju Island (South Korea), pp. 643–648, 2009.

[2] R. Cialdini. The Science of Persuasion. *Scientific American*, vol. 284, no. 2, pp. 76–81, 2001.

[3] P. Victor, C. Cornelis, M. De Cock and P. Pinheiro da Silva. Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, vol. 160, pp. 1367–1382, 2009.

[4] R.H. Fazio, J. Blascovich and D. Driscoll. On the functional value of attitudes. *Personality and social psychology bulletin*, vol. 18, pp. 338–401, 1992.

[5] J.S. Fogarty. Reactance theory and patient noncompliance. *Social Science and Medicine*, vol. 45, no. 8, pp. 1277–128, 1997.

[6] F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems*, vol. 78, no. 1, pp. 73–87, 1996.

[7] F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. Direct approach processes in group decision making using linguistic owa operators. *Fuzzy Sets and Systems*, vol. 79, pp. 175–190, 1996.

[8] E. Herrera-Viedma, L. Martinez, F. Mata and F. Chiclana. A consensus support system model for group decision-making problems with multi-granular linguistic preference relations. *IEEE Trans Fuzzy Syst*, vol. 13, pp. 644–658, 2005.

[9] E. Herrera-Viedma, S. Alonso, F. Chiclana, and F. Herrera. A consensus model for group decision making with incomplete fuzzy preference relations. *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 5, pp. 863–877, 2007.

[10] J. Kacprzyk and M. Fedrizzi. *Multiperson decision making models using fuzzy sets and possibility theory*. Kluwer Academic Publishers, Dordrecht, 1990.

[11] J. Kacprzyk, M. Fedrizzi and H. Nurmi. Group decision making and consensus under fuzzy preferences and fuzzy majority. *Fuzzy Sets and Systems*, vol. 49, pp. 21–31, 1992.

[12] J. Kacprzyk, H. Nurmi and M. Fedrizzi. Consensus under fuzziness. Kluwer Academic Publishers, Boston, 1997.

[13] J. Lu, G. Zhang and D. Ruan. Intelligent multi-criteria fuzzy group decision-making for situation assessments. *Soft Computing*, vol. 12, pp. 289–299, 2008.

[14] J. Montero. The impact of fuzziness in social choice paradoxes. *Soft Computing*, vol. 12, pp. 177–182, 2008.

[15] H. Nurmi. Fuzzy social choice: a selective retrospect. *Soft Computing*, vol. 12, pp. 281–288, 2008.

[16] I,J. Perez, F.J. Cabrerizo, S. Alonso, and E. Herrera-Viedma. A New Adaptive Consensus Reaching Process Based on the Experts' Importance. In: *Proceedings of the 13th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2010)*, Dortmund (Germany), pp. 474–483, 2010.

[17] M. Roubens. Fuzzy sets and decision analysis. *Fuzzy Sets and Systems*, vol. 90, no. 2 pp. 199–206, 1997.

[18] R. Sinha and K. Swearingen. Comparing recommendations made by online systems and friends. In: *Proceedings of the DELOS-NSF Workshop on Personalisation and Recommender Systems in Digital Libraries* 2001.

[19] R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988.

[20] R. Yager, Weighted maximum entropy owa aggregation with applications to decision making under risk, *IEEE Transactions on Systems, Man, and Cybernetics part A-Systems and Humans*, vol. 39, no. 3, pp. 555–564, 2009.

[21] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences*, vol. 8, pp. 199–249, 1975.

METODOLOGÍA OWA-TOPSIS PARA LA EVALUACIÓN DE TECNOLOGÍAS DE FABRICACIÓN DE CÉLULAS FOTOVOLTAICAS

E. Cables Pérez[1],M. S. Garcia-Cascales[2],M. T. Lamata Jiménez[1], J. M. Sánchez-Lozano[3]

[1]*Depto de Ciencias de la Computación e Inteligencia Artificial.*
Universidad de Granada.18071 Granada
{elio,mtl}@decsai.ugr.es

[2]*Depto deElectrónica, Tecnologíade Computadoras y Proyectos.*
[3]*Depto deExpresión Gráfica.*
Universidad Politécnica de Cartagena. Murcia
{socorro.garcia,juanm.sanchez}@upct.es.

Resumen

En todo problema de decisión donde sea necesario la obtención de un ranking es necesario conocer la matriz de decisión, los pesos de los criterios y el método por el cual se van a valorar las alternativas. En este trabajo se trata de encontrar la mejor tecnología para la fabricación de células fotovoltaicas. Para su evaluación, se ha consultado a los expertos acerca de cuáles son los criterios sobre los que se han de valorar las alternativas, encontrándose que unos son lingüísticos y otros numéricos, de ahí que se ha optado por el método de decisión TOPSIS. Para la obtención de los pesos se propone el operador OWA basado en una escala de razón.

Palabras Clave: OWA, escala de razón, TOPSIS, Comparación por pares, AHP.

1 INTRODUCCION

La toma de decisiones resulta un proceso habitual en la vida cotidiana del ser humano[9]. En multitud de ocasiones nos encontramos ante diferentes alternativas entre las que debemos seleccionar la que, a nuestro juicio, nos parece la mejor o la que satisface el mayor número de las necesidades requeridas, y el ámbito laboral no es una excepción. Los ejecutivos de las empresas y organizaciones, los responsables políticos, los directores de proyectos y en general los profesionales han de adoptar decisiones difíciles durante el transcurso de sus actividades. Estas decisiones según [13], se caracterizan por:

- Contemplar intereses contrapuestos.
- Poseer elementos de incertidumbre.
- Incluir distintas personas en la decisión.
- Exhibir elementos fácilmente valorables y elementos difícilmente valorables.

La actual crisis financiera que ha afectado a la economía mundial no ha dejado fuera ningún sector industrial y los últimos años han sido un gran desafío para la industria fotovoltaica. A pesar de un mercado de alta incertidumbre, la excesiva dependencia en el sector energético de las fuentes de energía de origen no renovable y el cambio climático global continúan resultando fundamentos de gran consistencia que obligan a no dejar de lado las fuentes de energía de origen renovable [4],[6].

El objetivo de alcanzar el Desarrollo Sostenible ha posibilitado que en las últimas décadas se produjeran importantes avances en investigación de nuevas tecnologías y estrategias de ahorro y eficiencia energética. La energía solar fotovoltaica se ha postulado como una excelente alternativa frente a otras fuentes de energía no renovable

El objetivo de este artículo es el estudio y análisis de los criterios de decisión a la hora de elegir la mejor tecnología de fabricación de células fotovoltaicas, teniendo en cuenta tanto criterios de tipo cuantitativo como son los económicos y/o técnicos y criterios de tipo cualitativo como son los sociales y/o ambientales. Los criterios de tipo cuantitativo se definirán mediante valores numéricos mientras que los criterios de tipo cualitativo serán definidos a través de etiquetas lingüísticas. Entre los diversos métodos de decisión multicriterio existentes en la actualidad, uno de los métodos que mejor se ajusta a los parámetros del problema y que permite desarrollar con habilidad y soltura tales características es el método TOPSIS, ya que un paso del algoritmo de cálculo en la normalización.

Entre los diversos métodos de decisión que se podrían utilizar para la obtención de los pesos de los criterios está el Proceso Analítico Jerárquico-AHP [16]. En el presente artículo se propone partir de los fundamentos del método AHP y utilizar otro proceso de cálculo (método OWA) cuya principal característica es que en lugar de realizarse comparaciones por pares entre la totalidad de los criterios, sólo se comparanlos criterios adyacentes, pues se considera que han sido ordenados previamente.

El artículo está organizado de la forma siguiente: la "metodología" engloba la constitución del problema de decisión y describe los métodos sugeridos OWA y TOPSIS. En el "problema de decisión sobre tecnologías de fabricación de células fotovoltaicas" se expondrá la aplicación de los citados métodos. Finalmente, en el apartado de "conclusiones" se describirán las conclusiones principales y los trabajos futuros.

2 OPERADOR OWA BASADO EN UNA ESCALA DE RAZÓN

En el proceso de agregación de la información en los problemas de decisión multicriterio, la utilización de pesos para indicar la importancia relativa de los criterios desempeña un papel esencial. De ahí, la necesidad de la búsqueda constante de métodos que faciliten su determinación a partir de condiciones dadas. Dentro de estos métodos se encuentran los operadores OWA entre los que debemos citar: los que orientan su solución a través de modelos de programación matemática [7, 14, 15, 17, 18,19,20,25]; la utilización de funciones analíticas [1,2, 12, 21]; los orientados por cuantificadores [22]; o el operador promedio priorizado [23] y otros muchos.

Otro enfoquea considerar es aquel que utiliza una escala de razón para indicar cuan superior es un criterio respecto a otro, según el orden de importancia establecido por el decisor. Por ello, el punto de partida es la ordenación de los criteriosde mayor a menor importanciay su valoración dada por el vector $V=[v_1,v_2,...,v_{n-1}]$. El vector $Z=[z_1,z_2,...,z_{n-1}]$ donde $z_j=1/v_j\in[0,1]$ representa la proporción de superioridad inversa del criterio C_j con respecto al C_{j+1}. A partir de esto la expresión (1) permite obtener un vector de pesos [3].

$$w_i=\begin{cases}\dfrac{1}{\left(\sum_{j=1}^{n-1}\prod_{k=1}^{j}z_k\right)+1} & si \quad i=1\\ w_1\prod_{k=1}^{i-1}z_k & si \quad i\neq 1\end{cases} \qquad (1)$$

Verificándose que:

1. Si $V=(1,1,1,...,1)$ entonces el vector de pesos W resultante es $W=\left(\frac{1}{n},\frac{1}{n},...,\frac{1}{n}\right)$.

2. Si $V=(v_1,v_2,...,0)$ entonces el vector de pesos W resultante es $W=(0,0,...,1)$

3. Si $V=(0,v_2,v_3,...,v_{n-1})$ entonces el vector de pesos W resultante es $W=(1,0,...,0)$.

Cualquier vector de pesos obtenido, refleja un nivel de proporción entre sus pesos adyacentes; por tanto, solo será necesario identificar el vector de razón de proporcionalidad que servirá de argumento de la expresión 1.

2.1 CÁLCULO DEL VECTOR DE PESOS

Teniendo en cuenta la formulación del modelo AHP, se desarrolla el cálculo del vector de pesos a través de las etiquetas de la tabla 1 y mediante el siguiente algoritmo.

Tabla 1: Escala de valores

IMPORTANCIA AL COMPARAR POR PARES.	ETIQUETA	VALOR
Igual	*I*	1
Moderada	*M*	3
Fuerte	*F*	5
Muy fuerte	*MF*	7
Sumamente fuerte.	*SF*	9

Nota: Los valores 2(*I-M*), 4(*M-F*), 6 (*F-MF*) y 8 (*MF-SF*) son utilizados para indicar el valor medio entre sus correspondientes juicios adyacentes.

Algoritmo

Para aplicar la nueva forma de cálculo del vector de pesos $W=[w_1,w_2,...,w_n]$ se aplica el algoritmo siguiente:

1.- Listar los criterios por orden descendente de importancia.

2.- Asignar la correspondiente valoración lingüística v_j al par (C_j,C_{j+1}), *j=1,2,…,n-1.*

3.- Hallar la proporción inversa $z_j=1/v_j$ de tal forma que $z_j\in[0,1]$.

4.- Aplicar la formula (1) a los valores z_j .

Ejemplo 1: Sean un conjunto de cinco criterios organizados de la forma siguiente $C_1\ ^{M-F}C_2\ ^{I}\ C_3\ ^{I-M}\ C_4\ ^{I-M}\ C_5$, indicando que el vector de valoración lingüística es $v_j=[4,1,2,2]$ y por tanto el vector de proporcionalidad inversa viene dado por $z_j=[1/4,1,1/2,1/2]$ luego, aplicando la expresión (1) , se obtiene:

$$w_1=\frac{1}{\left(\sum_{j=1}^{4}\prod_{k=1}^{j}z_k\right)+1}=\frac{1}{\frac{1}{4}+\frac{1}{4}\cdot 1+\frac{1}{4}\cdot 1\cdot\frac{1}{2}+\frac{1}{4}\cdot 1\cdot\frac{1}{2}\cdot\frac{1}{2}+1}=\frac{16}{27}$$

$$w_2=\frac{16}{27}\prod_{k=1}^{1}z_k=\frac{16}{27}\cdot\frac{1}{4}=\frac{4}{27},$$

$$w_3=\frac{16}{27}\prod_{k=1}^{2}z_k=\frac{16}{27}\cdot\frac{1}{4}\cdot 1=\frac{4}{27}$$

$$w_4 = \frac{16}{27}\prod_{k=1}^{3} z_k = \frac{16}{27}\cdot\frac{1}{4}\cdot 1\cdot\frac{1}{2} = \frac{2}{27}$$

$$w_5 = \frac{16}{27}\prod_{k=1}^{4} z_k = \frac{16}{27}\cdot\frac{1}{4}\cdot 1\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{27}$$

Por lo tanto, el vector de pesos resultante es $W = (0.593,\ 0.148,\ 0.148,\ 0.074,\ 0.037)$.

3 TOPSIS

El método TOPSIS es un modelo de decisión propuesto para ordenar preferencias por similitud a una solución ideal [8], también han trabajado sobre el tema autores como [6,11,26], y muchos más.

El principio básico del método consiste en encontrar la alternativa cuya distancia a la solución ideal positiva A^+sea menor y cuya distancia a la solución ideal negativaA^-sea mayor. Una solución ideal se define como una colección de puntuaciones o valores en todos los atributos considerados en la toma de decisión, pudiendo suceder que tal solución sea inalcanzable. El vector compuesto por los mejores valores del j-ésimo atributo respecto de todas las alternativas posibles es quien recibe el nombre de "solución ideal positiva" (SIP); recíprocamente, la "solución ideal negativa" (SIN) será aquella cuyo vector contenga los peores valores en todos los atributos.

El concepto intuitivo de alternativa ideal se corresponde con aquella alternativa que, sin dudarlo, siempre elegiría el decisor. De igual modo, la alternativa anti-ideal sería aquella que, sin dudarlo, nunca elegiría el decisor.

3.1. METODO ALGORÍTMICO PROPUESTO

El procedimiento TOPSIS puede expresarse en una serie de pasos que pueden verse en [5], y que nosotros reproducimos manteniendo los 5 primeros pasos y adaptando el resto al caso de variables lingüísticas y por ende de números difusos para el problema que resolvemos a continuación.

Paso 1:Identificación de los criterios de evaluación de forma que relacione las capacidades del sistema con los objetivos.

Paso 2: *Generación de alternativas* determinando el sistema finito de alternativas para lograr el objetivo.

Paso 3: *Evaluación de las alternativas* en función de los criterios.

Paso 4: *Identificación del peso de los criterios.* Es importante considerar que los criterios puedan tener distinta importancia.

Paso 5: *Construcción de la matriz de decisión* con números difusos.

Paso 6: L*a matriz de decisión normalizada* asociada. Para ello los valores de cada criterio los dividimos por la norma con objeto de que la escala sea la misma para todos los criterios.

$$\overline{n}_{ij} = \underline{x}_{ij} \Big/ \sqrt{\sum_{j=1}^{n} \left(\underline{x}_{ij}\right)^2},\ j = 1,\ldots,n,\ i = 1,\ldots,m \qquad (2)$$

utilizando para los números difusos las operaciones definidas en [10], siendo el resultado $\overline{n}_{ij} = \left(n_{ij}^a, n_{ij}^b, n_{ij}^c\right)$ otro número difuso.

Paso 7: Construcción de la matriz ponderada normalizada asociada.

$$\overline{v}_{ij} = \underline{w}_j \overline{n}_{ij},\ j = 1,\ldots,n,\ i = 1,\ldots,m, \qquad (3)$$

Donde $\overline{v}_{ij} = \left(v_{ij}^a, v_{ij}^b, v_{ij}^c\right)$

Paso 8: Determinación de la solución ideal positiva y la solución ideal negativa, respectivamente.

$$\overline{A}^+ = \left\{\overline{v}_1^+,\ldots,\overline{v}_n^+\right\} = \left\{\begin{matrix}\left(\max \overline{v}_{ij},\ j \in J\right),\\ \left(\min \overline{v}_{ij},\ j \in J'\right)\end{matrix}\right\} \qquad (4)$$

$$\overline{A}^- = \left\{\overline{v}_1^-,\ldots,\overline{v}_n^-\right\} = \left\{\begin{matrix}\left(\min \overline{v}_{ij},\ j \in J\right),\\ \left(\max \overline{v}_{ij},\ j \in J'\right)\end{matrix}\right\} \qquad (5)$$

donde J está asociado con aquellos criterios que indican ganancias o beneficios y J' está asociado con los criterios que indican pérdidas o costes.

Paso 9: Cálculo de la separación de cada alternativa con respecto a la solución ideal positiva y a la solución ideal negativa respectivamente:

$$\overline{d}_i^+ = \left\{\sum_{j\in J}\left(\overline{v}_{ij} - \overline{v}_j^+\right)^2\right\}^{\frac{1}{2}},\ i = 1,\ldots,m \qquad (6)$$

$$\overline{d}_i^- = \left\{\sum_{j\in J}\left(\overline{v}_{ij} - \overline{v}_j^-\right)^2\right\}^{\frac{1}{2}},\ i = 1,\ldots,m \qquad (7)$$

en este caso usamos la distancia euclídea, pero podríamos haber tenido en cuenta cualquier otra.

Paso 10: Cálculo de la proximidad relativa de cada alternativa a la solución ideal positiva y negativa mediante el índice de proximidad.

$$\overline{R}_i = \frac{\overline{d}_i^-}{\overline{d}_i^+ + \overline{d}_i^-},\ i = 1,\ldots,m \qquad (8)$$

Si $\rightarrow \overline{R}_i \cong 1 \rightarrow A_i = \overline{A}^+$ Si $\rightarrow \overline{R}_i \cong 0 \rightarrow A_i = \overline{A}^-$

4. PROBLEMA DE TECNOLOGIAS DE FABRICACIÓN DE CÉLULAS FOTOVOLTAICAS

Queremos hacer un estudio sobre las distintas tecnologías de fabricación de células fotovoltaicas.

Para ello el experto nos ha proporcionado los datos relativos a las alternativas a evaluar:

A1: Tecnología de silicio cristalino

A2: Tecnología de capa delgada inorgánica (Silicio amorfo)

A3: Tecnología con capa delgada inorgánica (CdTe y GIGS)

A4: Tecnología de capa delgada avanzada III-V con sistemas de seguimiento con concentración solar

A5: Tecnología de capa delgada avanzada de bajo coste (células híbridas y orgánicas).

Y de los criterios considerados para la evaluación del problema de decisión:

C1: Eficiencia en conversión de energía

C2: Coste de fabricación

C3: Emisiones de gases de efecto invernadero

C4: Tiempo de retorno energético

C5: Cuota de mercado

Tabla 2:Valoración de las alternativas para cada criterio

	C_1 (%)	C_2(€/Wp)	C_3	C_4	C_5
A_1	21.0 ± 3.0	1.440 ± 0.760	m	m	MA
A_2	11.5 ± 3.5	0.675 ± 0.275	m	mB	B
A_3	11.5 ± 3.5	0.900 ± 0.300	m	mB	B
A_4	32.5 ± 7.5	0.450 ± 0.250	A	B	B
A_5	9.0 ± 3.0	0.600 ± 0.200	B	B	B

Los criterios (C_1) y (C_2) son criterios cuantitativos y sus valores se han obtenido de los fabricantes de las distinta tecnologías fotovoltaicas.

Las valoraciones de las alternativas relativas a los criterios por el experto están reflejadas en la tabla 2. Es importante tener en cuenta que de estos criterios, la eficiencia (C_1) y la cuota de mercado (C_5) serán tanto mejores cuanto mayor valor tengan; ello acarrea que tengamos que maximizarlos. En cambio los criterios de coste de fabricación (C_2), emisiones de gases de efecto invernadero (C_3) y tiempo de retorno energético (C_4) serán mejores conforme su valor sea más reducido y lo que haremos será minimizarlos.

Tabla 3. Valoraciones de las etiquetas lingüísticas

DESCRIPCIÓN	ETIQUETA	NÚMEROS DIFUSOS
Muy bajo	MB	[0, 0, 1]
Bajo	B	[0, 1, 3]
Medio bajo	mB	[1, 3, 5]
Mediano	m	[3, 5, 7]
Medio alto	mA	[5, 7, 9]
Alto	A	[7, 9, 10]
Muy alto	MA	[9, 10, 10]

Observamos que estas valoraciones son numéricas para los criterios 1 y 2, mientras que C_3, C_4 y C_5 son lingüísticas. En la tabla 3 se exponen los valores lingüísticos y el correspondiente número difuso asociado.

En los problemas multicriterio no siempre se considera que todos los criterios tienen la misma importancia. En este caso, según el experto el orden de importancia de los criterios es $C_1 \succ C_2 \succsim C_3 \succ C_4 \succ C_5$ y el grado de importancia está dado por el vector [4,1,2,2].

Por lo que teniendo en cuenta el ejemplo 1, los valores de los pesos asociados a los criterios es:

$$W = (0.593,\ 0.148,\ 0.148,\ 0.074,\ 0.037)$$

Según esto, la matriz de decisión correspondería a la tabla 4. Si aplicamos el algoritmo del método TOPSIS obtendríamos la tabla normalizada (tabla 5), la ponderada (tabla 6) y a partir de esta última se calcularían las distancias d^+ y d^- a través de las formulas 6-7, respectivamente, obteniéndose los resultados que aparecen en la tabla 7.

Tabla 7: Valores correspondientes a d^+, d^- y R.

	d^+	d^-	R
A_1	0.2436	0.1991	0.4497
A_2	0.3328	0.1469	0.3063
A_3	0.3351	0.1323	0.2830
A_4	0.1000	0.3948	0.7979
A_5	0.3658	0.1852	0.3361

Si nos fijamos en los resultados obtenidos para d^+ vemos como la mejor alternativa es A_4 ya que la distancia a la alternativa ideal positiva es menor. Con respecto a d^- y al índice R, sucede lo contrario, la alternativa mejor sería la más alejada, pero la mejor alternativa sería también A_4.

5.CONCLUSIONES

Los pesos se han obtenido a través de un operador, y se han eliminado los problemas de consistencia que podrían haber surgido si se hubieran recopilado las 10 comparaciones entre los 5 criterios, necesarios si se hubiese considerado el método AHP. También conseguimos evitar las operaciones necesarias para la obtención de los valores y vectores propios de la matriz resultante de la comparación.

Se ha resuelto el problema de decisión multicriterio relativo a la obtención de la mejor tecnología de células fotovoltaicas cuando la evaluación se hace teniendo en cuenta los criterios, Coste de fabricación, Eficiencia en conversión de energía, Cuota de mercado, Emisiones de gases de efecto invernadero y del Tiempo de retorno energético. Con estas variables se obtiene que la mejor opción es A_4. El resultado obtenido resulta razonable ya que dicha tecnología corresponde a las células de capas delgadas avanzada compuestas por elementos pertenecientes a los grupos III-V de la tabla periódica por lo que, además de utilizar sistemas de seguimiento con concentración solar (características que les confieren un coste de fabricación en función a la potencia generada relativamente bajo), consiguen obtener una elevada eficiencia de forma que permiten situar dicha tecnología al frente del resto. La

segunda mejor alternativa (aunque con bastante diferencia respecto a A_4) es la alternativa A_1, el hecho de que resulte la segunda mejor opción se debe a que esta tecnología es la más utilizada en la actualidad con lo que posee una elevada cuota de mercado, además su eficiencia en conversión de energía es alta. Del resto de resultados obtenidos se observa que las alternativas restantes (A_2, A_3 y A_5) presentan valores similares por lo que no se puede afirmar con claridad que una sea predominante sobre las otras.

Agradecimientos

Este trabajo ha sido financiado por la DGICYT mediante el proyecto TIN2008-06872-C04-04 y por La Junta de Andalucía (P07-TIC02970), respectivamente.

Referencias

[1].B.S. Ahn: On the properties of OWA operator weights functions with constant level of orness, IEEE Transactions on Fuzzy Systems, 14, pp.511–515, 2006.

[2].B.S. Ahn, H.C. Park: Least-squared ordered weighted averaging operator weights. International Journal of Intelligent Systems, 23, pp. 33–49, 2008.

[3].E. Cables Pérez, M.T.Lamata:. OWAs basados en una escala de razón, CAEPIA.2010.

[4].F. Cavallaro: Fuzzy TOPSIS approach for assessing thermal-energy storage in concentrated solar power (CSP) systems. Applied Energy, 87, pp. 496-503, 2010.

[5].S.J. Chen, C.L. Hwang: Fuzzy Multiple Attribute Decision Making: Methods and Applications, Springer-Verlang, Berlin, 1992.

[6].M. S. García-Cascales, M. T. Lamata, M. Sánchez-Lozano. Evaluation of photovoltaic technologies in a process of multi-criteria decision making. Aceptado Annals of Operations Research.

[7].R. Fullér, P. Majlender: On obtaining minimal variablity OWA operator weights. Fuzzy Sets and Systems, 136. pp. 203-215, 2003.

[8].C.L. Hwang, K. Yoon: Multiple Attribute Decision Methods and Applications. Springer- Heidelberg (Ed). Berlin, 1981.

[9].R. Keeney, H.Raiffa: Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Wiley (Ed). New York, 1976.

[10]. G.J. Klir, B. Yuan: Fuzzy sets and fuzzy logic: theory and applications, Prentice Hall, 1995.

[11]. Y.J. Lai, T.Y. Liu, C.L. Hwang: TOPSIS for MODM, European Journal of Operational Research 76 (3), 486-500, 1994.

[12]. M.T. Lamata, E. Cables Pérez: OWA weights determination by means of linear functions. Mathware& Soft Computing,Vol. 16, pp. 107-122, 2009.

[13]. O. G. León, Tomar decisiones difíciles, Segunda Edición ed Universidad Autónoma de Madrid, 2001.

[14]. M. O'Hagan: Aggregating template rule antecedents in real-time expert systems with fuzzy set logic". In: Proc 22nd Annual IEEE Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA: IEEE and Maple Press. pp 681– 689, 1988.

[15]. P. Majlender: OWA operators with maximal Renyi entropy. Fuzzy Sets and System, 155. pp. 340-360, 2005.

[16]. T.L. Saaty: The Analytic Hierarchy Process. New-York, 1980.

[17]. Z.s. Xu, Q.L. Da: Approaches to obtaining the weights of the ordered weighted aggregation operators. J Southeast Univ. 33pp 94–96, 2003.

[18]. R.R. Yager: Families of OWA operators. Fuzzy Sets and Systems 59, 125-148, 1993.

[19]. R.R. Yager: On the maximum entropy negation of a probability distribution. Technical Report # MII-1303, Machine Intelligence Institute, Iona College, New Rochelle, NewYork, 1993.

[20]. R.R. Yager: On the completion of qualitative possibility mesures. IEEE Trans Fuzzy Syst. 1 pp.184–194, 1993.

[21]. R.R. Yager, D.P. Filev: Parameterized "andlike" and "orlike" OWA operators. Int J Gen Syst; 22, pp.297–316, 1994.

[22]. R.R. Yager: Quantifier Guided Aggregation Using OWA Operators. International Journal of Intelligent Systems, 11, pp. 49-73, 1996.

[23]. R.R. Yager: Prioritized aggregation operators. International Journal of Approximate Reasoning, 48, pp. 263–274, 2008.

[24]. R.R. Yager: Prioritized OWA aggregation. Fuzzy OptimDecis Making,8, pp. 245–262, 2009.

[25]. Y,.M Wang, C.Parkan: A minimax disparity approach for obtaining OWA operator weights. Information Sciences, 175. pp. 20-29, 2005.

[26]. M. Zeleny: Multiple Criteria Decision Making, McGraw-Hill, New York, 1982.

Tabla 4:Matriz de decisión

	C_1			C_2			C_3			C_4			C_5		
Pesos	0.5926			0.1481			0.1481			0.0741			0.0370		
A_1	18	21	24	0.68	1.44	24	3	5	7	3	5	7	9	10	10
A_2	8	11.5	15	0.4	0.675	15	3	5	7	1	3	5	0	1	3
A_3	8	11.5	15	0.6	0.9	15	3	5	7	1	3	5	0	1	3
A_4	25	32.5	40	0.2	0.45	40	7	9	10	0	1	3	0	1	3
A_5	6	9	12	0.4	0.6	12	0	1	3	0	1	3	0	1	3

Tabla 5: Matriz de decisión una vez normalizada

	C_1			C_2			C_3			C_4			C_5		
Pesos	0.5926			0.1481			0.1481			0.0741			0.0370		
A_1	0.3420	0.4892	0.7194	0.2359	0.7290	2.0232	0.1875	0.3990	0.8030	0.2774	0.7454	2.1106	0.7717	0.9806	1.1111
A_2	0.1520	0.2679	0.4496	0.1387	0.3417	0.8737	0.1875	0.3990	0.8030	0.0925	0.4472	1.5076	0.0000	0.0981	0.3333
A_3	0.1520	0.2679	0.4496	0.2081	0.4556	1.1036	0.1875	0.3990	0.8030	0.0925	0.4472	1.5076	0.0000	0.0981	0.3333
A_4	0.4750	0.7571	1.1990	0.0694	0.2278	0.6437	0.4375	0.7183	1.1471	0.0000	0.1491	0.9045	0.0000	0.0981	0.3333
A_5	0.1140	0.2097	0.3597	0.1387	0.3038	0.7357	0.0000	0.0798	0.3441	0.0000	0.1491	0.9045	0.0000	0.0981	0.3333

Tabla 6: Matriz de decisión ponderada junto con las alternativas A^+ y A^-.

	C_1			C_2			C_3			C_4			C_5		
Pesos	0.5926			0.1481			0.1481			0.0741			0.0370		
A_1	0.2027	0.2899	0.4263	0.0349	0.1080	0.2997	0.0278	0.0591	0.1190	0.0205	0.0552	0.1563	0.0286	0.0363	0.0412
A_2	0.0901	0.1588	0.2664	0.0206	0.0506	0.1294	0.0278	0.0591	0.1190	0.0068	0.0331	0.1117	0.0000	0.0036	0.0123
A_3	0.0901	0.1588	0.2664	0.0308	0.0675	0.1635	0.0278	0.0591	0.1190	0.0068	0.0331	0.1117	0.0000	0.0036	0.0123
A_4	0.2815	0.4486	0.7105	0.0103	0.0338	0.0954	0.0648	0.1064	0.1699	0.0000	0.0110	0.0670	0.0000	0.0036	0.0123
A_5	0.0676	0.1242	0.2132	0.0206	0.0450	0.1090	0.0000	0.0118	0.0510	0.0000	0.0110	0.0670	0.0000	0.0036	0.0123
A^+	0.2815	0.4486	0.7105	0.0103	0.0338	0.0954	0.0000	0.0118	0.0510	0.0000	0.0110	0.0670	0.0286	0.0363	0.0412
A^-	0.0676	0.1242	0.2132	0.0349	0.1080	0.2997	0.1699	0.1064	0.1699	0.0205	0.0552	0.1563	0.0000	0.0036	0.0123

EVALUACIÓN DE LA SOSTENIBILIDAD DE PROYECTOS CON LÓGICA BORROSA

Raúl Martín [1], Raquel Caro [2], Luis Garmendia [3], Alfonso Garmendia [4]

[1] raul.martin@estumail.ucm.es
[2] rcaro@doi.icai.upcomillas.es
[3] lgarmend@fdi.ucm.es
[4] algarsal@upvnet.upv.es

Resumen

La sostenibilidad constituye un criterio esencial para evaluar la calidad de un proyecto. En este sentido este trabajo presenta el desarrollo de una metodología y un programa para la evaluación de la sostenibilidad social, económica y medioambiental de proyectos para llegar a conseguir una evaluación global de sostenibilidad de dichos proyectos. Tradicionalmente los estudios y evaluación de proyectos se realizan solo desde el punto de vista económico. Se aplica la lógica borrosa a cada uno de los cálculos que se realizan en el análisis global de la sostenibilidad: valoración de indicadores, la evaluación cuantitativa y cualitativa del impacto que produce un proyecto en los diferentes factores medioambientales, sociales y económicos. Cabe destacar que dependiendo del tipo de proyecto se tendrá un peso de cada indicador y de cada factor distinto. Se ha tomando como base la normativa existente para la Evaluación de Impacto Ambiental

Palabras Clave: Impacto ambiental, lógica borrosa, soft computing, sostenibilidad medioambiental, sostenibilidad social, sostenibilidad económica

1 INTRODUCCIÓN: SOSTENIBILIDAD Y LÓGICA BORROSA

El problema del cambio climático y la necesidad de conseguir un desarrollo sostenible, están produciendo un cambio de paradigma no solo social, con la creciente concienciación ciudadana respecto a la problemática medioambiental, sino también económico e incluso empresarial.

Tradicionalmente los estudios y evaluación de proyectos se realizan solo desde el punto de vista económico [9, 10], sin embargo, esto en la actualidad ya no es suficiente [3, 6]. Resulta de vital importancia también evaluar en qué medida el proyecto en cuestión, además de reportar un crecimiento económico, es capaz de representar un impacto social favorable y de no impactar de manera agresiva o desfavorable sobre los recursos del medio ambiente que emplea, lo que en otras palabras se resume en como evaluar si el proyecto es o no sostenible [1, 11].

Por tanto, la sostenibilidad de un proyecto constituye un criterio esencial para evaluar su calidad. Solo aquellos proyectos que introduzcan cambios equitativos y aborden de forma duradera las causas de la vulnerabilidad estructural contribuirán a generar sistemas de sustento sostenibles y un desarrollo humano también sostenible. Para lograrla es preciso que las instituciones públicas, la comunidad o las familias destinatarias se impliquen y asuman la responsabilidad en el mantenimiento o gestión de las infraestructuras y bienes creados (que los bosques sean conservados, los sistemas de irrigación mantenidos, las carreteras reparadas, etc.).

La evaluación de la sostenibilidad de un proyecto se realiza a partir del impacto que se produce o se puede producir no sólo en el ámbito económico, sino también en lo social y en lo medioambiental [8].

Es difícil decidir en la evaluación de un proyecto si éste es o no sostenible de manera absoluta. Además, para hacer dicha evaluación, se parten de unos indicadores que permiten dar un valor al impacto producido en los factores sociales, económicos y medioambientales. Todas estas valoraciones no siempre tienen valores concretos, de ahí que la inclusión de la lógica borrosa en la sostenibilidad del proyecto sea tan apropiada y sea el elemento principal de nuestro estudio.

Hemos desarrollado un sistema en Java, y hemos utilizado la herramienta XFuzzy 3.0 para el diseño inicial de los conjuntos borrosos y para la implementación del motor de inferencia con reglas [5]. Debido a la gran cantidad de combinaciones posibles en algunas de las reglas, se ha tenido que realizar un desarrollo propio de generación de reglas y almacenarlas en ficheros de librerías para no saturar el sistema.

En el primer apartado es una introducción de porqué aplicar la lógica borrosa a la sostenibilidad. En la siguiente sección hacemos una introducción a la sostenibilidad y la evaluación de proyectos. El punto tres habla de los conjuntos borrosos que cubren las variables lingüísticas del sistema. El apartado cuarto explicamos que operadores, reglas y métodos de defuzzyficación utilizamos en el estudio. En el quinto se muestran ejemplos realizados y su comparación. Luego comentamos las conclusiones obtenidas durante este estudio y terminamos con los agradecimientos y la bibliografía.

2 EVALUACIÓN DE LA SOSTENIBILIAD

Hasta ahora, la evaluación de la sostenibilidad de un proyecto se realiza basándose en la evaluación de impacto ambiental, mediante métodos poco automatizados y en muchos casos, hasta manuales. La metodología más utilizada es la matriz de Leopold: se basa en una matriz donde en las columnas hay 100 acciones y en las filas 88 factores ambientales. Los cruces son efectos ambientales o posibles impactos. Las cuadrículas del cruce que presenten impactos significativos se dividen con una diagonal marcando en la parte superior la magnitud del impacto, valorada entre 0 y 10, y en la inferior la importancia también en una escala de 0 a 10. Sumando por filas, se obtiene el impacto producido sobre un determinado factor ambiental, y sumando por columnas el impacto producido por una cierta acción. Aunque para la evaluación de impacto ambiental si existe bibliografía al respecto, no hemos encontrado método informático estándar que evalúe la sostenibilidad de un proyecto. Nosotros lo hemos resuelto con lógico borrosa.

En la solución que aportamos, para ver si un proyecto es sostenible, primero hay que evaluar la sostenibilidad social, económica y medioambiental del proyecto, y con estos tres resultados, se determina la sostenibilidad global.

En la Evaluación de Impacto Ambiental [2] se definen las acciones que produce el proyecto, los factores a los que afectan y que indicadores los representan, para que el usuario pueda realizar la evaluación del impacto producido por estas acciones. De esta manera, el usuario asigna valores a los indicadores para la valoración cualitativa y cuantitativa, y con ambas, se evalúa el impacto producido en un factor medioambiental, social o económico.

El cálculo de la evaluación de cualquiera de los tres ámbitos se realiza de la misma forma, tomando como base la normativa existente para la Evaluación de Impacto Ambiental: primero hemos obtenido los diferentes factores sociales, medioambientales y económicos sobre los que puede afectar un proyecto. El usuario puede asignar un peso a los factores y a los indicadores que afectan a cada uno de ellos según el tipo de proyecto. Una vez identificado el proyecto e indicado el tipo, se habilitarán los efectos que son potencialmente víctimas de impacto. Dentro de cada efecto, el usuario debe asignar los valores a los diferentes indicadores según el daño o beneficio causado para evaluar la magnitud. El siguiente paso es evaluar la importancia, a partir de sus propiedades (extensión, recuperabilidad, signo, periodicidad, reversibilidad, momento, acumulación, intensidad, persistencia y efecto).

Con la importancia y la magnitud, se obtiene el impacto producido en cada uno de los factores de cada ámbito. Con estas valoraciones de los impactos, se obtiene la sostenibilidad por ámbitos, sostenibilidad social, económica y medioambiental, y a partir de estos tres datos, la sostenibilidad global del proyecto, pudiendo ver de manera sencilla, los puntos débiles de los proyectos.

Como hemos comentado, nuestro sistema ofrece la posibilidad de definir los indicadores para las dos valoraciones (cualitativa y cuantitativa). Dependiendo del tipo de proyecto, pueden utilizarse unos indicadores u otros, pudiendo indicar a su vez al usuario el peso con el que afecta un factor a un tipo de proyecto, y el peso con en el que afecta un indicador al mismo.

Con el fin de construir una herramienta válida para cualquier tipo de proyecto, existe un apartado de configuración en el que el usuario puede definir que indicadores usar y asociarlos a los diferentes factores del ámbito social, económico y medioambiental. Además, para cada tipo de proyecto, se puede indicar el peso de cada indicador y de cada factor, pues no se deben tener en cuenta en la misma medida los mismos indicadores y factores para la construcción de un puente sobre un río, que para la construcción de una fábrica de cerveza en un polígono industrial a 200 metros del núcleo urbano.

En el sistema que hemos desarrollado, hemos aplicado la lógica borrosa también a la evaluación de la magnitud (valoración cuantitativa), de tal forma que existan conjuntos borrosos para los indicadores y para que el cálculo de la magnitud sea una relación borrosa entre todos los indicadores de un factor [4]. Si un usuario añade un nuevo indicador, el sistema lo incorpora, creando conjuntos borrosos para las etiquetas lingüísticas que lo caracterizan y mostrándoselos al usuario para que éste pueda realizar una valoración dentro de la lógica borrosa, quedando resuelto el problema de la ambigüedad y falta de cotas en la valoración de los indicadores. El cálculo de la magnitud del impacto (evaluación cuantitativa) se realiza igualmente con una operación borrosa entre los conjuntos borrosos de los indicadores. Dependiendo del número de indicadores, que puede ser diferente de un tipo de proyecto a otro (puentes, túneles, centros comerciales…), se aplican unas

reglas u otras, variando incluso la manera de aplicar dichas reglas, pudiendo agrupar los indicadores y obteniendo el resultado en forma de árbol.

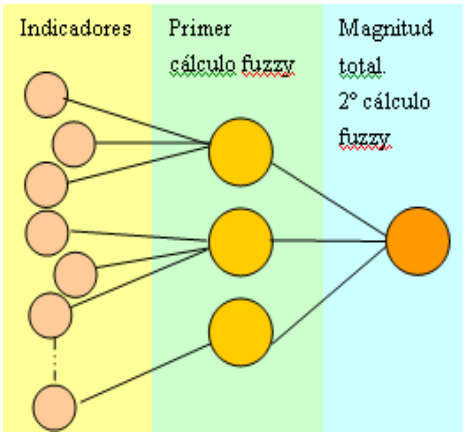


Figura 2.1: Cálculo de la magnitud del impacto en un factor

Para la evaluación cualitativa (importancia) de un impacto sobre un factor, nos hemos basado en la Evaluación de Impacto Ambiental, valorando las propiedades definidas en la Ley [9, 10] (extensión, recuperabilidad, signo, periodicidad, reversibilidad, momento, acumulación, intensidad, persistencia y efecto), de las cuales todas menos el signo son variables lingüísticas. El signo se utiliza para darle sentido positivo o negativo a los valores. Además de estos valores, en este trabajo hemos incorporado otras variables lingüísticas a tener en cuenta: espacios nacionales protegidos, patrimonio cultural, superación de niveles críticos de indicadores que catalogan el proyecto directamente como "No Sostenible".

El resultado de la sostenibilidad se muestra con los términos lingüísticos: compatible, moderado, severo y crítico, dando como resultado un informe, que se puede descargar en formato PDF. En dicho informe se muestran las diferencias de usar unos operadores u otros. El sistema permite utilizar varios tipos de operadores, los diferentes tipos de implicación (S-implicación, residuada y QM-implicación), y método de defuzzyficación, para poder comparar unos resultados con otros, convirtiendo esta herramienta, no sólo en una herramienta útil para la evaluación de proyectos, sino en una herramienta potente para estudio de la lógica borrosa. En cada una de las entradas, se representa gráficamente el grado de pertenencia de cada entrada a los conjuntos borrosos.

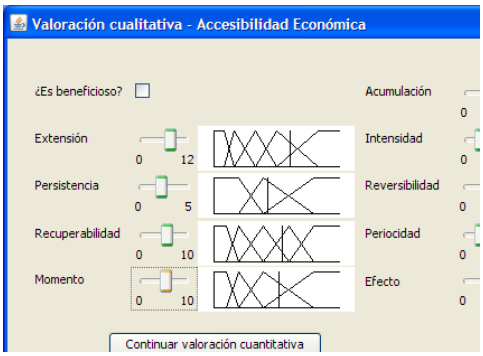


Figura 2.2: Fragmento de pantalla de la aplicación en la parte en que se ve el grado de pertenencia de cada valor a los conjuntos borrosos definidos en el sistema

3 CONJUNTOS BORROSOS

Hemos realizado un estudio para conseguir la evaluación de la sostenibilidad de proyectos aplicando la lógica borrosa, pudiendo utilizar datos no concretos y subjetivos. En el sistema que hemos implementado se realizan varios cálculos borrosos en diferentes puntos de la evaluación:

Cálculo de importancia (valoración cualitativa), cálculo de magnitud (valoración cuantitativa), impacto (a partir de la importancia y magnitud, impacto en un factor), sostenibilidad parcial (social, económica y medioambiente) y sostenibilidad global. Existen puntos donde el cálculo con una sola operación difusa es muy pesado para nuestros sistemas, por lo que aplicando técnicas de descomposición y reducción en problemas más sencillos, se dividen las entradas realizando cálculos parciales y, a partir de ahí, se hacen los cálculos combinando los resultados, y siempre aplicando reglas difusas.

Se definen conjuntos borrosos sobre los indicadores, su magnitud y su importancia. Hacemos una descripción de algunos de estos cálculos y mostramos sus conjuntos borrosos.

En la operación de cálculo de la magnitud, cada indicador se mide con los conjuntos borrosos definidos en la Figura 3.1 y el resultado nos da el valor de la magnitud, que será entrada de la siguiente operación "cálculo de impacto en un factor".

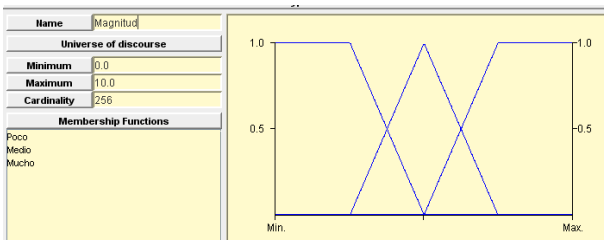


Figura 3.1: Conjuntos borrosos para los términos lingüísticos son "Poco", "Medio" y "Mucho" de un indicador.

En la operación que evalúa el impacto producido por las acciones del proyecto en un factor, intervienen como entrada la magnitud y la importancia. Esta se calcula a partir de sus parámetros (signo, momento, extensión, efecto, persistencia, recuperabilidad, reversibilidad, extensión, acumulación y periodicidad), cada uno con sus conjuntos borrosos, y obteniendo la importancia. Con la importancia y la magnitud, se obtiene el impacto producido en dicho factor.

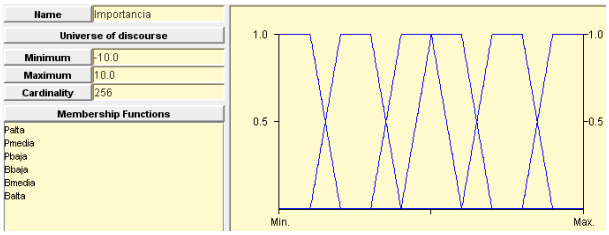


Figura 3.2: Conjuntos borrosos de la variable lingüística Impacto. Los términos lingüísticos son "Palta" perjuicio alto), "Pmedia", "Pbaja", "Bbaja" (Beneficio alto), "Bmedia" y "Balta". Con el fin de utilizar la variable "signo" y usarlo como constante multiplicativa (-1, +1), el número de conjuntos borrosos es no impar y el valoración media es ambigua, ya que es BBaja(0.0) y PBaja (-0.001).

Una vez que se ha calculado el impacto, el sistema realiza la siguiente operación, que consiste en el cálculo de la sostenibilidad parcial a partir del impacto en cada uno de los factores. Así, se podrá calcular la sostenibilidad social a partir de los impactos producidos en cada uno de los factores sociales. Lo mismo en el ámbito del medio ambiente y del ámbito económico.

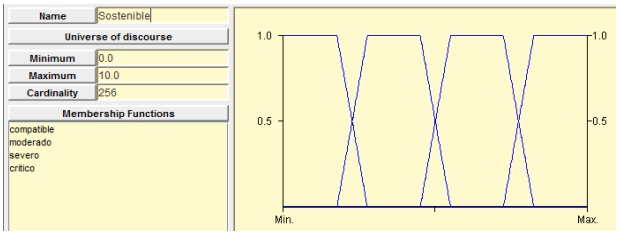


Figura 3.3: Conjuntos borrosos de la variable lingüística Sostenibilidad para evaluar si es "Compatible", "Moderado", "Severo" y "Crítico". Son las cuatro posibles repuestas a si un proyecto es sostenible

4 OPERADORES, REGLAS Y DEFUZZYFICACIÓN

4.1. OPERADORES Y DEFUZZYFICACIÓN

Para definir la intersección y unión de conjuntos borrosos (las premisas de las reglas son conjunciones de varios conjuntos borrosos), se utilizan las t-normas y t-conormas, cuyos estudios habían sido realizados por Karl Menger [7] en "Statistical Metrics", y posteriormente axiomatizadas por B. Schweizer y A. Sklar [12].

Los operadores utilizados por Zadeh [14] para definir la intersección de conjuntos borrosos son el mínimo y el producto. Nosotros hemos incluido el operador de Lukasiewicz por satisfacer propiedades algebraicas diferentes.

El usuarios selecciona que operadores utilizar en la evaluación. Además, en el informe de resultado se muestra una comparación de los resultados obtenidos en las operaciones de cálculo de importancia e impacto con los diferentes operadores.

Para las implicaciones también utilizamos distintos operadores. En la aplicación se puede escoger entre operadores de implicación residual de t-normas, S-implicaciones y QM-implicaciones [13].

Desde la aplicación, también se pueden seleccionar el método de defuzzyficación que se utiliza en nuestro sistema. Gracias a la defuzzyficación, podemos, a partir del conjunto borroso que representa el resultado, obtener un valor nítido y dar un valor a cada resultado de cada operación borrosa. Los métodos de defuzzyficación que hemos utilizado son: centro de masas, medio de los máximos, primer máximo y último máximo

4.2. REGLAS DE INFERENCIA

El motor de inferencia de nuestro sistema se basa en un conjunto de reglas de inferencia borrosa. La definición de las reglas del motor de inferencia se implementan utilizando la herramienta XFuzzy. En este paso, nos hemos encontrado con el problema del gran número de reglas en algunas operaciones debido al uso de muchos conjuntos borrosos para una variable lingüística o por el número de variables utilizadas en las premisas. Para solventar este problema, hemos hecho tres pasos: primero hemos creado una herramienta para crear todas las reglas y guardarlas en un archivo; el segundo paso es modificar el resultado obtenido de XFuzzy y adaptarlo para que lea las reglas del fichero, y el tercer paso es la técnica descomposición de tamaño, en la que hemos divido los cálculos para combinarlos hasta llegar a un resultado final.

En este trabajo hemos desarrollo un mecanismo automático por el que según el número de indicadores que están habilitados por cada tipo de proyecto, se generan las reglas automáticamente. El mecanismo consiste en, si son menos de 5 indicadores, se generan reglas para resolver estas operaciones con lógica borrosa. Si son mayores que cinco, simplificamos en cálculos de menor tamaño. En la Figura 4.1 mostramos un ejemplo de reglas.

```
if(acumulacion == Simple & extension == Puntual
& intensidad == Baja & persistencia == Fugaz &
reversibilidad == Largo plazo & recuperabilidad
== Mitigable & periodicidad == Periodico & mo-
```

```
mento == Critico & efecto == Indirecto tercia-
rio) -> valorCualitativo = media;

if(acumulacion == Simple & extension == Puntual
& intensidad == Baja & persistencia == Fugaz &
reversibilidad == Largo plazo & recuperabilidad
== Mitigable & periodicidad == Continuo & momen-
to == Largo plazo & efecto == Directo) -> valor-
Cualitativo = baja;

if(acumulacion == Simple & extension == Puntual
& intensidad == Baja & persistencia == Fugaz &
reversibilidad == Largo plazo & recuperabilidad
== Mitigable & periodicidad == Continuo & momen-
to == Inmediato & efecto == Directo) -> valor-
Cualitativo = baja;
```

Figura 4.1: Reglas de inferencia para la valoración cualitativa de un impacto en un factor.

5 EJEMPLOS

Hemos realizado tres pruebas con esta aplicación, Construcción de la carretera M-61, de sostenibilidad crítica, construcción de planta de fabricación de cerveza y cambio de red eléctrica, ambos con sostenibilidad moderada. La valoración real de los proyectos se puede observar en los documentos oficiales y públicos correspondientes: *http://www.ssreyes.org/es/portal.do?NM=3&IDM=139*, *http://www.madrid.org/cs/Satellite?blobcol=paginapdf&blobheade-de-r=application%2Fpdf&blobkey=id&blobtable=CM_Pagina_BOCM&blobwhere=1142456839486&ssbinary=true* y *http://www.boe.es/diario_boe/txt.php?id=BOE-A-2006-2179* respectivamente.

Vamos a mostrar cómo se analizó el caso de la construcción de la M-61 y comentando sus conclusiones.

Primero se introduce la descripción del proyecto. Entre estos datos están la localidad, el nombre del proyecto, la empresa que lo va a realizar, el tipo de financiación y el tipo de proyecto. Esto último activa los factores e indicadores que estarán habilitados y con qué peso cada uno.

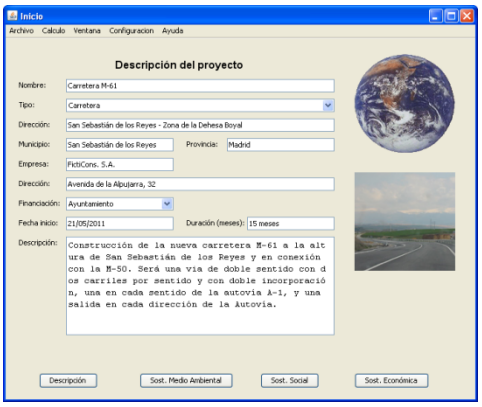


Figura 5.1: Ventana de descripción del proyecto

A continuación, el usuario debe pasar a las pestañas de sostenibilidad medio ambiental, sostenibilidad social y sostenibilidad económica. En cada una de estas pantallas, se muestran los factores para los que se analiza el impacto producido.

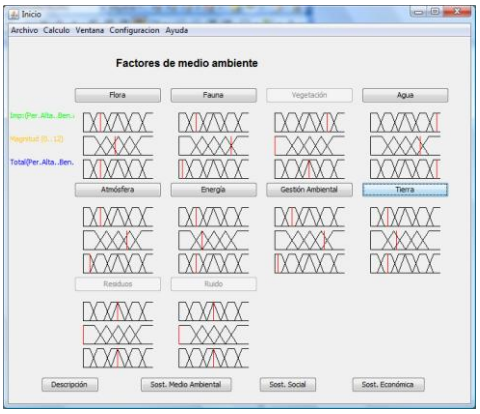


Figura 5.2: Factores de medio ambiente. Por factor, se muestra la importancia, la magnitud y el impacto producido en dicho factor.

Para que los datos de importancia y magnitud se vean completados, antes hay que valorar los datos de cada factor, que son los indicadores para la magnitud y las propiedades de la importancia, que se muestra en la figura 5.2.

El resultado final indica que el proyecto no es sostenible, dando un valor de sostenibilidad crítico, como se ve en la Figura 5.3. Dichos resultados finales se resumen en un informe en formato Pdf.

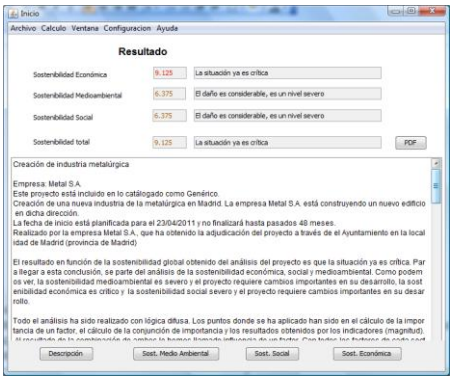


Figura 5.3: Resultado final de un ejemplo de cálculo de sostenibilidad

6 CONCLUSIONES

En este trabajo se presenta una metodología para la evaluación de la sostenibilidad social, económica y medioambiental de proyectos y de esta manera obtener una sostenibilidad global. Tradicionalmente la evaluación de proyectos se realiza solo desde el punto de vista económico. El hecho de que en el cálculo de la sostenibilidad se utilizan términos ambiguos ha favorecido el uso de la lógica borrosa frente a los métodos tradicionales de análisis de impacto. Se aplica la lógica borrosa a cada uno de los cálculos que se realizan en el análisis global de la sostenibilidad: valoración de indicadores, la evaluación cuantitativa y cualitativa del impacto que produce un proyecto en los diferentes factores medioambientales, sociales y económicos.

La posibilidad de un cálculo de sostenibilidad de proyectos que sea universal es una idea que de momento es lejana. No obstante el poder acercarse a ella puede convertir a un proyecto en una herramienta imprescindible de acuerdo a la idea de sostenibilidad global de la Agenda 21, en las Cumbres de la Tierra y en las Conferencias Europeas sobre Ciudades Sostenibles.

El software realizado se puede descargar para su uso en

www.fdi.ucm.es/profesor/lgarmend/SC/sostenibilidad.jar

7 AGRADECIMIENTOS

Agradecemos al ayuntamiento de San Sebastián de los Reyes por la información de los proyectos evaluados de la carretera M61 y de la fábrica de cerveza.

Agradecemos el soporte del proyecto TIN2009-07901, CAM GR35/10-A.

Agradecemos a Adela Salvador su soporte y generosidad aportando su experiencia en evaluación de impacto.

Referencias

[1] Baylis, J. y Smith, S. La globalización de la política mundial. (3ª ed). Oxford University Press. 454-455, 2005.

[2] Gallego, E.; González de Paula, L.; Garmendia, L. y Garmendia, A. "*Método de decisión borrosa de si un efecto es impacto ambiental y su carácter*". Jornadas internacionales de Didáctica de las Matemáticas en Ingeniería. ETSI Caminos, UPM. 2009

[3] Garmendia, A.; Salvador, A.; Crespo, C. y Garmendia, L. *Evaluación de impacto ambiental*. Prentice Hall. 2005.

[4] Gómez, R. A. *Lógicas no clásicas: principios y fundamentos*. Universidad EAFIT Escuela de Ciencias y Humanidades. Colección Académica. 2006.

[5] IMSE Centro Nacional de Microelectrónica. *Herramientas de CAD para Lógica Difusa. XFuzzy. 3.0*. http://www.imse.cnm.es. 2003.

[6] Martín, R. Directores: Garmendia L, Caro R., Garmendia A. Gestión de la Sostenibilidad utilizando Lógica Borrosa. *Trabajo Fin de Master de Investigación en Informática*. Facultad de Informática. UCM. 2011.

http://eprints.ucm.es/13054/1/memoriaSostFuzzy_5.0.0.pdf

[7] Menger, K. Statistical Metrics. En: *Proceedings of the National Academy of Sciences of the United States of America*. 1942.

[8] *Primera Conferencia Europea sobre Ciudades Sostenibles*. Carta de Aalborg. 1994

[9] *Real Decreto 1131/1988*, de 30 de septiembre, por el que se aprueba el Reglamento para la ejecución del Real Decreto legislativo 1302/1986, de 28 de junio, de evaluación de impacto ambiental.

[10] *Real Decreto Legislativo 1/2008*, de 11 de enero, por el que se aprueba el texto refundido de la Ley de Evaluación de Impacto Ambiental de proyectos.

[11] *Sexta Conferencia Europea sobre Ciudades Sostenibles*. Conferencia de Aalborg +10. 2004. http://www.aalborgplus10.dk

[12] Schweizer B. y Sklar. A. Probabilistic Metric Spaces, North-Holland, Amsterdam. 1983.

[13] Trillas, E. y Valverde, L. On mode and implication in approximate reasoning. *Approximate reasoning in expert systems*. Eds. M. M. Gupta. North-Holland. 157-166. 1985.

[14] Zadeh, L. A. Fuzzy sets. En *Proceedings of the IEEE (Information and Control)*, Vol. 8., pp.338-353, 1965.

USO DE OPERADORES OWA PARA MODELAR LA ACTITUD HACIA EL CONSENSO EN PROBLEMAS DE TOMA DE DECISIÓN EN GRUPO

Iván Palomares[1], **Jun Liu**[2], **Yang Xu**[3], **Luis Martínez**[1]

[1]*Depto. de Informática, Universidad de Jaén*, {ivanp,martin}@ujaen.es
[2]*School of Computing and Mathematics, University of Ulster*, j.liu@ulster.ac.uk
[3]*School of Mathematics, Southwest Jiaotong University*, xuyang@home.swjtu.edu.cn

Resumen

El uso de procesos de *consenso* para resolver problemas de Toma de Decisión en Grupo busca alcanzar un acuerdo mutuo antes de tomar una decisión. Un aspecto clave en estos procesos es la elección de una medida de consenso adecuada a las necesidades del grupo y al contexto del problema. Dado que dichas medidas implican normalmente agregación de información, en esta contribución presentamos una extensión del operador OWA, llamada *Attitude*-OWA (AOWA), para reflejar diferentes actitudes del grupo hacia el consenso.

Palabras Clave: Toma de decisión, consenso, OWA, agregación, actitud.

1 INTRODUCCIÓN

Los problemas de Toma de Decisión en Grupo (TDG) son habituales en la sociedad y en muchas organizaciones, donde varios individuos deben alcanzar una solución común a un problema consistente en una serie de alternativas, sobre las cuales cada individuo expresa sus preferencias [6]. Tradicionalmente, estos problemas se han resuelto mediante enfoques clásicos de TDG, tales como la unanimidad, regla de la mayoría, etc. [2]. Sin embargo, en muchos contextos es importante que el grupo alcance un acuerdo mutuo antes de tomar la decisión, de manera que todos los miembros acepten la solución obtenida. Este acuerdo puede conseguirse mediante el uso de procesos de *consenso* [12].

Diferentes interpretaciones se han propuesto en la literatura para el concepto de consenso, destacando el llamado enfoque de *soft consensus* basado en el concepto de mayoría difusa [6]. Asimismo, distintos modelos de consenso han sido propuestos, tales como modelos con información lingüística [3], modelos basados en estructuras de preferencia incompletas [5], modelos de consenso adaptativos [9], etc. Sin embargo, existen algunos aspectos relativos a los procesos de consenso que aún requieren un mayor estudio, como por ejemplo las *medidas de consenso* para determinar el nivel de acuerdo, las cuales suelen implicar una agregación de información [1]. Así, sería interesante definir un operador de agregación que permita reflejar la *actitud* del grupo hacia el consenso, para conocer la capacidad de cambio a la que están dispuestos los expertos para alcanzar un acuerdo y de esta forma optimizar el proceso de consenso.

En esta contribución presentamos una extensión del operador OWA [13] a la que llamaremos AOWA (*Attitude-OWA*), que permite integrar la actitud en la agregación de información en problemas de consenso en TDG. Dicho operador se basa en la medida de su grado de optimismo (*orness*), y en el empleo de cuantificadores lingüísticos [15], definidos a partir de la actitud deseada, para el cálculo de pesos. Además, se mostrará una aplicación práctica del operador en la resolución mediante consenso de un problema de TDG con un número de expertos elevado.

Este trabajo se estructura como sigue: en la Sección 2 se introducen los preliminares necesarios. En la Sección 3, presentamos el operador AOWA y mostramos cómo reflejar la actitud del grupo hacia el consenso. En la Sección 4 se presenta una simulación donde se integra la actitud en un sistema de apoyo al consenso. Por último, concluimos el trabajo en la Sección 5.

2 PRELIMINARES

En esta sección revisaremos los problemas de TDG y los procesos de consenso, además del operador OWA.

2.1 PROCESOS DE CONSENSO EN TDG

Formalmente, un problema de TDG se compone de los siguientes elementos [2, 6]:

- Un conjunto de *alternativas* o posibles soluciones al problema, $X = \{x_1, \ldots, x_n\} (n \geq 2)$.
- Un conjunto de individuos o *expertos*, que dan su opinión sobre las alternativas en X mediante una estructura de preferencia, $E = \{e_1, \ldots, e_m\} (m \geq 2)$.

Una de las estructuras de preferencia más utilizadas en TDG bajo incertidumbre es la *relación de preferencia di-*

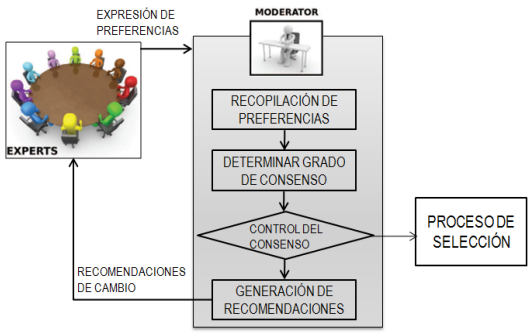


Figura 1: Esquema general del proceso de consenso.

fusa [4]. Una relación de preferencia difusa P_i asociada al experto e_i se caracteriza por una función de pertenencia $\mu_{P_i} : X \times X \to [0,1]$, y se representa, siempre que X sea finito, mediante una matriz de dimensión $n \times n$ como:

$$P_i = \begin{pmatrix} p_i^{11} & \cdots & p_i^{1n} \\ \vdots & \ddots & \vdots \\ p_i^{n1} & \cdots & p_i^{nn} \end{pmatrix}$$

donde cada valoración $p_i^{lk} = \mu_{P_i}(x_l, x_k)$ representa el grado de preferencia de la alternativa x_l sobre x_k según el experto e_i, de forma que $p_i^{lk} > 0.5$ indica preferencia de x_l sobre x_k, $p_i^{lk} < 0.5$ indica preferencia de x_k sobre x_l, y $p_i^{lk} = 0.5$ indica indiferencia entre x_l y x_k.

Tradicionalmente, el proceso de resolución de problemas de TDG se compone de una *fase de agregación*, en la que se combinan las preferencias de los expertos, y una *fase de explotación*, donde se obtiene una alternativa o subconjunto de alternativas como solución al problema [11].

Al resolver problemas de TDG puede aparecer el problema de que algunos expertos no estén de acuerdo con la decisión tomada porque piensen que sus opiniones no han sido suficientemente consideradas. Los *procesos de consenso* surgieron para solventar este problema, mediante la discusión de los expertos y la modificación de sus preferencias, con el objetivo de acercarlas entre sí y conseguir un nivel de acuerdo suficiente antes de tomar la decisión [12]. Existen diferentes interpretaciones de consenso, desde una visión rígida de consenso como acuerdo total (unanimidad) a otra más flexible que considera diferentes grados de acuerdo parciales [2]. Uno de los enfoques flexibles de consenso más conocidos es el de *soft consensus* [6], propuesto por Kacprzyk y basado en el concepto de mayoría difusa, según el cual existe consenso cuando "la mayoría de expertos participantes coincide en su opinión sobre las alternativas importantes".

El consenso es un proceso dinámico e iterativo, normalmente coordinado por una figura conocida como *moderador*, que evalua el nivel de acuerdo alcanzado e identifica aquellos expertos y opiniones que impiden alcanzar el consenso [8]. La Figura 1 muestra un esquema general del proceso de consenso [9]:

1. *Expresión y Recopilación de preferencias*: Cada experto e_i proporciona al moderador una relación de preferencia difusa P_i sobre X.
2. *Determinar grado de consenso*: Para cada par de expertos e_i, e_j $(i < j)$, se calcula un valor de similitud $s_{ij}^{lk} \in [0,1]$ en cada par de alternativas (x_l, x_k) como
$$s_{ij}^{lk} = 1 - |p_i^{lk} - p_j^{lk}| \quad (1)$$
Los valores de similitud se agregan para obtener el grado de consenso sobre cada par de alternativas c^{lk},
$$c^{lk} = \phi(s_{12}^{lk}, \ldots, s_{1m}^{lk}, s_{23}^{lk}, \ldots, s_{2m}^{lk}, \ldots, s_{(m-1)m}^{lk}) \quad (2)$$
donde la función $\phi : [0,1]^{\frac{m(m-1)}{2}} \to [0,1]$ representa el operador de agregación utilizado. Seguidamente, los grados de consenso c^{lk} se agregan para obtener un grado de consenso global $cr \in [0,1]$,
$$cr = \frac{\sum_{l=1}^{n} \sum_{k=1, k \neq l}^{n} c^{lk}}{n(n-1)} \quad (3)$$
3. *Control del consenso*: Se compara cr con el umbral de consenso mínimo deseado, μ. Si $cr \geq \mu$, se ha alcanzado consenso y el grupo pasa al proceso de selección; en caso contrario, se procede a la siguiente fase.
4. *Generación de Recomendaciones*: El moderador determina aquellas valoraciones, p_i^{lk}, de los expertos e_i que se encuentran más alejadas del consenso, y recomienda a dichos expertos su modificación para aumentar el grado de acuerdo.

2.2 OPERADORES OWA. CÁLCULO DE PESOS

Los operadores OWA (*Ordered Weighted Averaging*), propuestos por Yager [13], suponen una de las familias de operadores de agregación ponderados más utilizados en toma de decisiones.

Definición 1. *Dado un conjunto* $A = \{a_1, \ldots, a_n\}$, $a_i \in [0,1]$, *un operador OWA de dimensión n se define como una función* $OWA_W : [0,1]^n \to [0,1]$, *con un vector de pesos* $W = [w_1 w_2 \ldots w_n]^\top$, *donde* $w_i \in [0,1]$, $\sum_i w_i = 1$ *y*

$$OWA_W(a_1, \ldots, a_n) = \sum_{j=1}^{n} w_j b_j \quad (4)$$

donde b_j *es el* j*-ésimo mayor valor en* A.

Los operadores OWA son idempotentes, continuos, monótonos, neutros y compensativos [1]. Además, se pueden clasificar mediante el concepto *orness*, que es una medida de optimismo comprendida en el intervalo unitario, que nos indica a partir de W cuan cercano el operador se encuentra al operador *máximo* (OR) o al operador *mínimo* (AND).

$$orness(W) = \frac{1}{n-1} \sum_{i=1}^{n} (n-i) w_i \quad (5)$$

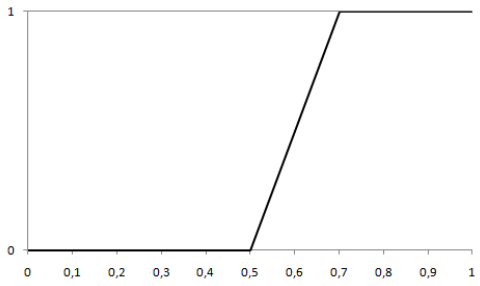


Figura 2: Cuantificador $Q =$"*la mayoría*".

Un operador OWA se puede considerar optimista si $orness(W) > 0.5$, pesimista si $orness(W) < 0.5$ y neutro si $orness(W) = 0.5$. Otra medida para distinguir entre operadores OWA con igual grado de optimismo es la *dispersión* o entropía, que indica la cantidad de información de A que es considerada en la agregación. Esta medida toma valores en el intervalo $[0, ln(n)]$, y se calcula como

$$Disp(W) = -\sum_{i=1}^{n} w_i \ln w_i \tag{6}$$

Existen diferentes métodos para determinar los pesos de un operador OWA, como por ejemplo a partir de cuantificadores lingüísticos [15]. En este trabajo utilizaremos un tipo de cuantificadores relativos conocidos como RIM (*Regular Increasing Monotone*) para el cálculo de pesos [7], ya que debido a su semántica implícita son apropiados para reflejar la noción de consenso como mayoría difusa [6].

Un cuantificador RIM Q es un conjunto difuso en [0,1], donde dada una proporción $r \in [0,1]$, $Q(r)$ indica el grado en que r satisface la semántica definida en Q. La Figura 2 muestra un ejemplo de cuantificador $Q =$*"la mayoría"*. Los cuantificadores RIM presentan las siguientes propiedades: (i) $Q(0) = 0$, (ii) $Q(1) = 1$ y (iii) si $r_1 > r_2$ entonces $Q(r_1) \geq Q(r_2)$. Para un cuantificador RIM lineal, su función de pertenencia $Q(r)$ se define a partir de dos parámetros $\alpha, \beta \in [0,1]$, $\alpha < \beta$ como

$$Q(r) = \begin{cases} 0 & \text{si } r \leq \alpha, \\ \frac{r-\alpha}{\beta-\alpha} & \text{si } \alpha < r \leq \beta, \\ 1 & \text{si } r > \beta. \end{cases} \tag{7}$$

Basándose en $Q(r)$, Yager propuso en [13] el siguiente método para calcular los pesos de un operador OWA mediante cuantificadores RIM:

$$w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right), i = 1, \ldots, n \tag{8}$$

Dado un cuantificador RIM Q, sea $[\alpha, \beta] \subseteq [0,1]$ el intervalo de crecimiento en la función de pertenencia Q(r) (tal que $Q(r) = 0, \forall r \leq \alpha$ y $Q(r) = 1, \forall r \geq \beta$), siendo

$$d = \beta - \alpha \tag{9}$$

la amplitud de dicho intervalo.

3 INTEGRACIÓN DE LA ACTITUD EN OPERADOR AOWA: APLICACIÓN EN PROCESOS DE CONSENSO

En esta sección se definirá un operador de agregación que nos permita reflejar la actitud del grupo al aplicar dicho operador a problemas de consenso en TDG.

3.1 OPERADOR AOWA (*Attitude*-OWA)

El procedimiento para reflejar la actitud en una agregación puede llevarse a cabo desde varios puntos de vista. En esta contribución, consideramos que la especificación del valor en la medida $orness(W)$ del operador OWA puede ayudarnos, al indicar el carácter más o menos optimista del operador, es decir, la actitud [1]. El operador AOWA será una extensión del OWA basado en la medida $orness(W)$ y en el uso de cuantificadores RIM para el cálculo de pesos.

Mediante los valores de $orness(W)$ y la amplitud d, que indica los valores que reciben un peso no nulo al agregar y se relaciona con la dispersión del operador OWA correspondiente (véase Sección 2.2), definimos el operador AOWA:

Definición 2. *Un operador AOWA de dimensión n en* $A = \{a_1, \ldots, a_n\}$ *es un operador OWA basado en dos parámetros de actitud* ϑ, φ*, tal que*

$$AOWA_W(A, \vartheta, \varphi) = \sum_{j=1}^{n} w_j b_j \tag{10}$$

donde b_j *es el* j*-esimo mayor valor en A;* $\vartheta = orness(W)$*,* $\varphi = d$ *(ver ec. (9)) son los parámetros que permiten expresar la actitud y W es el vector de pesos, obtenidos mediante un cuantificador RIM definido a partir de* $\vartheta, \varphi \in [0,1]$*.*

El operador AOWA mantiene las propiedades de idempotencia, continuidad, monotonía, neutralidad y compensación, propias del operador OWA. Un aspecto importante para la utilización de este operador es la relación entre los parámetros de actitud ϑ, φ y el cuantificador RIM asociado a dicha actitud, la cual mostramos a continuación.

Teorema 1. Dado un cuantificador RIM con una función de pertenencia $Q(r)$ diferenciable, el grado de optimismo o actitud del operador AOWA basado en dicho cuantificador se obtiene como

$$\vartheta = \int_0^1 Q(r)dr \tag{11}$$

Demostración. Basándonos en las ecs. (5) y (8), se tiene

$$\vartheta(n) = orness(W)(n) = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)\left[Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right)\right],$$

cuando n es suficientemente grande ($n \to \infty$),

$$\begin{aligned}\vartheta = \lim_{n\to\infty} \vartheta(n) &= \lim_{n\to\infty} \frac{1}{n-1}\sum_{i=1}^{n}(n-i)\left[Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right)\right] \\ &= \lim_{n\to\infty} \frac{1}{n-1}\sum_{i=1}^{n-1}(n-i)\left[Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right)\right]\end{aligned}$$

Considerando $P = \sum_{i=1}^{n-1}(n-i)\left[Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right)\right]$ y expandiendo los sumatorios, se obtiene

$$P=\sum_{i=1}^{n-1}\left[n\left[Q\left(\frac{i}{n}\right)-Q\left(\frac{i-1}{n}\right)\right]-i\left[Q\left(\frac{i}{n}\right)-Q\left(\frac{i-1}{n}\right)\right]\right]$$

$$=n\sum_{i=1}^{n-1}\left[Q\left(\frac{i}{n}\right)-Q\left(\frac{i-1}{n}\right)\right]-\sum_{i=1}^{n-1}i\left[Q\left(\frac{i}{n}\right)-Q\left(\frac{i-1}{n}\right)\right]$$

$$=nQ\left(\frac{n-1}{n}\right)-\left[-\left[\sum_{i=1}^{n-2}Q\left(\frac{i}{n}\right)\right]+(n-1)Q\left(\frac{n-1}{n}\right)\right]$$

$$=nQ\left(\frac{n-1}{n}\right)-(n-1)Q\left(\frac{n-1}{n}\right)+\sum_{i=1}^{n-2}Q\left(\frac{i}{n}\right)$$

$$=Q\left(\frac{n-1}{n}\right)+\sum_{i=1}^{n-2}Q\left(\frac{i}{n}\right)=\sum_{i=1}^{n-1}Q\left(\frac{i}{n}\right)$$

Luego,

$$\vartheta(n)=\frac{1}{n-1}\sum_{i=1}^{n-1}Q\left(\frac{i}{n}\right)$$

cuando $n\to\infty$, aplicando la definición de integral como límite se tiene [14]:

$$\vartheta=\lim_{n\to\infty}\frac{1}{n-1}\sum_{i=1}^{n-1}Q\left(\frac{i}{n}\right)=\int_0^1 Q(r)dr,$$

donde $r=i/n$. □

Dado que esta afirmación es aplicable a toda función $Q(r)$ diferenciable en $[0,1]$, es posible extenderla a diferentes tipos de cuantificadores, como mostramos a continuación.

Corolario 1. Dado un cuantificador RIM con función de pertenencia lineal como la mostrada en la ec. (7), cuando el número de elementos a agregar n es suficientemente alto, es posible calcular el grado de optimismo ϑ del operador AOWA basado en dicho cuantificador como [7, 14]

$$\vartheta=1-\alpha-\frac{\varphi}{2} \tag{12}$$

Demostración. A partir del teorema anterior, y basándonos en la ec. (7), se tiene

$$\vartheta=\int_0^1 Q(r)dr=\int_\alpha^1 Q(r)dr=A(Q)=\frac{1}{2}(\beta-\alpha)+1-\beta$$

donde el intervalo de integración $(\alpha,1]$ es el *soporte* del cuantificador, y $A(Q)$ es el área bajo $Q(r)$ en dicho intervalo. Por último, teniendo en cuenta que $\beta-\alpha=d=\varphi$,

$$\vartheta=\frac{1}{2}\varphi+1-(\alpha+\varphi)=1-\alpha-\frac{\varphi}{2}$$

se concluye que $\vartheta=1-\alpha-\varphi/2$. □

Este corolario permite que, dados ϑ y φ, sea posible calcular el valor del parámetro α, necesario para definir el cuantificador RIM asociado al operador AOWA, como

$$\alpha=1-\vartheta-\frac{\varphi}{2} \tag{13}$$

La relación entre los parámetros de actitud hace necesario definir una serie de restricciones en sus valores para asegurar la definición de un cuantificador válido a partir de ellos.

Teorema 2. Dado un cuantificador RIM lineal con parámetros $\alpha,\varphi\in[0,1]$, la actitud ϑ del operador AOWA asociado es válida únicamente si $\alpha+\varphi\leq 1$.

Demostración. Suponer $\alpha+\varphi>1$. Sabiendo que $\varphi=\beta-\alpha$, la ec. (12) implica,

$$\vartheta=1-\alpha-\frac{\varphi}{2}=1-\frac{\alpha+\beta}{2}$$

donde $\frac{\alpha+\beta}{2}$ es el valor central del intervalo de crecimiento del cuantificador, $[\alpha,\beta]$, de manera que

$$\alpha\leq\frac{\alpha+\beta}{2}\leq\beta$$

$$1-\alpha\geq 1-\frac{\alpha+\beta}{2}\geq 1-\beta$$

$$1-\alpha\geq\vartheta\geq 1-(\alpha+\varphi)$$

donde $\beta=\alpha+\varphi$. Si $\alpha+\varphi>1$, entonces ϑ puede ser negativo, luego $\alpha+\varphi$ debe ser menor o igual que 1 para asegurar la definición de una actitud válida. □

Para que se cumpla esta condición, las siguientes restricciones serán impuestas al proporcionar el valor de cada parámetro,

Corolario 2. Dado $\varphi\in[0,1]$, el valor de ϑ deberá cumplir

$$\frac{\varphi}{2}\leq\vartheta\leq 1-\frac{\varphi}{2} \tag{14}$$

Demostración. Según la ec. (13), α es negativo si $(\vartheta+\varphi/2)>1$. Es necesario que $\alpha\geq 0$, es decir,

$$1-\vartheta-\frac{\varphi}{2}\geq 0\Rightarrow\vartheta+\frac{\varphi}{2}\leq 1\Rightarrow\vartheta\leq 1-\frac{\varphi}{2}$$

Sin embargo, también es necesario que $\alpha+\varphi\leq 1$ (véase Teorema 2). Considerando de nuevo la ec. (13),

$$\alpha+\varphi=1-\vartheta-\frac{\varphi}{2}+\varphi\leq 1\Rightarrow$$

$$1-\vartheta+\frac{\varphi}{2}\leq 1\Rightarrow\vartheta\geq\frac{\varphi}{2}$$

Teniendo en cuenta ambas desigualdades, se concluye que $\frac{\varphi}{2}\leq\vartheta\leq 1-\frac{\varphi}{2}$. □

Corolario 3. Dado $\vartheta\in[0,1]$, el valor de φ deberá cumplir

$$\varphi\leq 1-|2\vartheta-1| \tag{15}$$

Demostración. Según la demostración del corolario anterior, $\alpha\geq 0$ implica

$$\vartheta+\frac{\varphi}{2}\leq 1\Rightarrow\varphi\leq 2(1-\vartheta)$$

que, pese a ser válido para $\vartheta\in[0.5,1]$, daría lugar a $\varphi>1$ y a violar el Teorema 2 si $\vartheta<0.5$. Considerando dicho teorema y la ec. (13), tenemos

$$\alpha+\varphi=1-\vartheta-\frac{\varphi}{2}+\varphi\leq 1\Rightarrow$$

$$1-\vartheta+\frac{\varphi}{2}\leq 1\Rightarrow\varphi\leq 2\vartheta$$

que es válido para $\vartheta\in[0,0.5]$, pero permite que $\varphi>1$ cuando $\vartheta>0.5$. Por tanto, según el valor de ϑ se deberán cumplir las siguientes desigualdades:

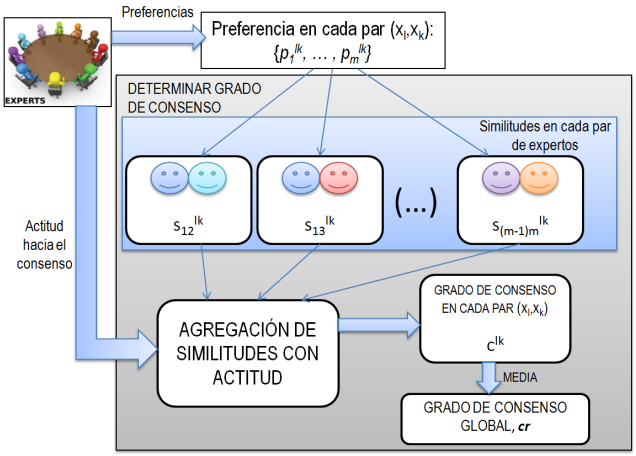


Figura 3: Procedimiento para determinar el grado de consenso basado en la actitud.

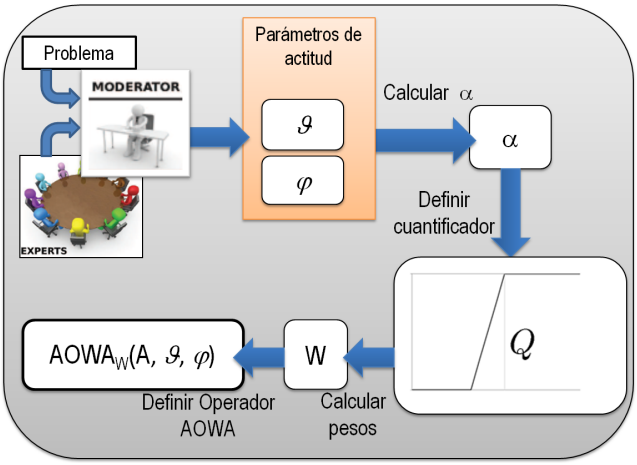


Figura 4: Proceso para definir un operador AOWA a partir de la actitud de grupo hacia el consenso.

$$\begin{cases} \varphi \leq 2\vartheta & \text{si } \vartheta \in [0,0.5] \\ \varphi \leq 2(1-\vartheta) & \text{si } \vartheta \in [0.5,1] \end{cases}$$

Nos planteamos finalmente encontrar una única expresión que integre ambas restricciones. Por un lado, tenemos

$$2\vartheta = 1-(-2\vartheta+1)$$

donde, para $\vartheta \in [0,0.5]$, el término $(-2\vartheta+1) \geq 0$. Por otro lado,

$$2(1-\vartheta) = 1-(2\vartheta-1)$$

donde, para $\vartheta \in [0.5,1]$, el término $(2\vartheta-1) \geq 0$. Por ello, mediante el valor absoluto del término $(2\vartheta-1)$ es posible integrar ambas restricciones como

$$\varphi \leq 1-|2\vartheta-1|$$

ya que $(2\vartheta-1)$ es negativo cuando $\vartheta \in [0,0.5]$. □

Si se respetan las restricciones impuestas en las ecs. (14) y (15), quedará garantizada la correcta definición de un operador AOWA para reflejar la actitud dada por ϑ y φ.

3.2 APLICACIÓN DE AOWA EN PROCESOS DE CONSENSO

Una vez presentado el operador AOWA, en este apartado aplicamos dicho operador en procesos de consenso para la resolución de problemas de TDG con muchos expertos, con el objetivo de reflejar la actitud del grupo hacia el consenso. En este contexto de aplicación, la actitud debe ser interpretada como explicamos a continuación:

- *Actitud Optimista*: Indica que los miembros del grupo se inclinan a cambiar sus preferencias para alcanzar un acuerdo, por lo que aquellas posiciones del grupo donde el nivel de acuerdo es mayor cobrarán mayor importancia en la agregación.
- *Actitud Pesimista*: Indica que los expertos son reacios a cambiar con facilidad sus opiniones para alcanzar un acuerdo, lo cual implica que aquellas posiciones del grupo en mayor desacuerdo cobrarán mayor importancia en la agregación.

La actitud dependerá de los expertos y del problema a resolver. En la Figura 3 se muestra en mayor detalle la fase de medición del grado de consenso explicada en la sección 2.1, incluyendo la agregación en la que proponemos reflejar la actitud mediante el operador AOWA (véase ec. (2)).

En la Figura 4 mostramos un esquema general para integrar la actitud en el proceso de consenso mediante AOWA. La expresión de la actitud por parte del grupo debe realizarse antes de comenzar el proceso de consenso (siendo el moderador responsable de recopilar e integrar los parámetros que definen dicha actitud). Para ello, el moderador puede considerar tanto el contexto y características del problema de decisión a resolver, como la idea de actitud que tenga cada experto, así como decidir si considerar solamente la actitud deseada por un subgrupo de expertos, debido a su nivel de conocimiento sobre el problema o su posición dentro del grupo.

4 SIMULACIÓN EXPERIMENTAL

En esta sección presentamos un ejemplo práctico de aplicación de la actitud hacia el consenso en un Sistema de Apoyo al Consenso desarrollado sobre una plataforma multiagente [10], que nos ha permitido simular la resolución de un problema de TDG mediante consenso, utilizando operadores AOWA basados en distintas actitudes, y con un alto número de expertos en el grupo.

Hipótesis: El empleo de diferentes actitudes hacia el consenso tiene un efecto directo sobre el proceso, donde el optimismo favorece una mayor convergencia al consenso tras un número de rondas de discusión menor, mientras que el pesimismo da lugar una convergencia menor y, por tanto, será necesario un número mayor de rondas de discusión.

El problema de TDG utilizado consiste en 4 alternativas, $X = \{x_1, x_2, x_3, x_4\}$, 50 expertos, $E = \{e_1, \ldots, e_{50}\}$, y un

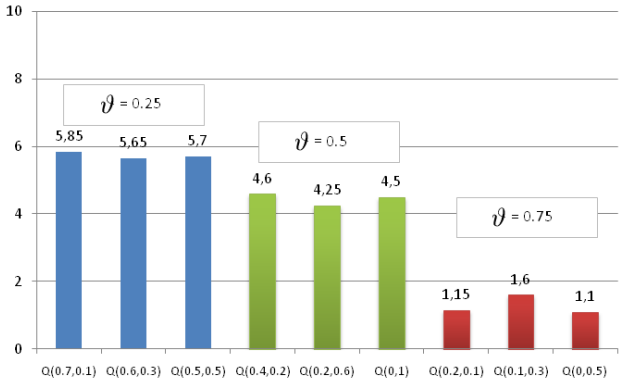


Figura 5: Promedio de rondas de discusión necesarias con operadores AOWA basados en diferentes actitudes dadas por ϑ y φ.

umbral de consenso $\mu = 0.85$. Se han definido nueve actitudes hacia el consenso (véase Tabla 1). Los experimentos han consistido en resolver mediante simulación el problema definido, un total de 20 veces para cada actitud.

Tabla 1: Parámetros de actitud y cuantificadores utilizados.

Actitud	ϑ	φ	α	$Q_{(\alpha,\varphi)}$
Pessimista	0.25	0.1	0.7	$Q_{(0.7,0.1)}$
	0.25	0.3	0.6	$Q_{(0.6,0.3)}$
	0.25	0.5	0.5	$Q_{(0.5,0.5)}$
Indiferente	0.5	0.2	0.4	$Q_{(0.4,0.2)}$
	0.5	0.6	0.2	$Q_{(0.2,0.6)}$
	0.5	1	0	$Q_{(0,1)}$
Optimista	0.75	0.1	0.2	$Q_{(0.2,0.1)}$
	0.75	0.3	0.1	$Q_{(0.1,0.3)}$
	0.75	0.5	0	$Q_{(0,0.5)}$

Los resultados de los experimentos, mostrados en la Figura 5, indican la convergencia hacia el consenso, mediante el número promedio de rondas de discusión necesarias para alcanzar μ con cada operador AOWA. Estos resultados confirman nuestra hipótesis de partida.

5 COMENTARIOS FINALES

En este trabajo, hemos presentado el operador AOWA, una extensión del operador OWA que permite reflejar la actitud en la agregación, y lo hemos aplicado a procesos de consenso para resolver problemas de TDG con un gran número de expertos, proponiendo un método que permite al grupo de decisión reflejar fácilmente la actitud que desean considerar a la hora de medir el consenso en dicho grupo.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los Proyectos de Investigación TIN-2009-08286, P08-TIC-3548 y fondos FEDER.

Referencias

[1] G. Beliakov, A. Pradera, T. Calvo: *Aggregation Functions: A Guide for Practitioners*. Springer-Verlag Berlin Heidelberg, 2007.

[2] C. Butler, A. Rothstein: *On Conflict and Consensus: A Handbook on Formal Consensus Decision Making*. Takoma Park, 2006.

[3] F. Herrera, E. Herrera-Viedma, J. Verdegay: A model of consensus in group decision making under linguistic assessments. *Fuzzy Set. Syst.* 78, pp. 73–87, 1996.

[4] E. Herrera-Viedma, F. Herrera, F. Chiclana: A consensus model for multiperson decision making with different preference structures. *IEEE T. Syst. Man Cy.,A* 32, pp. 394–402, 2002.

[5] E. Herrera-Viedma, S. Alonso, F. Chiclana, F. Herrera: A consensus model for group decision making with incomplete fuzzy preference relations. *IEEE T. Fuzzy Syst.* 15, pp. 863–867, 2007.

[6] J. Kacprzyk: Group decision making with a fuzzy linguistic majority. *Fuzzy Set. Syst.* 18, pp. 105–118, 1986.

[7] X. Liu, S. Han: Orness and parameterized RIM quantifier aggregation with OWA operators: A summary. *Int. J. Approx. Reason.* 48, pp. 77–97, 2008.

[8] L. Martínez, J. Montero: Challenges for improving consensus reaching process in collective decisions, *New Maths. Nat. Comp.* 3, pp. 203–217, 2007.

[9] F. Mata, L. Martínez, E. Herrera-Viedma: An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context. *IEEE T. Fuzzy Syst.* 17, pp. 279–290, 2009.

[10] I. Palomares, P.J. Sánchez, F.J. Quesada, F. Mata, L. Martínez: COMAS: A multi-agent system for performing consensus processes, *Advances in Intelligent and Soft Computing*, 91, pp. 125-132, Springer, 2011.

[11] M. Roubens: Fuzzy Sets and Decision Analysis. *Fuzzy Set. Syst.* 90, pp. 199–206, 1997.

[12] S. Saint, J.R. Lawson: *Rules for Reaching Consensus. A Modern Approach to Decision Making*. Jossey-Bass, 1994.

[13] R.R. Yager: On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE T. Syst. Man Cy.,B* 18, pp. 183–190, 1988.

[14] R.R. Yager: Quantifier guided aggregation using OWA operators. *Int. J. Int. Syst.* 11, pp. 49–73, 1996.

[15] L.A. Zadeh: A computational approach to fuzzy quantifiers in natural languages. *Comput. Maths. Appl.* 9, pp. 149–184, 1983.

MODELO LINGÜÍSTICO DE TOMA DE DECISIONES MULTICRITERIO CON EXPRESIONES LINGÜÍSTICAS COMPARATIVAS

Rosa M. Rodríguez[1], **Luis Martínez**[1], **Francisco Herrera**[2]

[1]*Dpto. de Informática, Universidad de Jaén*, {rmrodrig,martin}@ujaen.es
[2]*Dpto. de Ciencias de la Computación y I.A, Universidad de Granada*, herrera@decsai.ugr.es

Resumen

El interés de esta contribución se centra en problemas de toma de decisiones definidos bajo incertidumbre donde la información es vaga e imprecisa. Nuestro objetivo es facilitar a los expertos la expresión de sus valoraciones cuando dudan entre varios valores lingüísticos. Para ello introducimos el concepto de Conjunto de Términos Lingüísticos Difusos Dudosos que nos permite modelar expresiones lingüísticas comparativas generadas mediante una gramática libre de contexto. Posteriormente, presentamos un modelo lingüístico de toma de decisiones multicriterio que utiliza dichas expresiones lingüísticas comparativas.

Palabras Clave: gramática libre de contexto, expresiones lingüísticas comparativas, conjunto de términos lingüísticos difusos dudosos, toma de decisiones multicriterio.

1 INTRODUCCIÓN

La toma de decisión es un proceso habitual para los seres humanos en muchas actividades del mundo real como la ingeniería, organización, finanzas, medicina, etc. Los problemas de decisión normalmente se definen en contextos donde la información es vaga e imprecisa. Existen diferentes enfoques que manejan este tipo de información, tales como, conjuntos difusos [14], conjuntos difusos tipo 2 [3], teoría de la evidencia [13], etc. En este trabajo nos hemos centrado en el enfoque lingüístico difuso [15] que ha proporcionado buenos resultados en problemas de toma de decisiones definidos bajo incertidumbre [1, 12].

Hasta ahora estos enfoques utilizan un único término lingüístico para valorar alternativas o criterios en problemas de toma de decisiones. Sin embargo, en algunas situaciones es más adecuado utilizar más de una etiqueta lingüística. En la literatura se han presentado algunos enfoques que intentan evitar esta limitación y permiten utilizar más de un término lingüístico para valorar alternativas o criterios [5, 8]. En el contexto cuantitativo cuando suceden estas situaciones donde los expertos dudan entre varios valores se ha propuesto el uso de conjuntos difusos dudosos [10].

En esta contribución introducimos el concepto de Conjunto de Términos Lingüísticos Difusos Dudosos (CTLDD) [7], que nos permitirá modelar múltiples valores lingüísticos cuando los expertos duden entre varios de ellos. Mediante una gramática libre de contexto definiremos expresiones lingüísticas comparativas cercanas al modelo cognitivo de los seres humanos. Estas expresiones serán representadas mediante CTLDD y los expertos las podrán utilizar en problemas de decisión. Una vez introducidas estas herramientas, presentaremos un modelo lingüístico de toma de decisiones multicriterio, el cual se divide en tres fases: (i) Transformación, las expresiones lingüísticas comparativas y términos lingüísticos simples son transformados en CTLDD, (ii) Agregación, se calculan las opiniones colectivas para cada alternativa utilizando dos operadores de agregación simbólicos, y (iii) Explotación, a partir de una relación de preferencia construida mediante la comparación de intervalos lingüísticos obtenidos de las opiniones colectivas, se utiliza un grado de selección de no dominancia para obtener el conjunto solución de alternativas del problema de toma de decisión.

Este trabajo se estructura como sigue: en la Sección 2 se introducen brevemente algunos preliminares necesarios. En la Sección 3 se presenta el concepto de CTLDD y se define la gramática libre de contexto que genera las expresiones lingüísticas comparativas. En la Sección 4 se propone un modelo lingüístico de toma de decisiones multicriterio y se definen dos operadores de agregación simbólicos. Además se introduce un ejemplo ilustrativo de dicho modelo, y finalmente se concluye en la Sección 5.

2 PRELIMINARES

El modelo de toma de decisiones propuesto en esta contribución representa las expresiones lingüísticas mediante CTLDD. Estos conjuntos están basados en el enfoque lingüístico difuso [15] y conjuntos difusos dudosos [10], en

esta sección revisamos brevemente ambos enfoques.

2.1 ENFOQUE LINGÜÍSTICO DIFUSO

En muchos problemas de toma de decisiones donde la información es vaga e imprecisa es más adecuado el uso de información cualitativa que cuantitativa. En estas situaciones el uso del enfoque lingüístico difuso [15] ha proporcionado buenos resultados [1, 12], utilizando para ello variables lingüísticas [15]. Una variable lingüística es una variable cuyos valores no son números sino palabras o expresiones en lenguaje natural o artificial [15]. Un valor lingüístico es menos preciso que un número, pero es más cercano al lenguaje natural utilizado por los seres humanos.

Una variable lingüística se define según Zadeh mediante una sintaxis y una semántica. Existen diferentes posibilidades para elegir la sintaxis:

- Proporcionar un conjunto de términos distribuidos en una escala con un orden definido [12]. Por ejemplo, un conjunto de 7 términos lingüísticos podría ser: $S = \{s_0 : nada, s_1 : muy\ bajo, s_2 : bajo, s_3 : medio, s_4 : alto, s_5 : muy\ alto, s_6 : perfecto\}$
- Definir el conjunto de términos lingüísticos mediante una gramática libre de contexto [2, 15].

Para definir la semántica también existen diferentes formas:

- Una semántica basada en funciones de pertenencia y una regla semántica [2].
- Una semántica basada en una estructura ordenada del conjunto de términos lingüísticos [9, 12].
- Una semántica que utiliza elementos de las definiciones anteriores [4].

En esta contribución trabajaremos con un conjunto ordenado de términos lingüísticos que requiere que existan los operadores de Negación, Máximo y Mínimo.

2.2 CONJUNTOS DIFUSOS DUDOSOS

La definición de conjunto difuso dudoso fue introducida por Torra [10] para manejar situaciones de decisión en contextos cuantitativos donde los expertos dudan entre diferentes valores para valorar alternativas o criterios. Un conjunto difuso dudoso se define como

Definición 1 *[10] Sea X un conjunto de referencia, un conjunto difuso dudoso en X es una función h que devuelve un subconjunto de valores en el intervalo [0,1]:*

$$h : X \to P([0,1])$$

Un conjunto difuso dudoso también puede ser definido en términos de la unión de sus grados de pertenencia.

Definición 2 *Sea $M = \{\mu_1, \mu_2, ..., \mu_n\}$ un conjunto de n funciones de pertenencia, el conjunto difuso dudoso asociado a M, h_M, es definido como:*

$$h_M : M \to P([0,1])$$

$$h_M(x) = \bigcup_{\mu \in M} \{\mu(x)\}$$

Diferentes operaciones de conjuntos difusos dudosos tales como la unión, intersección, complemento, etc, pueden ser encontrados en [10].

3 EXPRESIONES LINGÜÍSTICAS COMPARATIVAS

Dado que el objetivo que perseguimos en esta contribución es la de facilitar la expresión y modelado de preferencias en problemas de decisión bajo incertidumbre, en esta sección presentamos la definición de CTLDD. Posteriormente, mostraremos las expresiones lingüísticas comparativas que podrán ser utilizadas para modelar preferencias, definiremos la gramática libre de contexto que las genera, e introduciremos una función para transformar dichas expresiones en CTLDD.

3.1 CONJUNTO DE TÉRMINOS LINGÜÍSTICOS DIFUSOS DUDOSOS

De forma similar a los conjuntos difusos dudosos [10] donde los expertos pueden dudar entre varios valores para valorar una alternativa o criterio, en contextos cualitativos puede suceder que los expertos duden entre varios valores lingüísticos para expresar sus valoraciones. Para manejar estas situaciones, introducimos el concepto de CTLDD.

Definición 3 *[7] Sea $S = \{s_0, \ldots, s_g\}$ un conjunto de términos lingüísticos, un CTLDD, H_S, se define como un subconjunto finito ordenado de términos lingüísticos consecutivos de S.*

Definición 4 *[7] Para obtener el límite superior e inferior de un CTLDD, H_S, se definen los operadores, upper bound, H_{S^+}, y lower bound, H_{S^-}:*

- $H_{S^+} = max(s_i, s_j) = s_i$, *si* $s_i \geq s_j$; $s_i, s_j \in H_S$
- $H_{S^-} = min(s_i, s_j) = s_i$, *si* $s_i \leq s_j$; $s_i, s_j \in H_S$

En problemas de toma de decisiones es necesario realizar comparaciones entre términos lingüísticos. Por esta razón, definimos el concepto de *envoltura* de un CTLDD que nos permitirá comparar dos CTLDD.

Definición 5 *[7] La envoltura de un CTLDD, $env(H_S)$, es un intervalo lingüístico cuyos límites son obtenidos mediante el límite superior e inferior:*

$$env(H_S) = [H_{S^-}, H_{S^+}], \;\; H_{S^-} <= H_{S^+}$$

Otras operaciones con CTLDD fueron definidas en [7].

3.2 EXPRESIONES LINGÜÍSTICAS COMPARATIVAS. GRAMÁTICA LIBRE DE CONTEXTO

El objetivo de la definición del concepto de CTLDD es facilitar a los expertos la expresión de sus valoraciones con expresiones lingüísticas cuando duden entre varios términos lingüísticos. Estas expresiones lingüísticas, ll, serán generadas mediante una gramática libre de contexto, G_H, y representadas mediante CTLDD. Dicha gramática, G_H, se define como

Definición 6 *[7] Sea G_H una gramática libre de contexto y $S = \{s_0, \ldots, s_g\}$ un conjunto de términos lingüísticos. Los elementos de $G_H = (V_N, V_T, I, P)$ son definidos como:*

$V_N = \{\langle término\ primario\rangle, \langle término\ compuesto\rangle, \langle relación\ unaria\rangle, \langle relación\ binaria\rangle, \langle conjunción\rangle\}$

$V_T = \{menor\ que, mayor\ que, entre, y, s_0, s_1, \ldots, s_g\}$

$I \in V_N$

Las reglas de producción son definidas mediante la forma extendida de Backus Naur, en la que la utilización de paréntesis indica elementos opcionales y el símbolo | indica elementos alternativos [2]. Las reglas de producción para la gramática libre de contexto, G_H, son las siguientes:

$P = \{I ::= \langle término\ primario\rangle | \langle término\ compuesto\rangle$

$\langle término\ compuesto\rangle ::= \langle relación\ unaria\rangle \langle término\ primario\rangle | \langle relación\ binaria\rangle \langle término\ primario\rangle \langle conjunción\rangle \langle término\ primario\rangle$

$\langle término\ primario\rangle ::= s_0 | s_1 | \ldots | s_g$

$\langle relación\ unaria\rangle ::= menor\ que | mayor\ que$

$\langle relación\ binaria\rangle ::= entre$

$\langle conjunción\rangle ::= y\}$

Dichas expresiones lingüísticas comparativas, ll, son representadas mediante CTLDD utilizando la función, E_{G_H}.

Definición 7 *[7] Sea E_{G_H} una función de transformación entre expresiones lingüísticas, $ll \in G_H$, y CTLDD, H_S, siendo S el conjunto de términos lingüísticos utilizados por G_H:*

$$E_{G_H} : ll \longrightarrow H_S$$

Las expresiones lingüísticas generadas mediante las reglas de producción de la gramática libre de contexto, G_H, son transformadas en CTLDD de diferentes formas:

- $E_{G_H}(s_i) = \{s_i / s_i \in S\}$
- $E_{G_H}(\text{menor que } s_i) = \{s_j / s_j \in S \text{ y } s_j \leq s_i\}$
- $E_{G_H}(\text{mayor que } s_i) = \{s_j / s_j \in S \text{ y } s_j \geq s_i\}$
- $E_{G_H}(\text{entre } s_i \text{ y } s_j) = \{s_k / s_k \in S \text{ y } s_i \leq s_k \leq s_j\}$

4 MODELO LINGÜÍSTICO DE TOMA DE DECISIONES MULTICRITERIO CON EXPRESIONES LINGÜÍSTICAS COMPARATIVAS

Esta sección presenta un modelo lingüístico de toma de decisiones multicriterio basado en procesos computacionales simbólicos, que permite a los expertos expresar sus valoraciones mediante expresiones lingüísticas comparativas o términos lingüísticos simples.

4.1 PROBLEMA DE TOMA DE DECISIÓN LINGÜÍSTICO MULTICRITERIO

Un problema de toma de decisión lingüístico multicriterio se define como un conjunto de alternativas $X = \{x_1, \ldots, x_n\}$, donde cada alternativa se valora mediante un conjunto de criterios $C = \{c_1, \ldots, c_m\}$, que pueden ser valorados mediante expresiones lingüísticas comparativas o términos lingüísticos simples.

Supongamos que tenemos un conjunto de términos lingüísticos $S = \{s_0, \ldots, s_g\}$, y una gramática libre de contexto G_H, que genera las expresiones lingüísticas comparativas, $ll_{i,j}$, para valorar los criterios, $c_i \in C = \{c_1, \ldots, c_m\}$, de cada alternativa, $x_j \in X = \{x_1, \ldots, x_n\}$.

4.2 MODELO LINGÜÍSTICO DE TOMA DE DECISIONES MULTICRITERIO

El modelo propuesto en esta contribución realiza los procesos de computación con palabras mediante intervalos lingüísticos. Dicho modelo se divide en 3 fases (ver Figura 1):

1. *Transformación de las expresiones lingüísticas comparativas en CTLDD.*

   Los expertos pueden expresar sus preferencias mediante expresiones lingüísticas comparativas, $ll \in G_H$. Dichas expresiones son transformadas en CTLDD utilizando la función de transformación E_{G_H}, (Definición 7).

   Ejemplo: Supongamos un conjunto de alternativas $X = \{x_1, x_2, x_3\}$ y un conjunto de criterios $C = \{c_1, c_2, c_3\}$ para cada alternativa. Sea $S = \{s_0 : nada\ (n), s_1 : muy\ bajo\ (mb), s_2 : bajo\ (b), s_3 : medio\ (m), s_4 : alto\ (a), s_5 : muy\ alto\ (ma), s_6 : perfecto\ (p)\}$ el conjunto de términos lingüísticos y

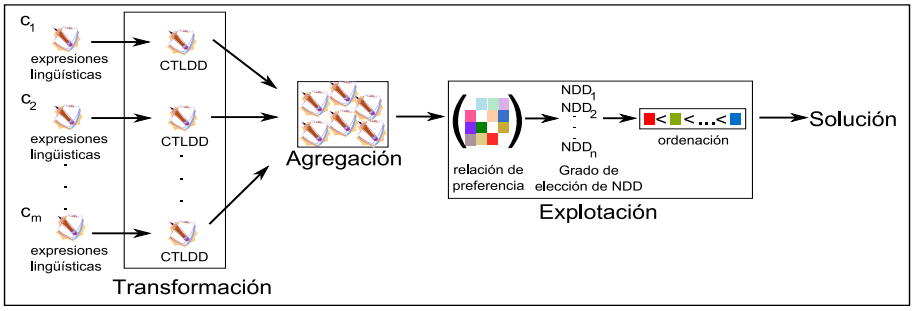


Figura 1: Esquema del modelo lingüístico de toma de decisiones multicriterio

G_H la gramática libre de contexto que genera las expresiones lingüísticas comparativas. En la Tabla 1 se visualizan las valoraciones proporcionadas para este problema.

Tabla 1: Valoraciones lingüísticas del problema de decisión

		criterios		
	ll	c_1	c_2	c_3
alt.	x_1	*entre mb y m*	*entre a y ma*	*a*
	x_2	*entre b y m*	*m*	*menor que b*
	x_3	*mayor que a*	*entre mb y b*	*mayor que a*

En la Tabla 2 se muestra la transformación de las expresiones lingüísticas comparativas en CTLDD.

Tabla 2: Valoraciones transformadas en CTLDD

		criterios		
	$H_S^j(x_i)$	c_1	c_2	c_3
alt.	x_1	$\{mb,b,m\}$	$\{a,ma\}$	$\{a\}$
	x_2	$\{b,m\}$	$\{m\}$	$\{n,mb,b\}$
	x_3	$\{a,ma,p\}$	$\{mb,b\}$	$\{a,ma,p\}$

2. *Agregación de las valoraciones representadas mediante CTLDD.*

Una vez que las valoraciones están representadas en CTLDD, se realiza el proceso de agregación utilizando para ello dos operadores de agregación simbólicos. El resultado de la agregación es un intervalo lingüístico para cada alternativa que representa la cobertura de los CTLDD agregados. Los operadores son:

i) Operador min_upper

Este operador de agregación simbólico para CTLDD obtiene como resultado el mínimo de los términos lingüísticos superiores.

Definición 8 *Sean* $\{H_S^j(x_i)/i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}\}$ *un conjunto de CTLDD y* $S=\{s_0,\ldots,s_g\}$ *un conjunto de términos lingüísticos, el operador min_upper,* $H_{S^+_{min}}$*, se define como:*

$$H_{S^+_{min}}(x_i) = min\{H^j_{S^+}(x_i) \ / \ j \in \{1,\ldots,m\}\}, \ i \in \{1,\ldots,n\}$$

Ejemplo: El resultado de utilizar este operador de agregación sobre el ejemplo presentado anteriormente puede verse en la Tabla 3.

Tabla 3: Valoraciones agregadas con, $H_{S^+_{min}}$,*min_upper*

$alternativas/H_{S^+_{min}}(x_i)$		
x_1	x_2	x_3
$\{m\}$	$\{b\}$	$\{b\}$

ii) Operador max_lower

Este operador devuelve el máximo de los términos lingüísticos inferiores de los CTLDD.

Definición 9 *Sean* $\{H_S^j(x_i)/i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}\}$ *un conjunto de CTLDD y* $S=\{s_0,\ldots,s_g\}$ *un conjunto de términos lingüísticos, el operador max_lower,* $H_{S^-_{max}}$*, se define como:*

$$H_{S^-_{max}}(x_i) = max\{H^j_{S^-}(x_i) \ / \ j \in \{1,\ldots,m\}\}, \ i \in \{1,\ldots,n\}$$

Ejemplo: Si utilizamos este operador de agregación sobre el ejemplo ilustrativo, los resultados son los siguientes (ver Tabla 4):

Tabla 4: Valoraciones agregadas con, $H_{S^-_{max}}$, *max_lower*

$alternativas/H_{S^-_{max}}(x_i)$		
x_1	x_2	x_3
$\{a\}$	$\{m\}$	$\{a\}$

iii) A continuación se construye un intervalo lingüístico para cada alternativa, $H'(x_i)$, utilizando para ello los términos lingüísticos obtenidos de los anteriores operadores de agregación:

$$H'_{max}(x_i) = max\{H_{S^+_{min}}(x_i), H_{S^-_{max}}(x_i)\}$$

$$p_{ij} = P(x_i > x_j) = \frac{max(0, Ind(H'_{S^+}(x_i)) - Ind(H'_{S^-}(x_j))) - max(0, Ind(H'_{S^-}(x_i)) - Ind(H'_{S^+}(x_j)))}{(Ind(H'_{S^+}(x_i)) - Ind(H'_{S^-}(x_i))) - (Ind(H'_{S^+}(x_j)) - Ind(H'_{S^-}(x_j)))} \quad (1)$$

$$H'_{min}(x_i) = min\{H_{S^+_{min}}(x_i), H_{S^-_{max}}(x_i)\}$$

$$H'(x_i) = [H'_{min}(x_i), H'_{max}(x_i)]$$

Ejemplo: Los intervalos lingüísticos de cada alternativa son mostrados en la Tabla 5:

Tabla 5: Intervalos lingüísticos para las alternativas

$alternativas/H'(x_i)$		
x_1	x_2	x_3
$[m,a]$	$[b,m]$	$[b,a]$

3. *Fase de explotación*

Una vez que la información lingüística ha sido agregada, se realiza la fase de explotación donde el conjunto de alternativas es ordenado a partir de los intervalos lingüísticos obtenidos, para elegir la mejor alternativa. Para ello seguimos los siguientes pasos:

i) *Construcción de una relación de preferencia*. Utilizando los intervalos lingüísticos asociados a cada alternativa construimos una relación de preferencia entre alternativas [6]. Para construir esta relación de preferencia adaptamos el método propuesto por Wang et al. en [11].

Definición 10 *[11] Sean $A = [a_1, a_2]$ y $B = [b_1, b_2]$ dos intervalos lingüísticos, el grado de preferencia de A sobre B (o $A > B$) se define como:*

$$P(A > B) = \frac{max(0, a_2 - b_1) - max(0, a_1 - b_2)}{(a_2 - a_1) + (b_2 - b_1)}$$

y el grado de preferencia de B sobre A como:

$$P(B > A) = \frac{max(0, b_2 - a_1) - max(0, b_1 - a_2)}{(a_2 - a_1) + (b_2 - b_1)}$$

Debemos tener en cuenta que $P(A < B) + P(B > A) = 1$ y $P(A > B) = P(B > A) = 0.5$ cuando $A = B$, es decir, $a_1 = b_1$ y $a_2 = b_2$.

Nota: Si los valores de los intervalos son iguales $(a_1 = a_2, b_1 = b_2)$ se compararán como números enteros. Por tanto, si $a_1 > b_1$ entonces $P(A > B) = 1$, $P(A < B) = 0$ y si $a_1 = b_1$ entonces $P(A > B) = P(A < B) = 0.5$.

Para poder utilizar esta función con intervalos lingüísticos debemos de adaptarla de la siguiente forma:

$Ind(s_i) = i, s_i \in S = \{s_0, \ldots, s_g\}$

$A = [Ind(s_i), Ind(s_j)], s_i \leq s_j$

$B = [Ind(s_k), Ind(s_l)], s_k \leq s_l$

$P(A > B) = \frac{max(0, Ind(s_j) - Ind(s_k)) - max(0, Ind(s_i - Ind(s_l))}{(Ind(s_j) - Ind(s_i)) + (Ind(s_l) - Ind(s_k))}$

Por tanto, la relación de preferencia obtenida es:

$$P_D = \begin{pmatrix} - & p_{12} & \cdots & p_{1n} \\ p_{21} & - & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & - \end{pmatrix}$$

donde p_{ij} es el grado de preferencia de la alternativa x_i sobre x_j y se calcula como indica la Ecuación (1); $i, j \in \{1, \ldots, n\}$; $i \neq j$; $Ind(s_l) = l$.

ii) *Aplicación de un grado de selección*

Existen diferentes funciones de selección para ordenar alternativas [6]. En este modelo de toma de decisiones proponemos el uso del grado de selección de no dominancia, *NDD*. Esta función indica el grado que una alternativa, x_i, no es dominada por las restantes alternativas. Se define como sigue

Definición 11 *[6] Sea $P = [p_{ij}]$ una relación de preferencia definida sobre un conjunto de alternativas X. El grado de no dominancia de la alternativa x_i, NDD_i, se obtiene aplicando la siguiente función:*

$$NDD_i = min\{1 - p^S_{ji}, j \neq i\}$$

donde $p^S_{ji} = max\{p_{ji} - p_{ij}, 0\}$ representa el grado que la alternativa x_i es estrictamente dominada por x_j.

iii) Finalmente, obtenemos el conjunto de alternativas no dominadas aplicando la siguiente función:

$$X^{ND} = \{x_i / x_i \in X, NDD_i = max_{x_j \in X}\{NDD_j\}\}$$

Ejemplo: Para aplicar la fase de explotación al ejemplo que estamos desarrollando seguimos los siguientes pasos:

i) Obtenemos los grados de preferencia utilizando la definición propuesta por Wang et al. [11].

$$P_D = \begin{bmatrix} - & 1 & 0.667 \\ 0 & - & 0.333 \\ 0.333 & 0.667 & - \end{bmatrix}$$

ii) Calculamos la relación de preferencia estricta, P^S_D, y aplicamos el grado de no dominancia para cada alternativa:

$$P_D^S = \begin{bmatrix} - & 1 & 0.334 \\ 0 & - & 0 \\ 0 & 0.334 & - \end{bmatrix}$$

$NDD_1 = min\{(1-0), 1-0)\} = 1$

$NDD_2 = min\{(1-1), (1-0.334)\} = 0$

$NDD_3 = min\{1-0.334), (1-0)\} = 0.664$

iii) Finalmente obtenemos el conjunto solución de alternativas:

$$X^{ND} = \{x_1\}$$

5 Conclusiones

El objetivo de esta contribución es facilitar la expresividad y modelado de preferencias en problemas de decisión bajo incertidumbre. Para ello ha sido introducido el concepto de CTLDD que permite modelar expresiones lingüísticas comparativas similares a las expresiones utilizadas por los seres humanos mediante una gramática libre de contexto. Finalmente, un modelo lingüístico de toma de decisiones multicriterio que utiliza dichas expresiones lingüísticas, y realiza los procesos de computación mediante intervalos lingüísticos ha sido presentado.

Agradecimientos

Este trabajo está parcialmente financiado por los proyectos de investigación TIN-2009-08286, P08-TIC-3548 y fondos FEDER.

Referencias

[1] B. Arfi. Fuzzy decision making in politics: A linguistic fuzzy-set approach (LFSA). *Political Analysis*, 13(1):23–56, 2005.

[2] G. Bordogna and G. Pasi. A fuzzy linguistic approach generalizing boolean information retrieval: A model and its evaluation. *Journal of the American Society for Information Science*, 44:70–82, 1993.

[3] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications*. Kluwer Academic, New York, 1980.

[4] F. Herrera, E. Herrera-Viedma, and L. Martínez. A fusion approach for managing multi-granularity linguistic term sets in decision making. *Fuzzy Sets and Systems*, 114(1):43–58, 2000.

[5] J. Ma, D. Ruan, Y. Xu, and G. Zhang. A fuzzy-set approach to treat determinacy and consistency of linguistic terms in multi-criteria decision making. *International Journal of Approximate Reasoning*, 44(2):165–181, 2007.

[6] S.A. Orlovski. Decision-making with fuzzy preference relations. *Fuzzy Sets and Sys.*, 1:155–167, 1978.

[7] R.M Rodríguez, L. Martínez, and F. Herrera. Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 2011, (Doi:10.1109/TFUZZ.2011.2170076).

[8] Y. Tang and J. Zheng. Linguistic modelling based on semantic similarity relation among linguistic labels. *Fuzzy Sets and Systems*, 157(12):1662–1673, 2006.

[9] V. Torra. Negation functions based semantics for ordered linguistic labels. *International Journal of Intelligent Systems*, 11:975–988, 1996.

[10] V. Torra. Hesitant fuzzy sets. *International Journal of Intelligent Systems*, 25(6):529–539, 2010.

[11] Y.M. Wang, J.B. Yang, and D.L. Xu. A preference aggregation method through the estimation of utility intervals. *Computers and Operations Research*, 32:2027–2049, 2005.

[12] R.R. Yager. An approach to ordinal decision making. *International Joumal of Approximate Reasoning*, 12(3-4):237–261, 1995.

[13] J.B. Yang, J. Liu, J. Wang, H-S. Sii, and H-W. Wang. Belief rule-base inference methodology using the evidential reasoning approach - rimer. *IEEE Transactions on Systems, Man, and Cybernetics-Part A*, 36(2):266–285, 2006.

[14] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[15] L. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences, Part I, II, III*, (8,9):199–249,301–357,43–80, 1975.

APPLICATION OF THE AHP IN CORPORATE REPUTATION MANAGEMENT

Ana María Casado Molina [1], José Ignacio Peláez Sánchez [2], Jesús Doña Fernández [3]

[1] *Department of Audiovisual Communication and Advertising University of Malaga, Spain, anacasado@lcc.uma.es*

2 Department of Languages and Computer Sciences University of Málaga,, Spain, jipelaez@uma.es

3 Department of Technologies and Information Systems, Health Management Area North of Málaga, Spain jesusm.dona.sspa@juntadeandalucia.es

Abstract

In recent years, companies and organizations make significant efforts in order to develop policies regarding Corporate Reputation. The loss of reputation directly impacts the performance of companies. Corporate Reputation is an intangible asset of great value for the enterprise, its identification and its management is essential for corporate sustainability along the time. The Analytical Hierarchy Process (AHP) is proposed as a potential decision making systems in Corporate Reputation Management and a hierarchical structure is constructed. By applying the AHP, the corporate reputation criteria can be prioritized and a descending-order list of their most important stakeholders and their issues (criteria) can be made in order to select the best corporate reputation strategy to perform the economic and social sustainability of the company along the time.

Keywords: Analytical Hierarchy Process, AHP, Corporate Reputation, Decision Making System, Sustainability, Corporate Management.

1 INTRODUCTION

The Corporate Reputation (CR) can be defined as an intangible asset that adds value to businesses, through the perceptions from the different stakeholders of the company's behavior (activities and communication). There are a lot of definitions about CR. Aguado [1], for example, defines it as established and substantiated prestige that an entity achieves in front of all its stakeholders; Villafañe [20], said that CR is when the stakeholders recognize the good company's behavior along the time through to carry out the commitments with their customers, employees, shareholders, community, etc; Fombrun [5, 6, 7], describes it as the perception that different stakeholders have of that institution along the time. All the definitions about Corporate Reputation include the following items [3]: (1) The companies establish a value-added and committed relationship with their stakeholders; (2) It is a perceptions or judgments of the stakeholders; (3) these perceptions are the result of how the companies meets the expectations of their stakeholders and their environment; and (4) it is built on long-term; the past, present and future of the company can influence.

The Analytical Hierarchy Process is a multi-criteria decision making method developed by Thomas L. Saaty in 1980 [4, 15, 16, 17]. This method is successfully applied to many cases that required complex decision analyses. The strength of this approach is that it organizes tangible and intangible factors in a systematic way, and provides a structured yet relatively simple solution to the decision-making problems [19].

The objective of this paper is to introduce the application of the AHP in corporate reputation management. The paper will briefly review the concepts and applications of the multiple criteria decision analysis, the AHP's implementation step, the information of corporate reputation and present AHP application on the Corporate Reputation management.

2 BACKGROUND

2.1. MULTIPLE CRITERIA DECISION ANALYSIS

The Chief Reputation Officers (CRO) or CEOs are faced with decision environments and problems in corporate reputation that are complex. The elements of the problems are numerous, and the interrelationships among the elements are extremely complicated. Relationships between elements of a problem may be highly nonlinear; changes in the elements may human value and judgment systems are integral element of project problems [11]. Therefore, the ability to make sound decisions is very important to the success of the corporate reputation along the time.

Multiple criteria decision-making (MCDM) approaches are major parts of decision theory and analysis. They seek

to take explicit account of more than one criterion in supporting the decision process [11]. The aim of MCDM methods is to help decision-makers learn about the problems they face and through exploring the context of the problem and the different alternatives to guide them in identifying a preferred course of action [2, 8, 14, 21, 22, 23].

2.2. THE ANALYTICAL HIERARCHY PROCESS

Saaty developed the following steps for applying the AHP:

1. Define the problem and determine its goal.
2. Structure the hierarchy from the top (the objectives from a decision-maker's viewpoint) through the intermediate levels (criteria on which subsequent levels depend) to the lowest level which usually contains the list of alternatives.
3. Construct a set of pair-wise comparison matrices (size *n* x *n*) for each of the lower levels immediately above by using the relative scale measurement (table 1) [4, 8, 20, 13]. The pair-wise comparisons are done in terms of which element dominates the other.

Table 1: Pair-wise comparison scale for AHP preferences

Numeral rating	Verbal judgements of preferences
1	Equally preferred
3	Moderately preferred
5	Strongly proferred
7	Very strongly preferred
9	Extremely preferred
2,4,6,8	Intermedia Values

4. There are *n (n-1)* / 2 judgments required to develop the set of matrices in step 3. Reciprocals are automatically assigned in each pair-wise comparison.
5. Hierarchical synthesis is now used to weight the eigenvectors by the weights of the criteria and the sum is taken over all weighted eigenvector entries corresponding to those in the next lower level of the hierarchy.
6. Having made all the pair-wise comparisons, the consistency is determined by using the eigenvalue, $\lambda_{max,}$ to calculate the consistency index *CI*, as follows, $CI=(\lambda max-n)/(n-1)$, where *n* is the matrix size.

   Judgment consistency can be checked by taking the consistency ratio $CR^*=CI/RI$, where *RI* is the random index. The judgments are consistent, if CR^* does not exceed 0.10 [15].
7. Steps 3-6 are performed for all levels in the hierarchy.

2.3. CORPORATE REPUTATION

Corporate Reputation is determinate by a number of inherent characteristics [12]. These are:

1. Its multidimensional nature. Diversity and heterogeneity of participating groups who form their perceptions and their different way of evolution depending on emotional appeals and expectations according to the group that evaluates (customers, employees, suppliers, etc.).

2. The assessments of the stakeholders about corporate issues are different and they cannot be treated as a whole. Each stakeholder groups have different priorities, expectations about corporate issues and, therefore, perceptions may defer in their weight. To manage corporate reputation is important to know which stakeholders contribute more in it, which ones are more priority for business success over time and their preferences.

3. Its complexity. CR valuation of the different variables of the company that affect these stakeholders cannot be treated as a whole. It is essential to take into account stakeholders' value by department or units, and furthermore consider the stakeholders perceptions value in each of the areas.

4. Its transversal nature, it is difficult to handle and slow accumulation. The Corporate Reputation is based on the stakeholder perceptions about their relationships with every department or business units of the companies. A company's reputation is the result of years of work, good behavior, good communication and good sustainable relations with their stakeholders over time. It is not easy to handle and needs to work with many information gathering systems.

5. Empirically verifiable. It is difficult to measure and will have greater value in those companies that are able to measure their return on investment. It becomes necessary to connect historical impact of reputational criteria and their stakeholders, with the invested resources by variable and stakeholders.

The effective management of Corporate Reputation is key to business survival and the sustainability of the companies over time. Table 2 shows the main value contributions, quantitative and qualitative [5,7,9,20].

Table 2: Corporate Reputation (CR) Value Contributions to Companies

RC Value For Companies	
Quantitatives	**Quaalitatives**
Reduce Costs	Atracts Innovation
Price Increase	Builds barriers to Competitors
Attracts Investments	Shield against crisis / Minimizes impact
Favours Stock Exchange quotation rise of shares	Quality commitment
Multiplies Brand value	Favours Differentiation
Attracts and Retains Employees / Talent	Reliable Relationships
Attracts Quality Suppliers	Corporate Decisions Acceptance
Customer Loyalty	New Markets Accessibility

3 APPLICATION OF THE AHP IN CR MANAGEMENT

In this paper, the direct perceptions of stakeholders will be used as an example of the possibility of using AHP in Corporate Reputation management in companies from ICT sector [3]. This example has been structured following the steps to applying the AHP (see section 2.2):

Step 1. Structuring the hierarchy for evaluation

The AHP method is used to make the arrangement of the problems as a hierarchy. In the Figure 1 (annex 1) show the CR hierarchy with their stakeholders.

The first group is the Supra decision-maker. This group is formed by the Executive Board and is in charge of determining the importance which the different stakeholders, internal or external group, have in the Corporate Reputation. The Second level in the hierarchy show the internal and external category of stakeholders (presidency, employees, customers…). Each category has different groups, (1…14), with their specific cardinality or number of people. In the third level, every group (1…14) establish the items that they consider to evaluate the Corporate Reputation. These items are sometimes established by the Supra decision-maker because they are strategic for the corporate issues.

When the hierarchy is structured, the number of people of every group is established as shown in the table 3.

Table 3: List of Stakeholders

Stakeholders	Cardinality of Experts
Presidency/Board of Administration	2
Employees	20
Internal Shareholders	4
External Shareholders	10
End-Users Customers	30
Small & Medium Enterprises Customers	12
Corporations Customers	5
Suppliers of Products	10
Suppliers of Services	10
Opinion Leadership/Mass Media	15
Opinion Leadership / On-Line Communication	15
Opinion Leadership / Unions	6
NGOs	8
Op. Leadership./Financial Analysts	3
Total	150

Step 2. Constructing the pair-wise comparison matrix

After structuring a hierarchy, the pair-wise comparison matrix for each group and decision makers is constructed (Table 4). During the pair-wise comparison, a nominal scale is used for the evaluation. The scale used in AHP for preparing the pair-wise comparison matrix is a discrete scale from 1 to 9, as presented in Table 1. For example, for the first level in the hierarchy, the group of Supra decision-makers (10 members) give their opinions by pair-wise comparison matrix,

[1 1/3; 3 1]	[1 1/4; 4 1]
[1 2; 1/2 1]	[1 1/5; 5 1]
[1 2; 1/2 1]	[1 1/3; 3 1]
[1 1/6; 6 1]	[1 2; 1/2 1]
[1 1/3; 3 1]	[1 1/7; 7 1]

Table 4: Pair-wise comparison matrix for 1 Supra decision-makers

CR	Internal Stakeholders	External Stakeholders	Priority
Internal Stakeholders	1	1/3	0.25
Extenal Stakeholders	3	1	0.75

Another example, in the third level is the Group 4 which is External Shareholders. These 10 members give their opinions as follows,

[1 2 9 4; 1/2 1 2 5; 1/9 1/2 1 2; 1/4 1/5 1/2 1]
[1 2 8 4; 1/2 1 2 5; 1/8 1/2 1 2; 1/4 1/5 1/2 1]
[1 2 7 4; 1/2 1 2 5; 1/7 1/2 1 2; 1/4 1/5 1/2 1]
[1 2 5 4; 1/2 1 2 5; 1/5 1/2 1 2; 1/4 1/5 1/2 1]
[1 2 4 3; 1/2 1 2 3; 1/4 1/2 1 3; 1/3 1/3 1/3 1]
[1 2 5 4; 1/2 1 2 6; 1/5 1/2 1 2; 1/4 1/6 1/2 1]
[1 4 5 6; 1/4 1 2 3; 1/5 1/2 1 2; 1/6 1/3 1/2 1]
[1 2 3 5; 1/2 1 1 5; 1/3 1 1 2; 1/5 1/5 1/2 1]
[1 4 3 4; 1/4 1 2 5; 1/3 1/2 1 2; 1/4 1/5 1/2 1]
[1 2 1 4; 1/2 1 2 5; 1 1/2 1 2; 1/4 1/5 1/2 1]

Step 3. Testing the consistency for each level

When all the people make their perceptions by their judgments, we check the consistency of their opinions. As aforementioned in section 2.2.

Step 4. Applying Group Decision Making

For each of these groups: Supra decision maker, Group 1, Group 2, Group 3, Group 4, Group 5, Group 6, Group 7, Group 8, Group 9, Group 10, 11, 12, 13 and Group 14, the process of the decision making shown in figure 2 applies, resulting in a sub-hierarchy for each group which is shown as a geometric mean of all the decision makers.

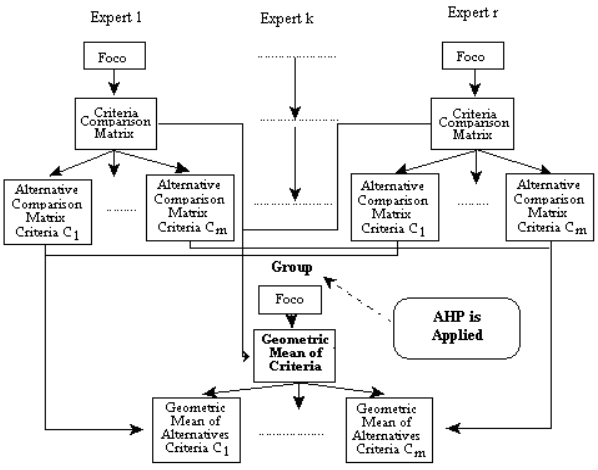


Figure 2: Decision Group Scheme with Individual Judgment Aggregation (IJA).

The final result is a hierarchy which is the geometric average of the hierarchies of the different groups.

Following the examples, all the individual matrix of the Supra decision-makers group (10 members) are collected in a aggregated matrix,

[1.0000 0.4516;
2.2144 1.0000]

Table 5 shows the aggregated matrix of the External Shareholders. The Growth Orientation (Criterion C) is strongly preferred than the Company's Investor (Criterion B). It's strongly preferred because the level of the scale is 5.359.

Table 5: Pair-wise comparison matrix for External Shareholders Group (19 members)

External Shareholders	A	B	C	D
Analyst Prescription (A)	1	2.2974	4.3208	4.1387
Company's Investor (B)	0.4353	1	1.8661	4.5975
Growth Orientation (C)	0.2314	0.5359	1	2.0828
Corporate Governance (D)	0.2416	0.2175	0.4801	1

Step 5. Calculating the priority of the Corporate Reputation Items

Finally, the priorities of the different items and stakeholders are determined by the Saaty's eigenvalue method [18]. The Table 6 shows the results after applying the different aggregation operators. In the first column show the priorities of the different items and stakeholders to manage Corporate Reputation.

Table 6. Stakeholders' Priorities for CR Management

Corporate Reputation	**PG-OWG**
Internal Stakeholders	0.3701
Presidency/Board of Administration	0.0968
Professional Reputation /Image	0.0559
Professional Leadership	0.0274
Participant Management	0.0134
Workforce (employees)	0.2733
Talent	0.1106
Corporate Culture	0.0569
Work Environment	0.0493
Internal Communication	0.0383
Quality of Work	0.0182
External Stakeholders	0.6299
Investors	0.2879
Internal Shareholders	0.1050
Prescription of Analysts	0.0480
Company's Investments	0.0309
Growth Orientation	0.0168
Good Corporate Governance	0.0093

External Shareholders	0.1829
Prescription of Analysts	0.0913
Company's Investments	0.0517
Growth Orientation	0.0251
Good Corporate Governance	0.0148
Customers	0.1939
End-Users	0.1013
Product & Services	0.0328
Price	0.0253
POS/Sales Force	0.0127
Value Added Services	0.0106
Corporate Brand	0.0088
Product Brand	0.0110
Small & Medium Enterprises	0.0536
Product & Services	0.0167
Price	0.0137
POS/Sales Force	0.0067
Value Added Services	0.0057
Corporate Brand	0.0047
Product Brand	0.0060
Corporates	0.0390
Products & Services	0.0114
Price	0.0101
POS/Sales Force	0.0052
Value Added Services	0.0042
Corporate Brand	0.0035
Product Brand	0.0046
Suppliers	0.0971
Suppliers of Products	0.0390
Corporate Image	0.0148
Transparent Recruitment	0.0132
Commitment with suppliers	0.0071
Quality of Services	0.0038
Suppliers of Services	0.0581
Corporate Image	0.0150
Transparent Recruitment	0.0239
Commitment with suppliers	0.0129
Quality of Services	0.0064
Opinion Leaderships	0.0510
Mass Media	0.0199
Influence on public opinion	0.0078
Prescription in sectoral groups	0.0121
Online Communities	0.0105
Influence on public opinion	0.0050
Precription in sectoral groups	0.0056
Unions	0.0098
Channel Interests	0.0038
Promote Labor Compliance	0.0035
CSR with employees	0.0025
NGOs (Enviroment, social,...)	0.0074
Influence on public opinion	0.0048
Prescription in sectoral groups	0.0026
Financial Analysts	0.0034
Financial Strength	7.9452e-04
Corporate Governance	6.9751e-04
Economic Results	3.3665e-04
Corporate Brand	4.0145e-04
Product Brand	3.0029e-04
Transparent Communication	3.1451e-04
CSR/Responsible Investments	3.1971e-04
Company's Strategic projection	2.3056e-04

Analyzing the Table 6, the priorities of the type of Stakeholders to manage the Corporate Reputation are External Stakeholders versus Internal Stakeholders. In the category of External Stakeholders the order of priorities are shareholders (0.2879), customers (0.1939), suppliers (0.0510) and opinion leaders (0.0510); in the Internal Stakeholder Category, the workforce (0.2733) is more important than Board of Directors (0.0968).

Applying AHP, we also obtain the priorities of each group of Stakeholders. In the example, the way to improve the Corporate Reputation with shareholders is to prepare actions which have relationship with their priorities and preferences. So the shareholders' priorities from highest to lowest for TIC companies are Analyst Prescription, company's investors, growth orientation and corporate governance. If companies define their corporate activities and communications taking into account the groups and their priorities, they manage and improve their corporate reputation.

AHP method does provide facilities for performing sensitivity analysis, where the Chief Reputation Officer can check the sensitivity of his judgments on the overall priorities of stakeholders by trying different values for his comparison judgments.

4 CONCLUSION

Corporate Reputation Management involves complex decision making situations that require discerning abilities and methods to make sound decisions. CRM needs to work with criteria and priorities that are determined by owner requirements and preferences of company's Supra decision-maker as well as the characteristics of all the individual stakeholders. This paper present the application of the method AHP as a decision-making process for the

corporate reputation problem, which allows the consideration of the multiple-criteria involved in the decision process. Also, a hierarchical structure of Stakeholders for corporate reputation evaluation is completely developed. Finally, an example of TIC company's stakeholders and their perception items is worked out to illustrate the application of AHP in corporate reputation management. AHP allows group decision-making and know the priorities of each group of stakeholder to carry out specific strategies for successful corporate reputation.

Acknowledgements

We would like to thank the Corporate Excellence - Center for Reputation Leadership, for providing the facilities and support to collect the information for the example which is published in this paper.

References

[1] Alcalá, M.A.: Reputación Corporativa. Fundamentos Conceptuales.Deloitte. http://www.mindvalue.com/reputacion.pdf

[2] Belton V.: Multiple criteria decision analysis-practically the only way to choose. In: HendryLC, Eglese RW, editors. *Operational research tutorial paper*., pp.53-102, 1990.

[3] Casado, A.M. (2011): The Chief Reputation Officer: Un Nuevo Modelo para la Reputación Corporativa. Tesis Doctoral, Universidad de Málaga. pp.156-157

[4] Expert Choice, Inc., Expert Choice software and manual. 4922 Elsworth Ave., Pittsburgh, PA 15213, USA.

[5] Fombrun, C. y Shanley, M. (1990*): "What's in a Name? Reputation Building and Corporate Strategy"*. Academy of Management Journal, vol. 33 , pp. 233-258.

[6] Fombrun, C.J. (1996): Reputation. Realizing Value from the Corporate Image. Harvard Business School Press, Boston.

[7] Fombrun, C. y Van Riel, C. (1997): "The Reputational Landscape" Corporate Reputation Review, vol. 1, nº 1. pp. 7-21.

[8] French S.: *Decision Theory: An Introduction to the mathematics of rationality*. Ellis Horwood, Chichester, 1988.

[9] Hill, K. y Knowlton, F (2002): Corporate Reputation. Wath, Oxford.

[10] Lamata M.T. Peláez J.I. 2002. A Method of Improving the Consistency of Judgements. Uncertainty, Fuzziness and Knowledge Based Systems. Vol. 10-6. Pag. 677-686.

[11] Lifson MW, Shaifer EF: Decision and risk analysis for construction management. Wiley, Nueva York, 1992.

[12] Martin de Castro, G. (2008): Reputación Empresarial y Ventaja Competitiva. Esic Editorial, Madrid.

[13] Peláez J.I., Lamata M.T. 2003. *A New Alternative Measure of Consistency for Positive Reciprocal Matrices*. Computers & mathematics with applications. Vol. Pag. 1839-1845.

[14] Russell JS. Surety bonding and owner-contractor prequalification:comparison. Journal of Professional Issues in Engineering, ASCE 1990; 116 (4), pp. 360-374.

[15] Saaty, TL: *The analytic hierarchy process*. McGraw-Hill, New York, 1980

[16] Saaty TL: Axiomatic Foundation of the Analytic Hierarchy Process. *Management Science*, 32, pp.841-855, 1986

[17] Saaty TL, Kearns KP: Analytical planning: the organization of systems. *The analytic hierarchy process series 1991*, vol 4RWS Publications Pittsburg USA.

[18] Saaty, T.L., (1994). Fundamentals of decision making and priority, Pittsburgh (PA): Rws. Publications.

[19] Skibniewsky MJ, Chao L.: Evaluation of advanced construction technology with AHP method. *Journal of Construction Engineering and Management*, 118(3), pp. 577-593, 1992.

[20] Villafañe, J (2004): La buena reputación. Claves sobre su valor intangible. Editorial Pirámide. Madrid.

[21] Von Winterfeldt D, Edwards W.: *Decision analysis and behavioral research*. Cambridge University Press, Cambridge, 1986.

[22] Watson SR, Buede DM.: *Decision synthesis: the principles and practice of decision analysis*. Cambridge University Press, Cambridge, 1987.

[23] Zeleny M.: Multiple criteria decision making. McGraw-Hill, Nueva York, 1990.

Annex 1

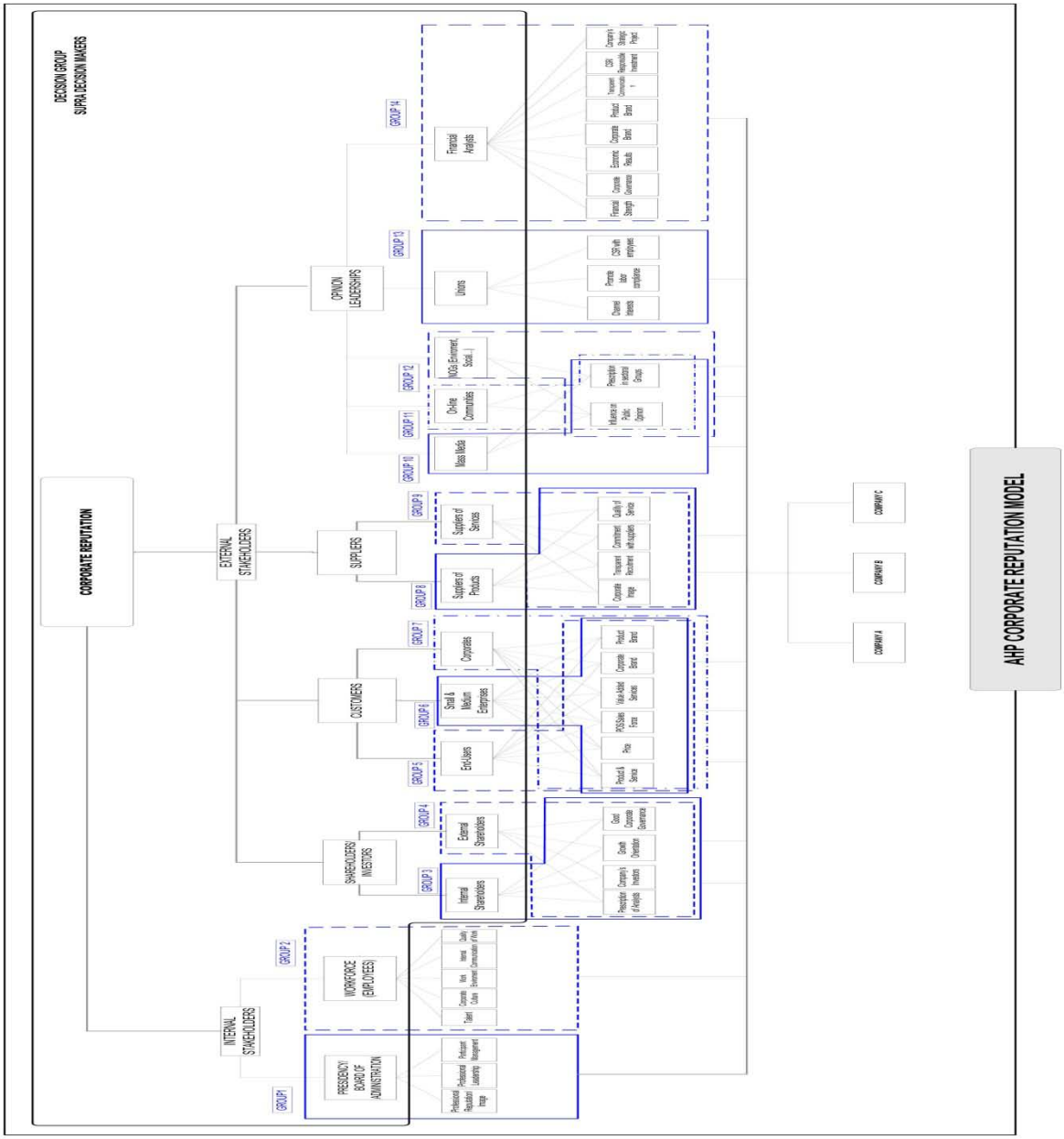


Figure 1: Hierarchy of Stakeholders for CR Evaluation (AHP Corporate Reputation Model)

TOMA DE DECISION MULTICRITERIO USANDO INTEGRALES DE CHOQUET INTERVALO-VALORADAS

Humberto Bustince[1], **Javier Fernandez**[1], **Aranzazu Jurio**[1], **Radko Mesiar**[2], **Anna Kolesárová**[3], **Benjamin Bedregal**[4]

[1]*Dept. Automatica y Computación, Universidad Pública de Navarra, Pamplona, Spain*, {bustince,fcojavier.fernandez,aranzazu.jurio}@unavarra.es

[2]*Department of Mathematics, Faculty of Civil Engineering, Slovak University of Technology, 81368 Bratislava, Slovakia and UTIA CAS Prague, Czech Republik*, mesiar@math.sk

[3]*Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, SK-812 37 Bratislava, Slovakia*, anna.kolesarova@stuba.sk

[4]*Departamento de Informatica e Matematica Aplicada Universidade Federal do Rio Grande do Norte Brazil*, bedregal@dimap.ufrn.br

Resumen

Proponemos un algoritmo de toma de decisión multicriterio para relaciones de preferencia intervalo-valoradas que se basa en el uso de una definición apropiada de integrales Choquet intervalo-valoradas. Para poder definir dichas integrales, proponemos un método de construcción de órdenes totales entre intervalos basado en el uso de funciones de agregación adecuadas.

Palabras Clave: Orden total, Agregación, Integral Choquet intervalo-valorada.

1 INTRODUCCIÓN

Los problemas de toma de decisión multicriterio (TDM) son muy habituales. Estos problemas pueden modelizarse como sigue: Dado un conjunto de alternativas $\{A_1,\ldots,A_n\}$ y un conjunto de criterios $\{x_1,\ldots,x_k\}$, escoger la mejor alternativa de acuerdo con dichos criterios. Una manera útil de tratar este tipo de situaciones es proporcionando un valor que mida la satisfacción por la alternativa A_i del criterio x_j. De esta foma, un problema de TDM puede describirse por medio de una matriz, denominada matriz de TDM, en la que el elemento en la fila i y columna j denota hasta qué punto la alternativa A_i satisface el criterio x_j. Una vez que se dispone de esta matriz, muchos algoritmos la utilizan para asignar un valor a cada una de las alternativas. Finalmente las alternativas se ordenan de acuerdo con estos valores. Cabe destacar que con esta aproximación un problema de TDM, una vez dada la matriz, es equivalente a, primero, encontrar una forma de agregar los valores en cada fila y, segundo, hallar un orden apropiado para los valores obtenidos.

La teoría difusa se ha mostrado muy útil para tratar los problemas de TDM. Recordemos que, dado un universo finito $U=\{u_1,\ldots,u_n\}$, un conjunto difuso sobre U viene dado por una función de pertenencia $f:U\to[0,1]$. Para utilizar esta teoría, es necesario normalizar los valores en la matriz de TDM al intervalo $[0,1]$.

Para agregar los valores de cada fila en la matriz de TDM, es recomendable introducir el valor esperado $E(f)$ de F. La aproximación original de Zadeh a este concepto en [11] se basaba en una medida de probabilidad P sobre U

$$P(u_i)=p_i \quad \text{y} \quad E(f)=\sum_{i=1}^{n} p_i f(u_i)\,.$$

Pero es posible considerar enfoques más generales basados en medidas difusas e integrales Choquet [3, 4].

Por otro lado, los conjuntos intervalo-valorados han sido muy estudiados en los últimos años [8, 10], ya que permiten tener en cuenta la incertidumbre a la hora de determinar la satisfacción de un criterio dado por la alternativa considerada [2]. Esta extensión hace necesario introducir el valor esperado para este tipo de objetos. Sin embargo, al contrario que en el caso real, no existe un orden total universalmente aceptado entre los intervalos en $[0,1]$.

El objetivo de este trabajo es discutir el uso de conjuntos intervalo-valorados en problemas de TDM. Para ello, introducimos el concepto de valor esperado para conjuntos intervalo-valorados basado en integrales Choquet intervalo-valoradas. [6]. Como la clasificación final de la alternativas viene dada por intervalos, también presentamos un método de construcción de órdenes totales entre intervalos por medio de funciones de agregación. Dejamos para más adelante el análisis de cuál de estos órdenes es el más adecuado para un problema de TDM dado.

Este trabajo se organiza como sigue. En la sección siguiente recordemos la aproximación estándar a la integral intervalo-valorada de Choquet a partir del concepto de integral de Aumanna [1]. En la sección 3, proponemos un método para obtener órdenes lineales sobre intervalos. En la sección 4 presentamos nuestra aproximación a la integral de Choquet intervalo-valorada. La sección 5 se dedica a nuestro algoritmo de TDM. Terminamos con las conclu-

siones y referencias.

2 INTEGRALES CHOQUET INTERVALO-VALORADAS

Un conjunto intervalo-valorado F viene dado por su función de pertenencia $F : U \to J([0,1])$, donde $J([0,1]) = \{[a,b]|, 0 \le a \le b \le 1\}$ La generalización de los números reales a los intervalos (cerrados) reales vino forzada por el desarrollo de los ordenadores y ha llevado al desarrollo de la aritmética intervalar [7]. Por ejemplo, la suma se define como $[a,b]+[c,d] = \{x+y|x \in [a,b], y \in [c,d]\} = [a+c, b+d]$.

Una idea similar condujo a Aumann [1] a su idea de integrales de funciones evaluadas sobre conjuntos. Ambos conceptos son de la misma naturaleza que el principio de extensión de Zadeh [12] y, en el marco de la integral Choquet, aparecen en diversos trabajos, como [6, 13].

Recordemos que el concepto de integral Choquet se define como sigue. Dada una medida difusa, es decir una función $m : 2^U \to [0,1]$ tal que $m(\emptyset) = 0$, $m(U) = 1$, $m(A) \le m(B)$ si $A \subseteq B \subseteq U$ la integral Choquet discreta (o esperanza) de f se define como

$$E(f) = C_m(f) = \sum_{i=1}^{n} f(u_{\sigma(i)})(m(\{u_{\sigma(i)}, \ldots, u_{\sigma(n)}\}) - \\ -m(\{u_{\sigma(i+1)}, \ldots, u_{\sigma(n)}\})),$$

donde $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ es una permutación tal que $f(u_{\sigma(1)}) \le f(u_{\sigma(2)}) \le \ldots \le f(u_{\sigma(n)})$, y $\{u_{\sigma(n+1)}, u_{\sigma(n)}\} = \emptyset$ por convenio. Esta definición puede extenderse al caso intervalo-valorado como sigue.

Definición 1 *Sea $F : U \to J([0,1])$ un conjunto intervalo-valorado difuso y $m : 2^U \to [0,1]$ une medida difusa. La experanza basada en la integral Choquet viene dada por* $\mathbf{C}_m(F)$

$$\begin{aligned} \mathbf{C}_m(F) &= \{C_m(f)|f : U \to [0,1],\ f(u_i) \in F(u_i)\} = \\ &= [C_m(f_*), C_m(f^*)], \end{aligned} \quad (1)$$

con $f_*, f^* : U \to [0,1]$ dadas por $f_*(u_i) = a_i$ y $f^*(u_i) = b_i$, con $[a_i, b_i] = F(u_i)$.

En [6, 13] se discuten algunas propiedades de esta integral. Por ejemplo, $\mathbf{C}_m(F)$ es comonótona aditiva, i.e, $\mathbf{C}_m(F+G) = \mathbf{C}_m(F) + \mathbf{C}_m(G)$ si $F, G : U \to J([0,1])$ son tales que $F(u_i) + G(u_i) \subseteq [0,1]$ para cada $u_i \in U$, y F, G son comonótonas, i.e., $(f^*(u_i) - f^*(u_j))\ (g^*(u_i) - g^*(u_j)) \ge 0$ y $(f_*(u_i) - f_*(u_j))\ (g_*(u_i) - g_*(u_j)) \ge 0$ para todo $u_i, u_j \in U$.

3 ORDENES TOTALES SOBRE INTERVALOS

La integral Choquet discreta C_m se basa en una permutación $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ con $F(u_{\sigma(1)}) \le F(u_{\sigma(2)}) \le \ldots \le F(u_{\sigma(n)})$. Para adaptar esta idea al caso intervalo-valorado necesitamos un orden total $\preceq$ sobre $J([0,1])$. En la literatura se encuentran diferentes ejemplos de órdenes totales. Por ejemplo:

(i) Orden de Yager y Xu basado en el score y el accuracy. $[a,b] \preceq M, G[c,d]$ si y solo si $a+b < c+d$ o $a+b = c+d$ y $d-c \le b-a$ ([14].

(ii) Orden lexicográfico: $[a,b] \preceq_{P1} [c,d]$ ($[a,b] \preceq_{P2} [c,d]$) si y solo si $a < c$ ($b < d$) o $a = c$ y $b \le d$ ($b = d$ y $a \le c$).

Estos dos ejemplos extienden el orden parcial usual entre intervalos $[a,b] \le [c,d]$ si y solo si $a \le c$ y $b \le d$. Introducimos ahora una clase de órdenes que también lo extienden y que incluye los ejemplos anteriores.

Lema 1 *Sean $A, B : [0,1]^2 \to [0,1]$ dos funciones de agregación [4] tales que $A(x,y) = A(u,v)$ y $B(x,y) = B(u,v)$ solo si $(x,y) = (u,v)$. Sea $\preceq_{A,B}$ en $J([0,1])$ definida por $[x,y] \preceq_{A,B} [u,v]$ si $A(x,y) < A(u,v)$ o $A(x,y) = A(u,v)$ y $B(x,y) \le B(u,v)$. Entonces $\preceq_{A,B}$ es un orden lineal en $J([0,1])$ con elemento mínimo $\{0\} = [0,0]$ y elemento máximo $\{1\} = [1,1]$.*

La prueba es trivial y la omitimos. Con esta definición, el orden de Yager y Xu se puede recuperar como el orden lineal $\preceq_{M,G}$ sobre $J([0,1])$, con M la media aritmética y G la media geométrica. El orden lexicográfico con respecto a la primera (segunda) variable se puede obtener tomando como A la proyección respecto a la primera (segunda) variable y como B la proyección respecto a la segunda (primera) variable.

Cabe indicar que los órdenes lineales definidos por pares diferentes pueden coincidir. Consideremos por ejemplo $\preceq_{Min,Max} \equiv \preceq_{P_1,P_1} \equiv \preceq_{P_1,B}$, con $P_1, P_2 : [0,1]^2 \to [0,1]$ proyecciones, $P_1(x,y) = x$, $P_2(x,y) = y$ y $B : [0,1]^2 \to [0,1]$ una función de agregación cancelativa arbitraria.

4 INTEGRALES (*A*,*B*)–CHOQUET INTERVALO-VALORADAS

Consideremos un orden total $\preceq_{A,B}$ definido como en el Lema 1. .

Definición 2 *Sea $F : U \to J([0,1])$ un conjunto intervalo-valorado difuso y $m : 2^U \to [0,1]$ una medida difusa. En las condiciones del Lema 1, la integral (A, B)–Choquet $\mathbf{C}_m^{A,B}(F)$ viene dada por*

$$\mathbf{C}_m^{A,B}(F) = \\ = \sum_{i=1}^{n} F(u_{\sigma_{A,B}(i)})(m(\left\{u_{\sigma_{A,B}(i)}, \ldots, u_{\sigma_{A,B}(n)}\right\}) - \\ -m(\left\{u_{\sigma_{A,B}(i+1)}, \ldots, u_{\sigma_{A,B}(n)}\right\})),$$

donde $\sigma_{A,B} : \{1,\ldots,n\} \to \{1,\ldots,n\}$ *es una permutación* $F(u_{\sigma_{A,B}(1)}) \preceq_{A,B} F(u_{\sigma_{A,B}(2)}) \preceq_{A,B} \ldots \preceq_{A,B} F(u_{\sigma_{A,B}(n)})$.

Si $F(u_i) = [a_i, b_i]$, i = 1, ..., n, entonces la identidad en la definición anterior puede reescribirse como

$$\begin{gathered}\mathbf{C}_m^{A,B}(F) = \\ = \Big[\sum_{i=1}^n a_{\sigma_{A,B}(i)} \cdot (m(\left\{u_{\sigma_{A,B}(i)},\ldots,u_{\sigma_{A,B}(n)}\right\}) - \\ -m(\left\{u_{\sigma_{A,B}(i+1)},\ldots,u_{\sigma_{A,B}(n)}\right\})), \\ \sum_{i=1}^n b_{\sigma_{A,B}(i)} \cdot (m(\left\{u_{\sigma_{A,B}(i)},\ldots,u_{\sigma_{A,B}(n)}\right\}) - \\ - m(\left\{u_{\sigma_{A,B}(i+1)},\ldots,u_{\sigma_{A,B}(n)}\right\}))\Big].\end{gathered}$$

El concepto de integral (A,B)–Choquet $\mathbf{C}_m^{A,B}$ extiende la integral Choquet discreta estándar. De hecho, si F es un subconjunto difuso de U, entonces $C_m(F) = \mathbf{C}_m(F) = \mathbf{C}_m^{A,B}(F)$ independientemente de A, B.

5 UN ALGORITMO DE TOMA DE DECISION MULTICRITERIO

Sean un conjunto de alternativas $\{A_1,\ldots,A_n\}$ y un conjunto de criterios $\{x_1,\ldots,x_k\}$. Supongamos que fusionamos esta información en una matriz intervalo-valorada de toma de decisión como sigue:

$$\begin{array}{c} \\ A_1 \\ A_2 \\ \\ A_n \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & & & x_k \end{array} \\ \begin{pmatrix} ([\underline{\mu_{A_1}}(x_1), \overline{\mu_{A_1}}(x_1)] & \ldots & [\underline{\mu_{A_1}}(x_k), \overline{\mu_{A_1}}(x_k)] \\ ([\underline{\mu_{A_2}}(x_1), \overline{\mu_{A_2}}(x_1)] & \ldots & [\underline{\mu_{A_2}}(x_k), \overline{\mu_{A_2}}(x_k)] \\ \ldots & \ldots & \\ ([\underline{\mu_{A_n}}(x_1), \overline{\mu_{A_n}}(x_1)] & \ldots & [\underline{\mu_{A_n}}(x_k), \overline{\mu_{A_n}}(x_k)] \end{pmatrix} \end{array}$$

donde $[\underline{\mu_{A_i}}(x_j), \overline{\mu_{A_i}}(x_j)]$ denota el grado de satisfacción por la alternativa A_i del criterio x_j. De esta forma, podemos entender cada alternativa como un conjunto intervalo-valorado sobre el referencial de criterios de modo que cada uno de los intervalos $\mu(A_i)(x_j) = [\underline{\mu_{A_i}}(x_j), \overline{\mu_{A_i}}(x_j)]$ proporciona el valor de pertenencia del criterio x_j al conjunto A_i.

El algoritmo que proponemos es el siguiente.

1. Fijar un orden lineal $\preceq_{A,B}$ sobre el conjunto $J([0,1])$.

2. Escoger una medida difusa m sobre el conjunto de criterios $\{x_1,\ldots,x_k\}$.

3. PARA cada fila $i = 1,\ldots,n$ de la matriz de TDM HACER

3.1 Ordenar los elementos en orden creciente $\{x_{(1)},\ldots,x_{x(k)}\}$ con respecto a sus pertenencias a A_i usando el orden lineal fijado en el paso 1.

3.2 Tomar $\underline{C(A_i)} = 0$ y $\overline{C(A_i)} = 0$.

3.3 PARA cada $j = 1,\ldots k$ HACER

3.3.1 $\underline{C(A_i)} = \underline{C(A_i)} + \underline{\mu_{A_i}}(x_{(j)})(m(\{x_{(j)},\ldots,x_{(k)}) - m(\{x_{(j+1)},\ldots,x_{(k)})$

3.3.2 $\overline{C(A_i)} = \overline{C(A_i)} + \overline{\mu_{A_i}}(x_{(j)})(m(\{x_{(j)},\ldots,x_{(k)}) - m(\{x_{(j+1)},\ldots,x_{(k)})$

FIN PARA

3.4 Tomar $\mathbf{C}(A_i) = [\underline{C(A_i)}, \overline{C(A_i)}]$.

FIN PARA

4. Tomar como mejor alternativa aquella para la que $\mathbf{C}(A_i)$ es mayor.

En este algoritmo, los pasos 1 y 2 son cruciales, ya que determinan su comportamiento. Con respecto al orden, en los siguientes ejemplos veremos cómo puede afectar a la decisión final. La elección de uno u otro debe basarse en la información que tenemos sobre las diferentes alternativas a evaluar.

En cuanto a las medidas, este es un punto muy complejo. En futuros trabajos pretendemos llevar a cabo estudios en profundidad sobre cómo determinar cuál es la mejor medida para un problema dado. Obsérvese que la medida debe tomar en cuenta la relación entre los diversos criterios y la importancia relativa concedida a cada uno de ellos.

Si estamos trabajando con un problema de TDM no intervalo-valorado, el algoritmo anterior se reduce al difuso que evalúa cada una de las filas por medio de una integral Choquet. En particular, si consideramos una medida simétrica, evaluamos cada fila por medio de un operador OWA, luego nuestro algorimo extiende, por ejemplo, el del voto ponderado [15].

Mostramos ahora algunos ejemplos de cómo funciona el algoritmo.

Ejemplo 1. Este ejemplo está tomado de [10], que a su vez se basa en un trabajo de [5]. Partimos de

$$\begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} ([0.45,0.65] & [0.50,0.70] & [0.20,0.45] \\ ([0.65,0.75] & [0.65,0.75] & [0.45,0.85] \\ ([0.45,0.65] & [0.45,0.65] & [0.45,0.80] \\ ([0.75,0.85] & [0.35,0.80] & [0.65,0.85] \end{pmatrix} \end{array}$$

Tomamos como orden lineal el obtenido tomando para A la media aritmética y para B la geométrica; es decir, el orden propuesto por Yager y Xu.

Necesitamos fijar la medida difusa sobre el conjunto de criterios. Empezamos utilizando la siguiente medida simétrica.

$$\begin{gathered} m(\emptyset) = 0 \\ m(\{x_1\}) = m(\{x_2\}) = m(\{x_3\}) = \tfrac{1}{3} \\ m(\{x_1,x_2\}) = m(\{x_1,x_3\}) = m(\{x_2,x_3\}) = \tfrac{2}{3} \\ m(\{x_1,x_2,x_3\}) = 1 \end{gathered}$$

m puede verse como una medida de cómo de cerca está el

conjunto medido del conjunto total, es decir, de la situación ideal en la que todos los criterios se satisfacen.

Si ordenamos respecto a $\preceq_{MG}$ las filas de la matriz obtenemos: Para A_1, $\{x_3,x_1,x_2\}$. Para A_2, $\{x_3,x_1,x_2\}$. Para A_3, $\{x_1,x_2,x_3\}$, y para A_4, $\{x_2,x_3,x_1\}$.

Los cálculos correspondientes proporcionan:

$$\mathbf{C}(A_1) = [0.38, 0.50]$$
$$\mathbf{C}(A_2) = [0.58, 0.78]$$
$$\mathbf{C}(A_3) = [0.45, 0.70]$$
$$\mathbf{C}(A_4) = [0.58, 0.83]$$

El orden final de alternativas es A_4,A_2,A_3,A_1. No coincide con el dado en [10],ya que la primera y la segunda alternativas aparecen intercambiadas. Esto se explica porque la aproximación de Ye es completamente diferente, ya que se basa en el uso de entropías y correlación.

Ejemplo 2. Consideremos la matriz siguiente, tomada de [9]:

$$\begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} ([0.70,0.70] & [0.80,0.90] & [0.90,0.90] \\ ([0.60,0.80] & [0.80,0.80] & [0.80,0.90] \\ ([0.60,0.90] & [0.50,0.90] & [0.80,0.80] \\ ([0.40,0.50] & [0.90,0.90] & [0.40,0.90] \end{pmatrix} \end{array}$$

Ordenando respecto a $\preceq_{MG}$ obtenemos para A_1, $\{x_1,x_2,x_3\}$; para A_2, $\{x_1,x_2,x_3\}$; para A_3, $\{x_2,x_1,x_3\}$; y para A_4, $\{x_1,x_3,x_2\}$. Consideremos la medida difusa dada en el mismo trabajo:

$$m(\emptyset) = 0$$
$$m(\{x_1\}) = m(\{x_2\}) = 0.4 \; ; \; m(\{x_3\}) = 0.3$$
$$m(\{x_1,x_2\}) = 0.6 \; ; \; m(\{x_1,x_3\}) = m(\{x_2,x_3\}) = 0.8$$
$$m(\{x_1,x_2,x_3\}) = 1 \; ,$$

que no es simétrica. Si desarrollamos los cálculos correspondientes, llegamos a

$$\mathbf{C}(A_1) = [0.81, 0.86]$$
$$\mathbf{C}(A_2) = [0.76, 0.83]$$
$$\mathbf{C}(A_3) = [0.64, 0.87]$$
$$\mathbf{C}(A_4) = [0.60, 0.73]$$

El orden final de alternativas es A_1,A_2,A_3,A_4, que coincide con el obtenido por Xu.

Consideremos ahora el orden $\preceq_{P2}$. En este caso, el orden de criterios para cada alternativa es: para A_1, $\{x_1,x_2,x_3\}$; para A_2, $\{x_1,x_2,x_3\}$; para A_3, $\{x_3,x_2,x_1\}$; y para A_4, $\{x_1,x_3,x_2\}$. Evaluando cada alternativa obtenemos

$$\mathbf{C}(A_1) = [0.81, 0.86]$$
$$\mathbf{C}(A_2) = [0.76, 0.83]$$
$$\mathbf{C}(A_3) = [0.66, 0.86]$$
$$\mathbf{C}(A_4) = [0.60, 0.82] \; ,$$

Luego llegamos al ránking A_1,A_3,A_2,A_4, que es diferente del obtenido anteriormente.

6 CONCLUSIONES

Hemos introducido el nuevo concepto de integral(*A, B*)–Choquet intervalo-valorada. Este concepto extiende el de integral Choquet clásica y se reduce a él en el caso de intervalos degenerados.

Hemos presentado una forma de generar órdenes totales sobre intervalos en $[0,1]$. Estos órdenes contienen como casos particulares los órdenes lexicográficos y el orden de Yager y Xu basado en el score y el accuracy.

Con estos conceptos hemos presentado un algoritmo de TDM que generaliza diversos algoritmos fuzzy, como el voto. Hemos dado algunos ejemplos.

En el futuro analizaremos el problema de la elección de la medida difusa y el problema de la elección del orden en función de la situación considerada.

Agradecimientos

Este trabajo ha sido apoyado por las becas P402/11/0378, APVV–0073–10 and VEGA 1/0080/10, y el proyecto TIN 2010-15055 del Gobierno de España.

Referencias

[1] Aumann, R. J. (1965). Integrals of set–valued functions. *J. Math. Anal. Appl.* **12**, pp 1–12.

[2] Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J. (2009) Interval-valued fuzzy sets constructed from matrices: Application to edge detection. *Fuzzy Sets Syst.*, **160**, pp. 1819-1840.

[3] Choquet, G. (1953–1954). Theory of capacities. *Ann. Inst. Fourier (Grenoble)* **5**, pp 131–292.

[4] Grabisch, M., Marichal, J.–L., Mesiar, R. and Pap, E. (2009). *Aggregation functions*. Cambridge University Press, Cambridge.

[5] Herrera, F., Herrera-Viedma, E. (2000). Linguistic decision analysis:steps for solving decision problems under linguistic information. *Fuzzy Sets and Systems* **115**, pp. 67–82.

[6] Jang, L.C. (2004). Interval–valued Choquet integrals and their applications. *J. Appl. Math. and Computing* **16**, pp 429–443.

[7] Moore, R.E.(1966). *Interval Analysis*. Prentice Hall.

[8] Tan, C., Chen, X. (2010). Intuitionistic fuzzy Choquet integral operator for multi-criteria decision making. *Expert Systems with Appl.* **37**, pp 149–157.

[9] Xu, Z.S. (2010) Choquet integrals of weighted intuitionistic fuzzy information. *Information Sciences* **180**, pp.726–736.

[10] Ye, J.(2010). Fuzzy decision-making method based on the weighted correlation coefficient under intuitionistic fuzzy environment. *European Journal of Operational Research* **205**, pp. 202–204.

[11] Zadeh, L.A. (1968). Probability measures of fuzzy events. *J. Math. Anal. Appl.* **23**, pp 421–427.

[12] Zadeh, L.A. (1975). The concept of a linguistic variable and its applications to approximate reasoning. *Inform. Sci.* **8**, Part I pp 199–251, Part II pp 301–357, *Inform. Sci.* **9**, Part III pp 43–80.

[13] Zhang, D., Wang, Z. (1993). On set–valued fuzzy integrals. *Fuzzy Sets and Systems* **56**, pp 237–247.

[14] Z. S. Xu and R. R. Yager. (2006) Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General System* **35**, pp 417–433.

[15] H. Bustince, M. Pagola, R. Mesiar, E. Hullermeier, F. Herrera. (2011). Grouping, Overlap and Generalized Bi-Entropic Functions for Fuzzy Modeling of Pairwise Comparisons. *IEEE Transactions on Fuzzy Systems* Doi 10.1109/TFUZZ.2011.2173581, available on line.

TYPE-1 OWA BASED MULTI-GRANULAR CONSENSUS MODEL

Francisco Mata[1], **Francisco Chiclana**[2], **Shang-Ming Zhou**[3]

[1]*Department of Computer Science, University of Jaén, 23700 Jaén, Spain*, fmata@ujaen.es
[2]*Faculty of Technology, De Montfort University, Leicester, LE1 9BH, UK*, chiclana@dmu.ac.uk
[3]*Health Information Research Unit, School of Medicine, Swansea University, SA2 8PP, UK*, smzhou@ieee.org

Abstract

In the multi-granular context, different theoretical models have been proposed in the literature to address the consensus reaching process, in which transformation functions were used to unify the multi-granular linguistic information. A drawback of such unification process is that the consensus reaching process loses information because the transformations functions used are not bijective. In this paper we present a type-1 OWA based multi-granular consensus model that makes the unification step unnecessary and therefore does not suffer from the loss of information in the necessary aggregation step of these GDM problems.

Keywords: Linguistic variable, Muti-granular information, Type-1 OWA operator, Consensus.

1 INTRODUCTION

The linguistic approach mostly used nowadays in human decision processes has its origins in the proposal that Zadeh outlined in his 1973 seminal paper [16] and further elaboration in [17]. In essence, this approach relies on the use of the *linguistic variable*, i.e. a variable whose *primary* values are words instead of numbers. The meaning of the primary values, although assumed to be 'subjective and context-dependent,' is to be specified before other terms are generated from them using linguistic hedges in conjunction with the negator and the connective operators [16, 17]. The main purpose of using linguistic values instead of numbers is that linguistic characterizations are, in general, less specific than numerical ones, but much closer to the way that humans express and use their knowledge [5].

A linguistic variable is formally represented by a 5-tuple $\langle L, T(L), U, S, M\rangle$ [17] where (i) L is the name of the variable; (ii) $T(L)$ is a finite term set of (primary) labels or words (a collection of linguistic values); (iii) U is a universe of discourse or base variable; (iv) S is the syntactic rule which generates the terms in $T(L)$; and (v) M is a semantic rule which associates with each linguistic value X its meaning $M(X) : U \rightarrow [0,1]$. Usually, $T(L)$ is denoted as L when there is no risk of confusion.

The semantic rule, also known as 'compatibility function' [17], associates with each element of the base variable its compatibility with each linguistic value. This interpretation of the meaning of a linguistic label coincides with that of a fuzzy set, and therefore linguistic labels can be considered and formally represented as fuzzy subsets of their base variable. Therefore, the nature of the base variable will dictate the general method to use when manipulating linguistic values. A non numerical base variable makes the definition of the compatibility function 'difficult to be formalized in explicit terms' [17]. As a result, it turns out to be problematic when implemented at present in computer programmes. Thus, it is fair to say that most, if not all, important linguistic decision models in the literature assume that the base variable is a subset of the set of real numbers, and therefore numeric in nature. Indeed, these linguistic decision models usually start associating the linguistic values (labels) to be used with membership functions (triangular, trapezoidal, Gaussian ...) to represent their meanings.

Figure 1 (taken from [5]) illustrates a representation of the linguistic variable '*Height*', whose corresponding set of primary linguistic values is $T(Height) = \{Very\ Low, Low, Medium, High, Very\ High\}$. We can see how the semantic rule associates each of the primary linguistic values X to its fuzzy subset $M(X)$ of U. As mentioned above, a crucial aspect that will determine the validity of the linguistic approach is the determination of correct membership functions for the linguistic term set.

The ideal situation in group decision making (GDM) problems within a linguistic context would be one where all the experts use the same linguistic term set to express their preferences about the alternatives. However, in some cases,

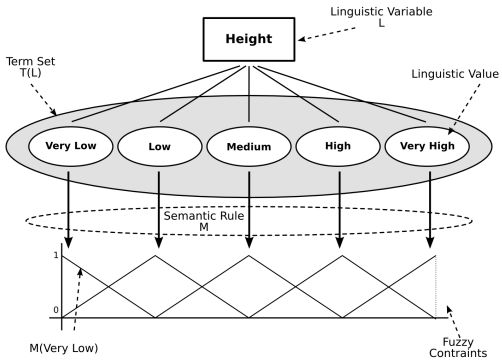


Figure 1: Example of the linguistic variable "Height" [5]

experts may belong to distinct research areas and will, therefore, have different backgrounds and levels of knowledge. A consequence of this is that the expression of preferences will depend on linguistic term sets with different granularity. As a result, the development of adequate tools to manage and model multi-granular linguistic information becomes essential [6, 8, 9].

In this paper we present a consensus model with multi-granular linguistic preferences that follows the recommendation above, i.e. a model that makes use of the whole membership functions used to represent the linguistic terms within the problem and that is built on the use of the extension principle of fuzzy sets [17]. This proposal takes further the model presented in [9], whose methodology was based on transformation functions to unify the multi-granular linguistic information. A drawback of such unification process is that the consensus reaching process loses information because the transformations functions, although made use of the membership functions used to represents the linguistic preference values, are not bijective. Indeed, a different mathematical representation (membership functions) of the linguistic preference values could be sought leading to the same result after the unification process. However, the justification for the necessity of such unification phase was that there were no tools available at that moment to directly compute over linguistic sets of different granularity, while it was a necessary step to manipulate the linguistic sets with different granularity in order to make the computation of both consensus degrees and proximity measures possible [9]. Interestingly, new tools have emerged that allow for the direct manipulation of different elements belonging to linguistic sets of different granularity. In particular, the type-1 OWA operator introduced in [18], and developed by applying the extension principle to Yager's OWA operator [15], has been proved to be a useful tool to aggregate the linguistic opinions or preferences in human decision making with linguistic weights [19]. Thus, our aim is to present a type-1 OWA based multi-granular consensus model that makes the unification step unnecessary and therefore does not suffer from the loss of information in the necessary aggregation step of these GDM problems.

The rest of the paper is set out as follows. Section 2 is devoted to the preliminary concepts needed to present the consensus model: the multi-granular linguistic GDM problem is described in Subsection 2.1, while Subsection 2.2 presents the type-1 OWA operator, its fast implementation and how to derive linguistic weights to be used in type-1 OWA aggregation given a type-2 linguistic quantifier. Section 3 details the architecture of the type-1 OWA based multi-granular consensus model. Finally, in Section 4 we draw our conclusions and suggest further research.

2 PRELIMINARIES

2.1 MULTI-GRANULAR INFORMATION IN GROUP DECISION MAKING

GDM problems are classically described as decision situations where, given a set of feasible alternatives $X = \{x_1, x_2, \ldots, x_n\}$ $(n \geq 2)$, a set of experts $E = \{e_1, e_2, \ldots, e_m\}$ $(m \geq 2)$ try to achieve a common solution together. In a linguistic context, experts' opinions or preferences are mathematically modelled by means of preference relations [13], $\mathbf{P_{e_i}} = (p_i^{lk})$, $l,k \in \{1,\ldots,n\}$, where $p_i^{lk} = \mu_{\mathbf{P_{e_i}}}(x_l, x_k)$ is a fuzzy subset of the unit interval $[0,1]$ and that will represent the expert e_i's preference of the alternative x_l over x_k.

As mentioned above, experts may have different backgrounds and levels of knowledge about the problem to solve and therefore they could prefer to use linguistic term sets with different granularity according to their expertise [6, 7, 14]. The granularity of a linguistic term set should be small enough so as not to impose useless precision levels on experts but big enough to allow a discrimination of the assessments in a limited number of degrees.

In this contribution we deal with GDM problems where each expert e_i may use a distinct linguistic term set with different cardinality and/or semantics $S_i = \{s_0^i, \ldots, s_g^i\}$.

2.2 TYPE-1 OWA OPERATOR

Unlike Yager's OWA operator that aggregates crisp values [15], the type-1 OWA operator is able to aggregate type-1 fuzzy sets with uncertain weights, with these uncertain weights being also modelled as type-1 fuzzy sets. As a generalisation of Yager's OWA operator, and based on the extension principle, a type-1 OWA operator is defined as follows [18]:

Definition 1 *Given n linguistic weights* $\left\{W^i\right\}_{i=1}^n$ *in the form of type-1 fuzzy sets defined on the domain of discourse*

$[0,1]$, *a type-1 OWA operator is a mapping,* Φ,

$$\begin{array}{rcl} \Phi\colon \tilde{P}(\mathbb{R})\times\cdots\tilde{P}(\mathbb{R}) & \longrightarrow & \tilde{P}(\mathbb{R}) \\ (A^1,\cdots,A^n) & \mapsto & Y \end{array}$$

such that

$$\mu_Y(y) = \sup_{\substack{\sum_{k=1}^{n}\bar{w}_i a_{\sigma(i)}=y \\ w_i\in U, a_i\in X}} \left(\begin{array}{c} \mu_{W^1}(w_1)\wedge\cdots\wedge\mu_{W^n}(w_n) \\ \wedge\mu_{A^1}(a_1)\wedge\cdots\wedge\mu_{A^n}(a_n) \end{array}\right) \tag{1}$$

where $\bar{w}_i=\dfrac{w_i}{\sum_{i=1}^{n}w_i}$*;* σ *is a permutation function such that* $a_{\sigma(i)}\geq a_{\sigma(i+1)}$, $\forall i=1,\cdots,n-1$*; and* $\tilde{P}(\mathbb{R})$ *is the set of fuzzy sets on* $\mathbb{R}$.

A Direct Approach to performing type-1 OWA operation was suggested in [18]. However, this approach is computationally expensive, which inevitably curtails further applications of the type-1 OWA operator to real world decision making. So a fast approach to type-1 OWA operations has been developed based on the α-level of fuzzy sets [19].

2.2.1 α-Level Type-1 OWA Operator.

Definition 2 *Given the n linguistic weights* $\left\{W^i\right\}_{i=1}^{n}$ *in the form of type-1 fuzzy sets defined on the domain of discourse* $[0,1]$*, then for each* $\alpha\in[0,1]$*, an* α*-level type-1 OWA operator with* α*-level weight sets* $\left\{W_\alpha^i\right\}_{i=1}^{n}$ *to aggregate the* α*-level of type-1 fuzzy sets* $\left\{A^i\right\}_{i=1}^{n}$ *is given as*

$$\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)=\left\{\left.\frac{\sum_{i=1}^{n}w_i a_{\sigma(i)}}{\sum_{i=1}^{n}w_i}\right| w_i\in W_\alpha^i,\ a_i\in A_\alpha^i,\forall i\right\} \tag{2}$$

where $W_\alpha^i=\{w|\mu_{W_i}(w)\geq\alpha\}$, $A_\alpha^i=\{x|\mu_{A_i}(x)\geq\alpha\}$*, and* σ *is a permutation function such that* $a_{\sigma(i)}\geq a_{\sigma(i+1)}$, $\forall\, i=1,\cdots,n-1$.

According to the Representation Theorem of type-1 fuzzy sets, the α-level sets $\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)$ obtained via Definition 2 can be used to construct the following type-1 fuzzy set on $\mathbb{R}$

$$G=\bigcup_{0<\alpha\leq 1}\alpha\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right) \tag{3}$$

with membership function

$$\mu_G(x)=\bigvee_{\alpha:x\in\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)_\alpha}\alpha \tag{4}$$

2.2.2 Representation Theorem of Type-1 OWA Operators.

The two apparently different aggregation results in (1) and (3) obtained according to Zadeh's Extension Principle and the α-level of type-1 fuzzy sets, respectively, are equivalent as proved in [19]:

Theorem 1 *Given the n linguistic weights* $\left\{W^i\right\}_{i=1}^{n}$ *in the form of type-1 fuzzy sets defined on the domain of discourse* $[0,1]$*, and the type-1 fuzzy sets* $A^1,\cdots,A^n$*, then we have that*

$$Y=G$$

where Y is the aggregation result defined in (1) and G is the result defined in (3).

Theorem 1 is called the *Representation Theorem of Type-1 OWA Operators*. Therefore, an effective and practical way of carrying out type-1 OWA operations is to decompose the type-1 OWA aggregation into the α-level type-1 OWA operations and then reconstruct it via the above representation theorem. This α-level approach has been proved to be much faster than the direct approach [19], so it can be used in real time decision making and data mining applications.

2.2.3 α-Level Type-1 OWA of Fuzzy Numbers.

When the linguistic weights and the aggregated sets are fuzzy number, the α-level type-1 OWA operator produces closed intervals [19]:

Theorem 2 *Let* $\left\{W^i\right\}_{i=1}^{n}$ *be fuzzy numbers on* $[0,1]$ *and* $\left\{A^i\right\}_{i=1}^{n}$ *be fuzzy numbers on* $\mathbb{R}$*. Then for each* $\alpha\in U$, $\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)$ *is a closed interval.*

Based on this result, the computation of the type-1 OWA output according to (3), G, reduces to compute the left end-points and right end-points of the intervals $\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)$:

$$\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)_- \text{ and } \Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)_+,$$

where $A_\alpha^i=[A_{\alpha-}^i,A_{\alpha+}^i]$, $W_\alpha^i=[W_{\alpha-}^i,W_{\alpha+}^i]$.

For the left end-points, we have

$$\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)_-=\min_{\substack{W_{\alpha-}^i\leq w_i\leq W_{\alpha+}^i \\ A_{\alpha-}^i\leq a_i\leq A_{\alpha+}^i}}\sum_{i=1}^{n}w_i a_{\sigma(i)}\Big/\sum_{i=1}^{n}w_i \tag{5}$$

while for the right end-points, we have

$$\Phi_\alpha\left(A_\alpha^1,\cdots,A_\alpha^n\right)_+=\max_{\substack{W_{\alpha-}^i\leq w_i\leq W_{\alpha+}^i \\ A_{\alpha-}^i\leq a_i\leq A_{\alpha+}^i}}\sum_{i=1}^{n}w_i a_{\sigma(i)}\Big/\sum_{i=1}^{n}w_i \tag{6}$$

It can be seen that (5) and (6) are programming problems. Solutions to these problems, so that the type-1 OWA aggregation operation can be performed efficiently, are available from [19].

3 TYPE-1 OWA BASED MULTI-GRANULAR CONSENSUS MODEL

Consensus reaching processes can be defined as iterative processes composed by several rounds where experts express and discuss about their preferences with the aim to achieve a minimum level of agreement before making a decision. In real-world problems, these processes are guided by a human moderator who is in charged of supervising all process phases and carries out the necessary actions to drive the consensus process toward success.

In the multi-granular context, the different theoretical models proposed in the literature [2, 10, 11] have been proposed to address the consensus reaching process, in which transformation functions were used to unify the multi-granular linguistic information. As pointed our earlier, a drawback of such unification process is that the consensus reaching process loses information because the transformations functions are not bijective. These proposed models did not address the issues of direct manipulation of different elements belonging to linguistic sets of different granularity, specially in the necessary aggregation step of GDM problems, because there were no mathematical tools available at that moment. Interestingly, the introduction of the type-1 OWA operator provides such a needed tool for direct manipulating linguistic sets with different granularity in consensus decision making.

In this section we present, and illustrate in 2, a consensus reaching model based on type-1 OWA operator to address GDM problems defined in multi-granular linguistic contexts. Specifically our proposed model includes the following steps:

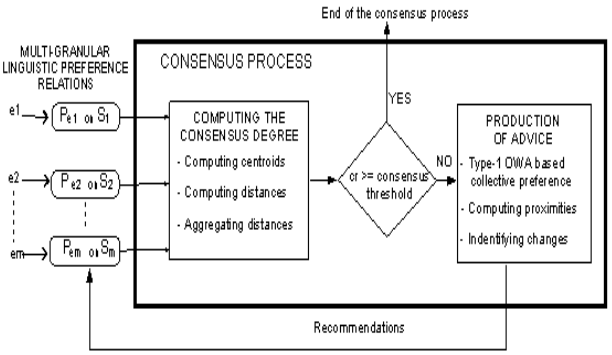


Figure 2: Type-1 OWA based multi-granular consensus model phases

1. **Computing the consensus degree**. The consensus degree will represent the level of agreement among experts, i.e., how close experts' preferences are. In order to evaluate the agreement, the model computes and aggregates the 'distance' representing disagreement between the preferences of each pair of experts on each pair of alternatives. It is well known that the set of fuzzy sets of the unit interval is not totally ordered. Therefore, for our purpose, following the methodology proposed in [4], each fuzzy set will be associated to a numeric representative score, in our case its centroid, and the distance between two fuzzy sets will be set as the distance between their respective scores.

   The centroid of a type-1 fuzzy set A in a continuous domain X is calculated as,

$$C_A = \frac{\int_x x \cdot \mu_A(x) dx}{\int_x \mu_A(x_i)}. \tag{7}$$

   The centroid when the domain X is discretised into n points is

$$C_A = \frac{\sum_{i=1}^{n} x_i \cdot \mu_A(x_i)}{\sum_{i=1}^{n} \mu_A(x_i)}. \tag{8}$$

   The centroid of linguistic label p_i^{lk} will be denoted by c_i^{lk}.

   For each pair of experts e_i, e_j $(i < j)$, a similarity matrix, $SM_{ij} = (sm_{ij}^{lk})$, is obtained as

$$sm_{ij}^{lk} = 1 - \mid c_i^{lk} - c_j^{lk} \mid \tag{9}$$

   where sm_{ij}^{lk} represent the similarity between the preferences of the experts e_i and e_j on the pair of alternatives (x_l, x_k). The closer sm_{ij}^{lk} to 1 the more similar p_i^{lk} and p_j^{lk} are considered, while the closer sm_{ij}^{lk} to 0 the more distant p_i^{lk} and p_j^{lk} are considered.

   A consensus matrix, $CM = (cm^{lk})$, is calculated by aggregating all similarity matrices, where each aggregated pair of alternatives is computed as:

$$cm^{lk} = \theta(sm_{12}^{lk}, sm_{13}^{lk}, \ldots, sm_{1m}^{lk}, sm_{23}^{lk}, \ldots, sm_{2m}^{lk}, \ \ldots, sm_{(m-1)m}^{lk})$$

   for $l,k \in \{1,\ldots,n\}$ and where θ is an aggregation operator. Note that different aggregation operators can be used according to different consensus strategies (more details on this can be found in [12]).

   The level of agreement is computed at three different levels [9]:

   i) Pairs of alternatives,

$$cp^{lk} = cm^{lk}, \ \forall\, l,k = 1,\ldots,n \ \wedge \ l \neq k,$$

   where cp^{lk} represents the level of agreement by all experts on the pair of alternatives (x_l, x_k).

ii) Alternatives,

$$ca^l = \frac{\sum_{k=1,\, l\neq k}^{n} cp^{lk}}{(n-1)}, \qquad (10)$$

where ca^l represents the level of agreement by all experts on the alternative x_l.

iii) Preference relation,

$$cr = \frac{\sum_{l=1}^{n} ca^l}{n}, \qquad (11)$$

where cr represents the global agreement among all experts.

2. **Consensus control.** The agreement cr obtained in the previous phase is checked here against a consensus threshold agreed by the set of expert previous to the application of the consensus process. If cr is greater or equal than this consensus threshold then the consensus reaching process is considered over. Otherwise, the experts need to discuss and attempt to get their opinions closer for their level of agreement to increase and achieve the threshold value set.

3. **Advice generation.** In this phase, the model suggests how experts should change their opinions in order to bring them preferences closer and to increase the level of agreement. Three tasks are carried out in order to achieve this:

   (a) *Computing the collective preference and the experts' proximity values.* Firstly, a collective preference $\mathbf{P_{e_c}} = (p_c^{lk})$ is calculated by aggregating all experts' preference relations $\{\mathbf{P_{e_1}}, \ldots, \mathbf{P_{e_m}}\}$ at level of pairs of alternatives. This aggregation is carried out by using the type-1 OWA operator.

   $$\mathbf{P_c} = \Phi(\mathbf{P_{e_1}}, \ldots, \mathbf{P_{e_m}}). \qquad (12)$$

   The proximity between each individual linguistic preference relation and the collective linguistic one is obtained, $PM_i = (pm_i^{lk})$,

   $$pp_i^{lk} = 1 - \mid c_i^{lk} - c_c^{lk} \mid \qquad (13)$$

   where pp_i^{lk} represents the proximity between the preference of expert e_i and the collective one on the pair of alternatives (x_l, x_k). These proximity values are used by the model to identify the furthest individual preference values from the collective ones, and therefore the preference values that should be considered by the group of experts for possible changes.

   (b) *Identification of preferences to be changed.* The identification of the preference values to be subject of modification is done by considering both the consensus degrees and the proximity values. The alternatives with a level of agreement by all experts, ca^l, lower than the consensus threshold are identified, and for these alternatives the preference values with a level of agreement by all experts, cp^{lk}, lower than the threshold will be subject of a recommendation of change:

   $$SC = \{(l,k) \in \{1, \ldots, n\} \mid \max\{ca^l, cp^{lk}\} < cr\}$$

   The recommendation of change will be produced for just those experts that are furthest from the whole group in the identified elements of SC. This is done by comparing, for each $(l,k) \in SC$, pp_i^{lk} with a collective threshold $\overline{pp}^{lk}$ computed by aggregating all individual proximity values:

   $$\overline{pp}^{lk} = \theta(pp_1^{lk}, \ldots, pp_m^{lk}). \qquad (14)$$

   Those experts for which $pp_i^{lk} < \overline{pp}^{lk}$ will be provided with a recommendation of the direction of change of their current preference values associated to $(l,k) \in SC$.

   (c) *Direction changes.* For each preference to be changed, the model will suggest increasing or decreasing its current assessment [11]:

   DR.1. If $(c_i^{lk} - c_c^{lk}) < 0$, then expert e_i should increase the linguistic assessment associated to the pair of alternatives (x_l, x_k).

   DR.2. If $(c_i^{lk} - c_c^{lk}) > 0$, then expert e_i should decrease the linguistic assessment associated to the pair of alternatives (x_l, x_k).

   DR.3. If $(c_i^{lk} - c_c^{lk}) = 0$, then expert e_i should not modify the linguistic assessment associated to the pair of alternatives (x_l, x_k).

4 CONCLUSIONS

In this contribution we have presented a consensus modelled for multi-granular linguistic GDM problems that is based on the use of the type-1 OWA operator. This operator allows the direct aggregation of fuzzy sets, therefore making it superfluous to unify the multi-granular information as proposed in previous models. An advantage of the proposed method is that it avoids the loss of information that the mentioned methods incurred. One further issue is to carry out a comparative study of the proposed method with previous ones, in order to ascertain the extent of the influence that the type-1 OWA operators brings about in the measuring of consensus.

Acknowledgements

This work has been supported by the Research Projects UJA2009/12/25, TIN2009-08286 and P08-TIC-3548.

References

[1] Z. Chen and D. Ben-Arieh. "On the fusion of multi-granularity linguistic label sets in group decision making," *Computers and Industrial Engineering* 51(3):526–541, 2006.

[2] F. Chiclana, F. Mata, L. Martínez, E. Herrera-Viedma and S. Alonso, "Integration of a consistency control module within a consensus model," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16(Supplement 1): 35–53, 2008.

[3] F. Chiclana and S.-M. Zhou, "The type-1 OWA operator and the centroid of type-2 fuzzy sets." In *EUSFLAT-LFA 2011 Proceedings*, 2011, pp. 15–20.

[4] J. L. García-Lapresta, "Weighting individual opinions in group decision making." In *Proceedings of MDAI 2007, Lecture Notes in Artificial Intelligence (LNAI) 4617,* 80-91, 2007.

[5] F. Herrera, E. Herrera-Viedma, S. Alonso and F. Chiclana, "Computing with words in decision making: foundations, trends and prospects," *Fuzzy Optimization and Decision Making* 8: 337–364, 2009.

[6] F. Herrera, E. Herrera-Viedma and L. Martínez, " A fusion approach for managing multi-granularity linguistic term sets in decision making," *Fuzzy Sets and Systems* 114: 43–58, 2000.

[7] F. Herrera and L. Martínez, "A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making," *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics* 31(2): 227–234, 2001.

[8] E. Herrera-Viedma, O. Cordón, M.Luque, A.G. López and A. N. Muñoz, "A model of fuzzy linguistic IRS based on multi-granular linguistic information," *International Journal of Approximate Reasoning* 34(3): 221–239, 2003.

[9] E. Herrera-Viedma, F. Mata, L. Martínez and F. Chiclana, "A consensus support system model for group decision-making problems with multi-granular linguistic preference relations," *IEEE Transactions on Fuzzy Systems* 13(5): 644–658, 2005.

[10] J. Kacprzyk, M. Fedrizzi and H. Nurmi, *Consensus Under Fuzziness*. Kluwer Academic Publishers, 1997.

[11] F. Mata, L. Martínez and E. Herrera-Viedma, "An adaptive consensus support system model for group decision-making problems in a multigranular fuzzy linguistic context," *IEEE Transactions on Fuzzy Systems* 17(2): 279–290, 2009.

[12] F. Mata, L. Martínez and J. C. Martínez, "A preliminary study of the effects of different aggregation operators on consensus processes." In *Proceedings of the 9th International Conference on Intelligent Systems Design and Applications, ISDA'09*, Pisa, Italy, 2009, pp. 821–826.

[13] T. Tanino, "Fuzzy preference orderings in group decision making," *Fuzzy Sets and Systems*, 12: 117–131, 1984.

[14] S. Wang, "Applying 2-tuple multigranularity linguistic variables to determine the supply performance in dynamic environment based on product-oriented strategy," *IEEE Transactions on Fuzzy Systems* 16(1): 29–39, 2008.

[15] R. R. Yager, "On ordered weighted averaging aggregation operators in multi-criteria decision making," *IEEE Transactions on Systems, Man and Cybernetics* 18: 83–190, 1988.

[16] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. on Systems, Man and Cybernetics* 3: 28–44, 1973.

[17] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning – I," *Information Sciences* 8: 199–249, 1975.

[18] S. -M. Zhou, F. Chiclana, R. I. John and J. M. Garibaldi, "Type-1 OWA operators for aggregating uncertain information with uncertain weights induced by type-2 linguistic quantifiers," *Fuzzy Sets and Systems* 159(24): 3281–3296, 2008.

[19] S. -M. Zhou, F. Chiclana, R. I. John and J. M. Garibaldi, " Alpha-level aggregation: a practical approach to type-1 OWA operation for aggregating uncertain information with applications to breast cancer treatments," *IEEE Transactions on Knowledge and Data Engineering* 23(10): 1455–1468, 2011.

MÉTODO DIFUSO PARA IDENTIFICACIÓN DE VARIABLES RELEVANTES EN PROSPECTIVA TECNOLÓGICA

Pablo Villacorta, **Dagoberto Castellanos**[2], **Antonio Masegosa**

Grupo de Modelos de Decisión y Optimización, CITIC-UGR,
Dpto. de Ciencias de la Computación e IA, Universidad de Granada,
{pjvi,admase}@decsai.ugr.es
[2], dcastellanos@ugr.es

Resumen

La prospectiva tecnológica puede definirse como un conjunto de estudios que se llevan a cabo con el fin de anticipar cuál va a ser el futuro en una determinada área. Una metodología ampliamente utilizada para ello es el Método de los Escenarios de Godet, que incluye un módulo para realizar el conocido como análisis estructural (MICMAC). Su finalidad es determinar las variables más importantes en el sistema que se va a estudiar, tomando como entradas las relaciones de influencia existentes entre ellas, las cuales vienen cuantificadas numéricamente por un grupo de expertos en forma de matriz. Proponemos una extensión difusa de este método que permite a los expertos juzgar las relaciones mediante etiquetas lingüísticas y obtener resultados más fácilmente intepretables. Esto es una importante ventaja sobre el método tradicional, ya que en este tipo de estudios, la interpretabilidad de las salidas y los resultados intermedios es un aspecto primordial. Se ha realizado una implementación en Java de nuestro procedimiento difuso con una interfaz web accesible públicamente, y que demuestra la facilidad de uso del enfoque planteado. Los primeros resultados obtenidos usando un caso real son interesantes desde el punto de vista práctico.

Palabras Clave: Prospectiva tecnológica, Método de escenarios, MICMAC, Variable lingüística.

1 INTRODUCCIÓN

La prospectiva tecnológica puede definirse como un conjunto de estudios que se llevan a cabo con el fin de anticipar cuál va a ser el futuro en una determinada área. El objetivo de un estudio de este tipo es identificar acciones que se puedan llevar a cabo en la actualidad para que ese futuro se parezca al más deseado por quien encarga el estudio. En el mundo empresarial actual, esta anticipación puede significar en muchas ocasiones la diferencia entre el éxito y el fracaso.

Para conseguirla, es necesario identificar los posibles escenarios futuros. Una de las metodologías más frecuentes es el Método de los Escenarios [6, 4, 5, 7] , propuesto por M. Godet, del Laboratorio de Investigación en Prospectiva, Estrategia y Organización (LIPSOR). Se ha utilizado en problemas de prospectiva [1] pero también en otros contextos como diseño de interfaces de usuario [10]. Esta metodología determina las posibles situaciones futuras mediante la definición de escenarios. Un escenario es una representación de un estado futuro y los pasos que llevan hasta él. Los escenarios se desarrollan a través de información obtenida mediante paneles de expertos, encuestas, etc.

El Método de Escenarios de Godet establece diferentes etapas para cada una de las cuales proporciona herramientas específicas. Estas etapas son el análisis estructural, análisis de los actores, análisis morfológico y análisis multicriterio. En todas estas técnicas, la información que se suele utilizar proviene habitualmente de expertos, y a veces es difícil cuantificarla con exactitud. Por este motivo, el empleo de técnicas de Soft computing que traten con información imprecisa o expresada en palabras para muchos de estos cálculos puede ser una opción adecuada y natural, como ya se ha indicado en trabajos existentes [2, 9]. En este estudio nos hemos centrado en la etapa de análisis estructural por ser cronológicamente la primera dentro de la metodología. Propondremos una extensión al método de análisis estructural basada en una variable lingüística para valorar las influencias de unas variables sobre otras. Dicha influencia es especificada por un grupo de expertos humanos y se utiliza como entrada de dicho método. La utilización de etiquetas en lugar de números facilita la interacción con los expertos y permite que expresen más fácilmente sus ideas. Además, el significado o matiz de cada etiqueta puede ser capturado con precisión gracias a las estructuras matemáticas subyacentes y ajustables que permiten realizar los cálculos, como

se describirá más adelante.

El artículo se organiza como sigue. En primer lugar, en la sección 2 se explicará el método MICMAC propuesto originalmente por Godet. A continuación, se motivará en la sección 3 una extensión basada en etiquetas lingüísticas y conjuntos difusos para realizar los cálculos del método. La sección 4 muestra un ejemplo de aplicación sobre la interfaz web desarrollada a tal efecto. Finalmente, se presentarán las conclusiones en la sección 5.

2 EL MÉTODO MICMAC

El método de Godet para el análisis estructural se denomina MICMAC (Matriz de Impactos Cruzados Multiplicación Aplicada a una Clasificación). Su finalidad es descubrir qué variables son más relevantes para el sistema en estudio, basándose en la influencia de unas variables en otras. Puede utilizarse solo o como parte de un proceso de prospectiva más complejo. En la web[1] se puede encontrar una descripción en español muy concisa de las etapas que lo conforman y que se explican a continuación.

En primer lugar, un panel de expertos propone una lista de n variables que son las que en principio intervienen en el estudio. Con ellas los expertos construyen una matriz de dimensiones n x n, en base a sus conocimientos y a su experiencia. Se conoce como *Matriz de Influencias Directas* (MDI), y cada elemento MDI_{ij} indica el grado de influencia que la variable i-ésima ejerce sobre la variable j-ésima. Este valor es un entero entre 0 y 4 con el siguiente significado:

- 0 si la variable i no tiene influencia sobre j
- 1 si la variable i tiene poca influencia sobre j
- 2 si la variable i tiene una influencia media sobre j
- 3 si la variable i tiene mucha influencia sobre j
- 4 si la variable i tiene una influencia potencial sobre j

Una relación potencial significa que la influencia no está clara actualmente pero podría hacerse patente bajo ciertas condiciones. Por ello Godet descarta este tipo de relaciones durante el proceso de análisis (se toma como 0 para un tipo de análisis, y se sutituye por 1, 2 ó 3, a elección del usuario, para otro tipo de análisis). La diagonal principal de la matriz es 0 ya que se considera que una variable no ejerce influencia sobre sí misma. Según Godet, en estudios prospectivos reales, sólo alrededor del 30 % de las casillas de la matriz son distintas de 0.

Podemos obtener medidas de interés a partir de la matriz MDI. La *influencia directa* de una variable i sobre el resto se define como la suma de todos los valores de la fila i de la matriz. En dicha suma, las relaciones potenciales son tomadas como 0. De forma análoga, se define la *dependencia directa* de una variable i respecto al resto de variables como la suma de todos los valores de la columna i. En ambos casos, las relaciones potenciales son tomadas como 0. Así, se tienen dos medidas distintas asociadas a cada variable. Con dicha información, es posible ordenar las variables para construir dos rankings independientes, uno en base a la influencia de la variable y otro en base a la dependencia. Ambos sirven como una primera indicación de la importancia de cada variable. El cálculo de la influencia y la dependencia directas a partir de la matriz MDI es lo que se conoce como *Método Directo*.

Otras medidas interesantes utilizadas en MICMAC son la *influencia indirecta* y la *dependencia indirecta*. Se calculan sobre otra matriz diferente a la MDI mediante el conocido como *método indirecto*, que es un proceso iterativo en 2 etapas:

1. Inicialización. Sean R_f y R_d los rankings de influencia y dependencia directa construidos con el método directo. Sea M_{nxn} una matriz inicializada con los mismos valores de la matriz MDI.
2. Paso iterativo:
   - Hacer $M \leftarrow M$ x MDI, y calcular sobre M los nuevos rankings de influencia y dependencia. Sean S_f y S_d dichos rankings
   - Comparar S_f con R_f, y S_d con R_d.
   - Si ambas comparaciones coinciden, terminar. En caso contrario, actualizar los rankings: hacer $R_f \leftarrow S_f$ y $R_d \leftarrow S_d$ e ir al paso 2 de nuevo.

Se espera que los rankings coincidan en aproximadamente 8 iteraciones, lo que significa que la matriz que estaremos usando para calcular los rankings finales será MDI^8, la octava potencia de la matriz MDI. La influencia y dependencia indirectas también son indicadoras de la relevancia global de una variable dentro del sistema.

2.1 CLASIFICACIÓN DE LAS VARIABLES

Las medidas asociadas a cada variable descritas en la sección anterior permiten elaborar gráficos muy informativos para resumir la información, como diagramas 2D de influencia/dependencia, donde cada punto del espacio representa una variable cuyas coordenadas cartesianas son el par (influencia, dependencia). La Figura 1 muestra un ejemplo de ello. El eje horizontal representa el grado de dependencia y el vertical el grado de influencia. No se daran detalles por motivos de espacio, pero la clasificación mostrada está tomada de [3] donde se describe el papel que suele tener una variable situada en cada uno de los cuadrantes.

[1] http://es.laprospective.fr/Metodos-de-prospectiva/Los-programas/67-MICMAC.html

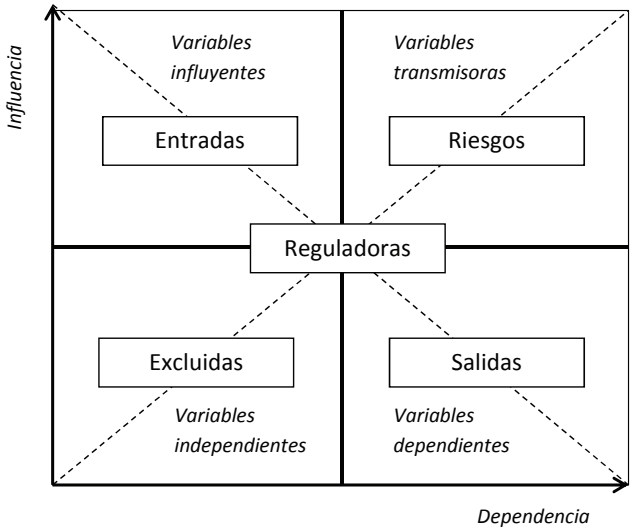


Figura 1: Representación de la influencia y dependencia de las variables mediante ejes cartesianos

3 EL MÉTODO MICMAC DIFUSO

La idea principal es facilitar la implantación de MICMAC para permitir al usuario dar valoraciones cualitativas en vez de cuantitativas en la matriz, y utilizar dichos valores cualitativos en los cálculos para poder obtener también salidas cualitativas, más informativas. Revisaremos brevemente

3.1 VARIABLES LINGÜÍSTICAS

En numerosas ocasiones los expertos no son capaces de expresar numéricamente sus juicios. Por tanto, el uso de variables lingüísticas es una buena solución. Desde que Zadeh introdujo los conjuntos difusos y los extendió con el concepto de variable lingüística [11], la popularidad de este campo ha crecido enormemente. Nos interesan las variables lingüísticas para establecer una escala ordinal, junto con su representación matemática asociada como número difuso. Se utiliza un modelo computacional lingüístico basado en funciones de pertenencia [8], donde los cálculos se efectúan con la aritmética difusa habitual para números difusos triangulares (productos y sumas componente a componente) .

Una variable lingüística es una variable cuyos valores son palabras o frases en un lenguaje natural o artificial. Formalmente, es una quintupla {X, T(X), U, G, M} donde

1. X es el nombre de la variable
2. T(X) es el conjunto de términos de X, es decir las etiquetas lingüísticas que X puede tomar como valores
3. U es el universo del discurso
4. G es una regla sintáctica para generar los elementos de T(X)
5. M es una regla semantica para asociar significado a los valores lingüísticos

Para aplicarlo al MICMAC, es necesario que T(X) sea un conjunto ordenado, es decir que exista un orden estricto entre los posibles valores lingüísticos de X de manera que todos sean comparables entre sí. En lo que respecta a los números (conjuntos) difusos que soportan a estas etiquetas, nos centraremos por el momento en números difusos triangulares. La función de pertenencia de un conjunto difuso triangular A viene definida por tres números reales $a \leq b \leq c$ y adopta la forma

$$f_A(u) = \begin{cases} (u-a)/(b-a) & a \leq x \leq b \\ (c-u)/(c-b) & b < x \leq c \\ 0 & \text{en otro caso} \end{cases} \quad (1)$$

El número difuso con la función de pertenencia anterior puede representarse como una tripleta (a,b,c) cuyas componentes se utilizan para las operaciones aritméticas.

3.2 ANÁLISIS ESTRUCTURAL DIFUSO

En general, el análisis estructural es más fácil cuando se describe la influencia y dependencia en términos de etiquetas lingüísticas. En nuestro caso, utilizaremos la variable lingüística X = "Influencia de una variable sobre otra". Los expertos utilizarán esta variable para definir la matriz de influencias MDI, y los cálculos se realizarán gracias a los números difusos que soportan a dichas etiquetas. Los números difusos permiten efectuar operaciones matemáticas de un modo análogo al habitual con números reales. Dado que la base del método MICMAC son operaciones aritméticas sencillas, es posible llevarlo a cabo también utilizando números difusos.

Para ello en primer lugar se debe definir el conjunto T(X) de etiquetas lingüísticas para valorar la influencia, que serán utilizadas por los expertos para construir la matriz difusa MDI. En una primera aproximación podemos usar las etiquetas {Sin influencia, Débil, Fuerte, Muy fuerte} (Figura 2) para replicar las mismas utilizadas por Godet en su enfoque numérico. No obstante, en siguientes estudios probablemente debería tomarse un número impar de etiquetas para que haya el mismo número distribuidas en torno a una etiqueta central. Tanto la expresión concreta (valores de los números a, b y c) de las funciones de pertenencia de los números difusos triangulares que soportan a estas etiquetas como el universo del discurso pueden estar definidos por defecto, o pueden dejarse a la elección del usuario. Así, la matriz MDI pasa a ser una matriz difusa como se muestra en la Tabla 1.

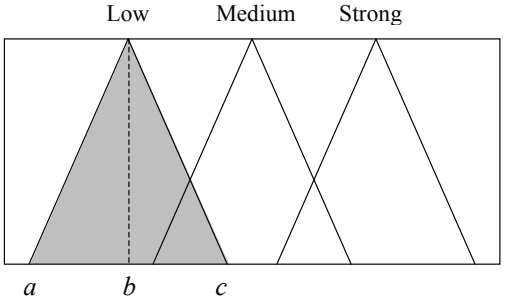


Figura 2: Etiquetas lingüísticas con sus números difusos triangulares asociados

Tabla 1: Matriz MDI difusa de etiquetas lingüísticas

	V_1	V_2	...	V_n
V_1	Sin influencia	Débil	$\cdots$	Muy fuerte
V_2	Fuerte	Sin influencia	$\cdots$	Fuerte
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
V_n	Baja	Sin influencia	$\ldots$	Sin influencia

Una vez introducidos estos conceptos, el desarrollo del método MICMAC es análogo al que ya se ha explicado, utilizando números difusos en las operaciones en lugar de números enteros. El problema de la re-traducción de los números difusos resultantes de nuevo a etiquetas lingüísticas se ha solventado definiendo un nuevo universo del discurso para los resultados y efectuando un reescalado. En principio, se definen tantas etiquetas de salida como etiquetas diferentes haya en la matriz MDI de entrada. Se definen las funciones de pertenencia de las etiquetas de salida basándose en el concepto de matriz *ideal* difusa. Se construye una matriz ideal diferente para cada una de las posibles etiquetas de salida, sustituyendo todas las etiquetas no nulas de la matriz MDI por esa etiqueta. Aplicando el algoritmo del MICMAC a la matriz tras estos reemplazamientos se obtienen los 3 números que definirán la función de pertenencia A modo de ejemplo, la matriz ideal para la etiqueta Débil se construye tomando la matriz MDI original y sustituyendo todos los elementos no nulos por la etiqueta Débil.

4 EJEMPLO PRÁCTICO

Para poder comparar los resultados, se ha utilizado como ejemplo el incluido en el software MICMAC desarrollado por el grupo LIPSOR y que puede descargarse libremente[2]. Dicho ejemplo es un estudio de prospectiva sobre los determinantes de los espacios rurales en el horizonte temporal de 2010. El sistema utiliza 50 variables; enumeramos a continuación algunas de ellas a título informativo.

[2]http://es.laprospective.fr/Metodos-de-prospectiva/Descarga-de-aplicaciones.html

1. Metropolización
2. Organización del mercado internacional
3. Demanda de alimentos
4. Contribución de la migración
5. Mercado laboral
6. Tercera edad
7. Política social

4.1 IMPLEMENTACIÓN DEL MÉTODO CON INTERFAZ WEB

El método MICMAC difuso se ha implementado y se ha incorporado una interfaz web para que los usuarios puedan suministrar sus datos de entrada y obtener las salidas del método en su navegador. La Figura 3(a) muestra un fragmento de la interfaz mencionada donde se está introduciendo como entrada el ejemplo anterior. Aunque no se muestra, más abajo en la misma página se le pide al usuario que escriba la matriz de etiquetas, y por otro lado la lista de etiquetas lingüísticas que se han utilizado en la matriz, ordenadas de menor a mayor para establecer así el orden en dicho conjunto, necesario para realizar los cálculos. El universo del discurso y las funciones de pertenencia se definen automáticamente de manera adecuada al número de etiquetas indicado por el usuario, siempre utilizando números difusos triangulares. En un futuro, se permitirá que el propio usuario especifique estos detalles si lo desea. De esta manera, es posible dar una valoración lingüística a la influencia y a la dependencia directa e indirecta de cada variable, en lugar de obtener simplemente un número como ocurre con el método original de Godet.

4.2 COMPARACIÓN DE LOS RESULTADOS CON LOS DEL MÉTODO ORIGINAL

Se han comparado los resultados obtenidos con MICMAC difuso y los obtenidos con el método original según la implementación proporcionada por los autores. Los rankings obtenidos en cada caso se detallan en las Tablas 2, 3, 4, 5, donde por motivos de espacio sólo se muestran las 8 primeras variables de cada ranking.

En primer lugar se observa que la coincidencia entre los rankings directos es total salvo las posiciones 7 y 8 del ranking de influencia directa que están intercambiadas. Esto ya indica que el cálculo con etiquetas lingüísticas funciona adecuadamente y que tiene sentido la extensión difusa del método. Además, se observa una coincidencia también muy cercana en los rankings indirectos, a pesar de que el método de cálculo difuso lleva a cabo operaciones bastante diferentes a las empleadas con números crisp. Los intercambios que hay se producen en posiciones bastante cercanas respecto a las originales, confirmando que el cálculo

Original	Influenc.	Difuso	Influenc.
1. Estrategia	59	1. Estrategia	Muy fuerte
2. Mercado	52	2. Mercado	Muy fuerte
3. Migración	43	3. Migración	Fuerte
4. Turismo	41	4. Turismo	Fuerte
5. Oferta trab.	38	5. Oferta trab.	Fuerte
6. Progreso	35	6. Progreso	Débil
7. Imagen	31	7. Juventud	Débil
8. Juventud	30	8. Imagen	Débil

Tabla 2: Ranking de influencia directa

Original	Influenc.	Difuso	Influenc.
1. Estrategia	329E3	1. Estrategia	Muy fuerte
2. Oferta trab.	306E3	2. Turismo	Fuerte
3. Mercado	287E3	3. Mercado	Fuerte
4. Turismo	277E3	4. Migración	Fuerte
5. Mundo rural	268E3	5. Juventud	Fuerte
6. Migración	248E3	6. Oferta trab.	Débil
7. Conc. local	234E3	7. Mundo rural	Débil
8. Juventud	222E3	8. Conc. local	Débil

Tabla 3: Ranking de influencia indirecta

Original	Depend.	Difuso	Depend.
1. Estrategia	48	1. Estrategia	Muy fuerte
2. Entretenim.	45	2. Entretenim.	Fuerte
3. Turismo	40	3. Turismo	Fuerte
4. Sociedad	39	4. Sociedad	Fuerte
5. Mercado	34	5. Mercado	Débil
6. Imagen	33	6. Imagen	Débil
7. Conflictos	31	7. Conflictos	Débil
8. Cultura	29	8. Cultura	Débil

Tabla 4: Ranking de dependencia directa

Original	Depend.	Difuso	Depend.
1. Estrategia	336E3	1. Estrategia	Muy fuerte
2. Sociedad	329E3	2. Sociedad	Muy fuerte
3. Cultura	327E3	3. Entretenim.	Muy fuerte
4. Imagen	324E3	4. Imagen	Fuerte
5. Entretenim.	320E3	5. Conflictos	Fuerte
6. Conflictos	318E3	6. Mercado	Débil
7. Turismo	303E3	7. Turismo	Débil
8. Mercado	284E3	8. Cultura	Débil

Tabla 5: Ranking de dependencia indirecta

con números difusos triangulares resulta efectivo ya que se llegan a conclusiones muy similares y no se encuentra ninguna diferencia brusca en variables aisladas, incluso después de un proceso largo de cálculos encadenados como es el método indirecto. Sin embargo, nótese que las etiquetas facilitan la valoración por parte de los expertos, proporcionan valoraciones lingüísticas absolutas de influencia y dependencia de cada variable y, ademas, es posible ajustar los parámetros de cada número difuso triangular para que represente con precisión el concepto recogido por cada etiqueta, por lo que el método propuesto es muy ventajoso.

5 CONCLUSIONES

Se ha presentado una extensión difusa al método MICMAC para análisis estructural en prospectiva tecnológica. El método permite a los expertos dar valoraciones lingüísticas para expresar sus opiniones acerca de la influencia de unas variables sobre otras. La teoría de conjuntos difusos permite llevar a cabo un proceso de cálculo similar al original utilizando estas etiquetas, y se ha comprobado que los resultados a los que llegan ambos métodos son muy similares. Por otro lado, se ha implementado el método MICMAC difuso en un sitio web mediante servlets de manera que pueda ser usado libremente desde un navegador. Además, se ha obtenido una valoración también lingüística de la influencia y la dependencia absolutas de cada variable, tanto con el método directo como con el indirecto. Esto facilita la comprensión de la importancia relativa de las variables, más allá de la posición que ocupen en cada ranking, y justifica desde otra perspectiva la utilización del método difuso. Como trabajo futuro se propone el desarrollo de un método para la representación gráfica de la influencia y dependencia difusas de las variables, con el objetivo de obtener gráficos similares a la Fig. 1 construidos sobre ejes fuzzificados.

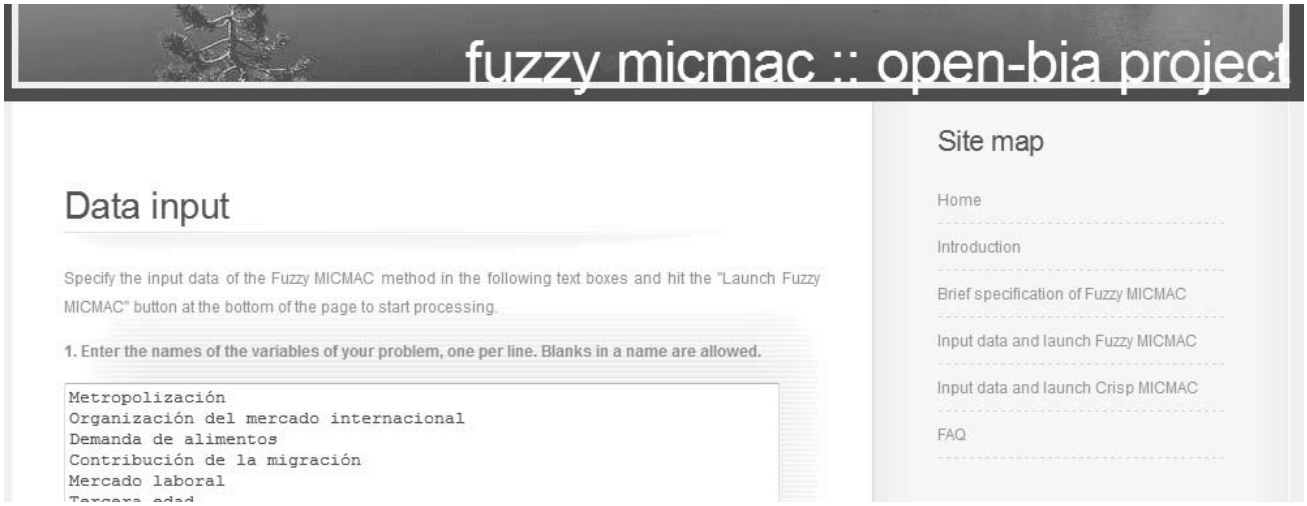
Agradecimientos

Este trabajo ha contado con el soporte de los proyectos TSI-020513-2009-74 del Ministerio de Industria, Turismo y Comercio, TIN2008 - 06872 - C04 - 04 del Ministerio de Ciencia e Innovación, y P07 - TIC2970 de la Junta de Andalucía.

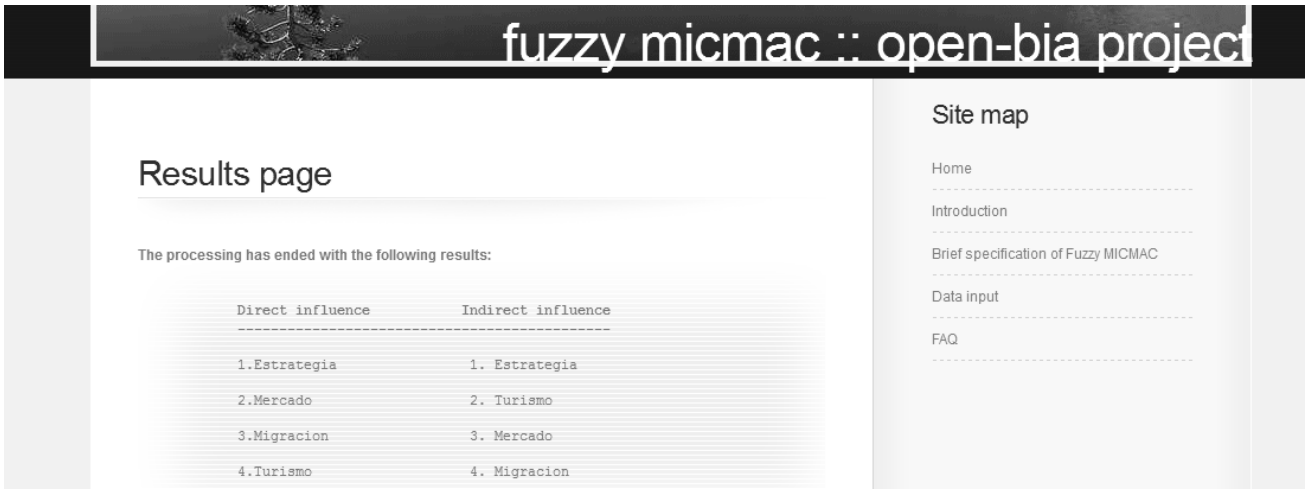
Referencias

[1] R. Bettencourt. Strategic prospective for the implementation of employment policies in the Azores. *Technological Forecasting and Social Change*, 77(9):1566 – 1574, 2010.

[2] D.Castellanos, A. Masegosa, P. Villacorta, P. Novoa, and D. Pelta. Improving scenario method for technology foresight by soft computing techniques. In *Proc. of the 4th International Seville Conference on Future-Oriented Technology Analysis (FTA)*, 2011.



(*a*)



(*b*)

Figura 3: Captura de pantalla de la interfaz web desarrollada para el método MICMAC difuso: (a) página de entrada de datos, (b) página de visualización de resultados

[3] J. C. Duperrin and M. Godet. Methode de hierarchisation des elements d un sisteme. *Rapport Economique du CEA, R.*, pages 45–41, 1973.

[4] P. Durance and M. Godet. Scenario building: uses and abuses. *Technological Forecasting and Social Science*, 77(9):1488 – 1492, 2010.

[5] M. Godet. *Scenarios and Strategic Management*. Butterworth-Heinemann, 1987.

[6] M. Godet. How to be rigorous with scenario planning. *Prospective*, 2(1):5–9, 2000.

[7] M. Godet. *Creating Futures: Scenario Planning as Strategic Management Tool*. Economica, 2006.

[8] F. Herrera, S. Alonso, F. Chiclana, and E. Herrera-Viedma. Computing with words in decision making: foundations, trends and prospects. *Fuzzy Optimization and Decision Making*, 8(4):337–364, 2009.

[9] M. Lamata, D. Pelta, and J. Verdegay. Empleo de técnicas de soft computing en prospectiva tecnológica. In *Actas del III Simposio sobre Lógica Fuzzy y Soft Computing (LFSC2010)*, pages 335–341, 2010.

[10] Y.-C. Lee, Y. H. Chao, and S.-B. Lin. Structural approach to design user interface. *Computers in Industry*, 61(7):613 – 623, 2010.

[11] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning–i. *Information Sciences*, 8(3):199 – 249, 1975.

LÓGICA DIFUSA EN SISTEMAS DE FUSIÓN DE INFORMACIÓN VISUAL: APLICACIONES, EXTENSIONES Y PROPUESTAS

Juan Gómez-Romero, Jesús García, Miguel A. Patricio, Miguel A. Serrano, José M. Molina

Grupo de Inteligencia Artificial Aplicada (GIAA), Departamento de Informática, Universidad Carlos III de Madrid
Avd. de la Universidad Carlos III, 22 28270 *Colmenarejo, España*
{jgromero, mserrano, jgherrer, mpatrici}@inf.uc3m.es, molina@ia.uc3m.es

Resumen

La fusión de información tiene como objetivo combinar datos sensoriales y contextuales para obtener una interpretación global de las situaciones, evaluar el riesgo o impacto de las mismas y proporcionar soporte para la toma de decisiones. En trabajos previos hemos presentado un sistema basado en ontologías que facilita la implementación de procedimientos de fusión de información visual de alto y bajo nivel. Este artículo describe varias aproximaciones para el tratamiento de información imprecisa en este tipo de sistemas y plantea varias líneas de trabajo futuro en la confluencia de la Lógica Difusa, la fusión de información y la toma de decisiones.

Palabras Clave: Fusión de datos e información, Tracking difuso, Ontologías difusas, RCC difuso, Videovigilancia.

1 INTRODUCCIÓN

La investigación en fusión de datos e información (*Data and Information Fusion*, DIF) tiene como objetivo desarrollar teorías y métodos que permitan combinar datos procedentes de varios sensores e información relacionada, mejorando de esta manera los resultados que podrían obtenerse utilizando un único sensor [14]. El modelo de referencia JDL (*Joint Directors of Laboratories*) clasifica los procedimientos de fusión de datos e información en cinco niveles, que se corresponden a distintas fases del procedimiento de transformación de las señales de entrada en conocimiento útil para la toma de decisiones [13, 23]: (i) extracción de características de la señal (nivel 0, o *L0*); (ii) identificación de entidades (*L1*); (iii) reconocimiento de situaciones (*L2*); (iv) evaluación de escenarios (*L3*); (v) mejora del proceso (*L4*). Existen numerosas aplicaciones de fusión de datos e información en diferentes ámbitos; en este trabajo nos centraremos en los sistemas basados en redes de sensores visuales, que procesan los datos obtenidos por conjunto de cámaras de vídeo con el objetivo de identificar posibles amenazas y proporcionar soporte a los usuarios que tienen que organizar una respuesta ante tales eventos.

En las últimas décadas, la atención de los investigadores se ha centrado en los procesos de fusión catalogados en los niveles L0 y L1, denominados genéricamente *fusión de bajo nivel*. Esto ha dado lugar a un gran número de teorías, algoritmos y herramientas tanto para adquirir y procesar los datos obtenidos de sensores distribuidos como para estimar las propiedades de los objetos de la escena (funciones de densidad, vectores de movimiento, etc.). Los procesos de *fusión de alto nivel* (niveles L2 y L3), por su parte, tienen como objetivo descubrir y describir las relaciones cualitativas existentes entre los objetos a partir de las medidas numéricas calculadas en L0 y L1. Estas relaciones, por lo general, denotan asociaciones abstractas, como por ejemplo acciones, actividades, intenciones o amenazas, y se representan de manera simbólica. El objetivo final de la fusión de alto nivel es obtener conocimiento relevante sobre la situación actual (*comprender*), calcular el impacto asociado a la misma (*evaluar*) y proporcionar todo este conocimiento al usuario para ayudar en el proceso de decisión (*soporte*).

La fusión de bajo nivel, y en particular las aplicaciones que utilizan cámaras como sensores de entrada, deben enfrentarse a numerosos problemas derivados de la complejidad de los escenarios y de las limitaciones propias de los sensores. Una de las tareas más comunes es el *tracking* (o seguimiento), que trata de identificar, segmentar y etiquetar de forma consistente los objetos móviles de la escena (*tracks* o pistas) y estimar sus propiedades cinéticas (posición, velocidad, trayectoria, etc.). Las técnicas habituales para el cálculo de las pistas están basadas en métodos de predicción e inferencia estadística, como el filtro de Kalman o los sistemas de partículas. Sin embargo, estos métodos numéricos resultan insuficientes cuando los objetos presentan interacciones (oclusiones, uniones, separaciones) o modificaciones (deformaciones), o cuando se producen cambios en el escenario (iluminación). Estas circunstancias dan lugar a errores de discontinuidad, identificación o estimación de propiedades de las pistas [25]. Asimismo, existe un salto conceptual significativo entre las características de los objetos obtenidas

por estos métodos y las representaciones dispersas (*sparse representations*) que se utilizan en otras tareas más complejas, como compresión de imágenes o reconocimiento de eventos. En consecuencia, existen diferentes aproximaciones que tratan de describir los elementos de la escena con representaciones más expresivas, como por ejemplo las etiquetas espacio-temporales [17].

En cuanto a la fusión de alto nivel, es necesario que los sistemas desarrollen capacidades cognitivas superiores de procesamiento e interpretación de información, especialmente si se tiene en cuenta la complejidad de los escenarios, la imposibilidad de predecir las situaciones posibles y la necesidad de incorporar información de contexto.

Por estos motivos, se han propuesto diferentes formalismos simbólicos para representar y razonar con información abstracta en sistemas de fusión. Así, las aproximaciones cognitivas estudian la creación de modelos cuantitativos expresados en lenguajes basados en Lógica capaces de representar los objetos y las relaciones de un dominio de aplicación. Dichos modelos deben ser capaces de manejar información vaga o incierta, que aparecerá en numerosas aplicaciones debido a diferentes causas: las propia imprecisión de los sensores, la incorporación de conocimiento impreciso proporcionado por usuarios humanos, la agregación de resultados o criterios, etc. Las técnicas basadas en Lógica Difusa pueden ser explotadas para mejorar el rendimiento y aumentar la capacidad de los sistemas de fusión de bajo y alto nivel.

En este trabajo describimos una arquitectura y una herramienta general para la fusión de información procedente de redes de sensores visuales e investigamos varias extensiones basadas en Lógica Difusa en (i) el bajo nivel, para la optimización del seguimiento, y (ii) el alto nivel, para la representación y razonamiento con información imprecisa. Asimismo, se proponen varias líneas de trabajo futuro encaminadas a mejorar el funcionamiento del sistema mediante mecanismos de representación e inferencia difusa.

El artículo está organizado en seis secciones. Después de esta introducción, en la sección 2, describimos la arquitectura general del sistema de fusión de información visual. En la sección 3, se explica el diseño y la implementación de una técnica difusa que mejora notablemente la asociación de regiones (*blobs*) y pistas en el cálculo del seguimiento. En la sección 4, se estudian varias aproximaciones para la gestión de conocimiento impreciso en la representación de escenas y, en particular, de las relaciones espaciales. En la sección 5 se presentan varias líneas de trabajo futuro que amplían las propuestas de las secciones anteriores. Finalmente, en la sección 6 se resumen las conclusiones a las que hemos llegado con este trabajo.

2 ARQUITECTURA PARA FUSIÓN DE INFORMACIÓN VISUAL

En trabajos anteriores, hemos presentado un sistema general para la interpretación de escenas a partir de datos obtenidos mediante una red de sensores visuales [8]. Este sistema utiliza un conjunto de ontologías interrelacionadas para representar la información sensorial y contextual de la escena. Este modelo simbólico proporciona soporte para dos tareas de razonamiento: (i) obtención de una interpretación de alto nivel de la situación actual; (ii) generación de retroalimentación al procedimiento de seguimiento de bajo nivel para mejorar su precisión.

La arquitectura básica del sistema se muestra en la figura 1. Esta arquitectura corresponde al esquema de explotación de información de contexto descrito en [7], en el que se propone la implementación de una capa de procesamiento de alto nivel (capa de contexto) que utiliza información de contexto para comprobar la consistencia de las hipótesis actuales, generar nuevas hipótesis, extender la información disponible con recursos externos y ajustar los procedimientos de fusión de bajo nivel.

La entrada de la capa de contexto son los resultados del seguimiento, que se describe en la sección 3. Los datos de seguimiento son representados de forma simbólica en la sub-ontología de más bajo nivel de abstracción, y a partir de los mismos se inicia el razonamiento para obtener la interpretación de alto nivel de la escena. Este proceso consta de varios pasos: abstracción (para transformar los datos numéricos en instancias de la ontología), correspondencia (identificación de objetos a partir de mediciones físicas), reconocimiento (detección de relaciones entre los objetos) y evaluación (valoración del riesgo).

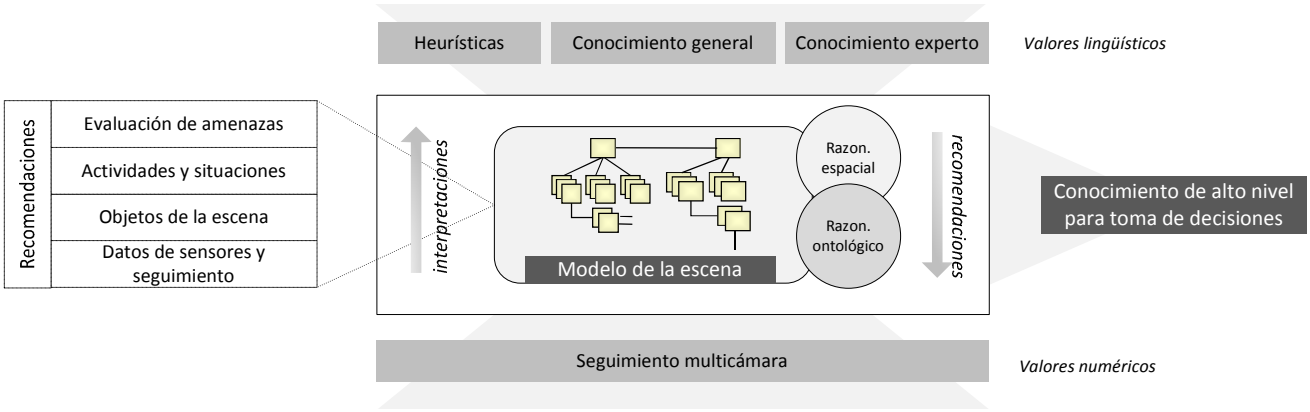


Figura 1. Arquitectura del sistema contextual para fusión de información visual

Como se indicó en la introducción, el nivel L4 del modelo JDL engloba los procesos de adaptación de la captura de información sensorial y su procesamiento para optimizar el resultado de la fusión. La figura 1 muestra cómo nuestro sistema contempla esta posibilidad, implementando un procedimiento de retroalimentación que lo dota de capacidad de adaptación a las situaciones: una vez interpretada la situación actual, el sistema genera recomendaciones que son enviadas al módulo de seguimiento de bajo nivel, modificando su comportamiento de acuerdo al escenario. Por ejemplo, el sistema es capaz de prever situaciones de oclusión entre objetos a partir de la información de contexto almacenada; cuando esta situación es inminente, la capa de contexto genera una recomendación para que el proceso de seguimiento mantenga el tamaño de las pistas, incluso aunque en las imágenes se produzca una reducción parcial o total del mismo.

Un componente fundamental de la arquitectura es el módulo de razonamiento espacial, que se ocupa de la detección, representación y actualización de las relaciones topológicas cualitativas del modelo. Este módulo utiliza, a su vez, una estructura de datos auxiliar para optimizar los cálculos con las geometrías de las entidades. La representación y el razonamiento con la información del modelo, así como las extensiones difusas, se describen en la sección 4.

Este sistema ha sido aplicado en problemas de videovigilancia [6] e inteligencia ambiental [9]. La adaptación del sistema a diferentes dominios de aplicación requiere, en primer lugar, extender las ontologías generales predefinidas para representar los objetos y situaciones concretas de cada problema. A partir del vocabulario definido en estas ontologías, pueden crearse reglas y otros mecanismos de razonamiento para inferir situaciones a partir de datos de seguimiento. La salida del sistema, en ambos casos, es una interpretación de la escena que facilita a los usuarios el análisis de la situación y la toma de decisiones.

3 PROCESAMIENTO DE SEÑALES DE BAJO NIVEL: *TRACKING* DIFUSO

El proceso de seguimiento tiene como entrada las imágenes de la secuencia de vídeo y obtiene como salida la estimación de la posición y del tamaño de las pistas. Este procedimiento consta de una serie de etapas en las que se realizan sucesivamente las siguientes tareas [1]: (1) detectar entidades en movimiento en la secuencia de video; (2) asociar *blobs* y *tracks*[1]; (3) creación y borrado de pistas; (4) generación y análisis de trayectorias.

1 Un *blob*/*región* se define como un conjunto de píxeles que componen una región conexa de la imagen. Un *track/pista* se define como un conjunto de *blobs* que se corresponden con un objeto de la escena con propiedades cinéticas y visuales: tamaño, posición, velocidad, color, etc. Generalmente, los *blobs* y los *tracks* se representan utilizando los rectángulos mínimos que los delimitan.

Uno de los principales problemas del seguimiento es la asociación de regiones y pistas: una vez que las regiones han sido identificadas en el paso (1), es necesario determinar a qué pista pertenece cada uno de ellas, o bien si es necesario crear una nueva pista. Un algoritmo de seguimiento minimiza los errores de asignación, que pueden aparecer cuando las regiones se confunden con el fondo de la imagen o se encuentran demasiado juntas.

Para evitarlo, en nuestro sistema hemos implementado un procedimiento de asociación basado en lógica difusa denominado FRA (Fuzzy Region Assignment) [5, 16]. El método utilizado se basa en la creación de un sistema de reglas difusas de tipo Mamdani que representan diferentes heurísticas para resolver la asignación en situaciones conflictivas. Las heurísticas se basan en medidas de solapamiento entre regiones, deformación, densidad de píxeles y conflictos. El sistema de reglas calcula como resultado, para cada par región-pista, un valor agregado de confianza que mide la certeza con la que la región puedo asociarse a la pista. Este valor es utilizado en la siguiente fase del seguimiento para componer las regiones en movimiento de la imagen y actualizar convenientemente las pistas. Los resultados obtenidos con el proceso de seguimiento difuso, tanto en experimentos previos como durante su utilización en el sistema de fusión contextual, demuestran que, aunque el esfuerzo de crear las reglas puede ser notable, los resultados obtenidos son sustancialmente mejores que si no se utiliza esta extensión.

4 PROCESAMIENTO DE INFORMACIÓN DE ALTO NIVEL: REPRESENTACIÓN Y RAZONAMIENTO DIFUSO

4.1. REPRESENTACIÓN DE ESCENAS CON IMPRECISIÓN E INCERTIDUMBRE

Como señalábamos en la sección 2, el modelo utilizado para la representación del conocimiento sobre los objetos y las acciones de la escena está formado por un conjunto de ontologías organizadas en varias capas de abstracción de acuerdo al modelo JDL, que abarcan desde la codificación básica de datos obtenidos en el seguimiento hasta la representación abstracta de situaciones y riesgos. Las ontologías permiten representar las propiedades de las entidades del dominio y razonar automáticamente con ellas, por ejemplo para llevar a cabo tareas de comprobación de consistencia de la base de conocimiento o clasificación de conceptos.

El uso de ontologías en este modelo ofrece varias ventajas: (i) la representación abstracta de la información hace que el sistema sea más fácilmente interpretable y, por lo tanto, resulte más sencillo para los usuarios interactuar con él; (ii) las ontologías facilitan la reutilización, extensión y compartición de los modelos, lo que reduce el

esfuerzo requerido para la adaptación del sistema en diferentes dominios; (iii) las ontologías se basan en la teoría de Lógica de Descripciones, que ofrece un sustrato formal en el que se basan los procesos de razonamiento; (iv) existen lenguajes estándares para ontologías (principalmente, el lenguaje de ontologías web OWL [10]) y herramientas software que permiten trabajar con ellas (por ejemplo, razonadores e interfaces de programación).

No obstante, las ontologías estándar se caracterizan por no permitir la representación y el razonamiento con conocimiento impreciso o incierto. Recientemente, se han propuesto varias extensiones a los formalismos clásicos de ontologías para la incorporación de estos tipos de conocimiento, entre las que destacan las ontologías difusas y las ontologías con incertidumbre [15]. Las ontologías difusas extienden la semántica de las ontologías clásicas estableciendo que los conceptos denotan conjuntos difusos y las relaciones denotan relaciones binarias difusas, de forma que un individuo puede pertenecer a un concepto en un grado en [0, 1] y dos individuos pueden estar relacionados en un grado en [0, 1]. De forma similar, los axiomas pueden cumplirse en un grado; por ejemplo, dados dos conceptos difusos, puede definirse un axioma de subsunción difuso, que representa una relación de inclusión entre conjuntos difusos.

Utilizando ontologías difusas, es posible que el sistema pueda representar e interpretar las escenas en términos imprecisos. Una ventaja adicional de la representación y razonamiento basados en conceptos y reglas difusas es que facilita la generación de retroalimentación desde el nivel alto al nivel bajo. Así, puede ampliarse el sistema de razonamiento en los niveles más bajos descrito en la sección 3 (formación de los objetos a partir de la imagen y asociación de datos a objetos) a partir de información extraída en los niveles más altos (relaciones entre objetos, actividades y movimientos predichos, etc.).

Nuestras experiencias preliminares en la extensión del modelo de escenas actual a un modelo difuso nos llevan a afirmar que el principal problema no está tanto en la representación difusa, sino en la necesidad de disponer de herramientas apropiadas para razonar, ya que los algoritmos de inferencia no difusos (principalmente los algoritmos *tableau*) no son válidos. En este sentido, en la literatura pueden encontrarse dos aproximaciones diferentes. La primera consiste en desarrollar e implementar nuevos algoritmos de razonamiento, como por ejemplo fuzzyDL [2]. La segunda propone definir procedimientos de reducción para transformar una ontología difusa en una ontología no difusa que preserve la semántica original, de forma que sea posible utilizar los razonadores existentes [3]. En la sección 5 se comentan algunas direcciones de trabajo futuro para incorporar y explotar convenientemente las funcionalidades que las ontologías difusas pueden aportar al sistema de fusión.

En cuanto al conocimiento incierto, existen dos fuentes principales de incertidumbre en el sistema. Por una parte, los datos adquiridos por los sensores tienen una incertidumbre asociada, debido al funcionamiento imperfecto de los sensores. Por otra parte, la detección y análisis de situaciones conlleva cierta incertidumbre en el propio procedimiento de búsqueda de hipótesis que expliquen los datos de entrada. Si bien los motores de inferencia actuales permiten realizar este razonamiento de carácter abductivo, en general no incluyen soporte para razonamiento con incertidumbre.

El manejo de incertidumbre puede incorporarse al sistema directamente en las propias ontologías o mediante una capa adicional que se ocupe de analizar los objetos y situaciones que no están perfectamente catalogadas en el modelo. En el primer caso, existen varios trabajos que han desarrollado lenguajes de representación y herramientas de razonamiento que añaden esta capacidad a las ontologías clásicas [4, 12]. En el segundo, uno de los paradigmas que se puede aplicar es el modelo de argumentación basado en creencias (*belief-based argumentation system*, BAS) [20]. Este modelo combina una representación simbólica con teoría de creencias, permitiendo realizar razonamiento abductivo no monótono con incertidumbre. Actualmente, estamos llevando a cabo algunos experimentos en esta línea y, como se menciona en la sección 5, esperamos ofrecer resultados próximamente.

4.2. RAZONAMIENTO ESPACIAL DIFUSO

La representación y el razonamiento sobre relaciones topológicas tienen especial importancia en aplicaciones de videovigilancia, ya que la situación espacial de los objetos puede utilizarse para deducir otras relaciones más complejas hasta, finalmente, obtener una interpretación de alto nivel de la escena. Además, la representación de relaciones espaciales mediante ontologías no es inmediata y requiere la habilitación de mecanismos especiales.

RCC (*Region Connection Calculus*) es un formalismo basado en lógica de primer orden para representación y razonamiento con relaciones espaciales cualitativas [18]. Existen diferentes subconjuntos de RCC que incluyen sólo una parte de las relaciones definidas en el modelo general. Entre ellos, uno de los más utilizados es RCC-8, que incluye ocho relaciones: DC (desconectado de), EC (conectado externamente a), PO (superposición parcial), TPP (parte interior tangencial), NTPP (no parte interior tangencial), TPPi (inversa de TPP), NTPPi (inversa de NTPP) y EQUAL (igual). La principal característica de RCC en lo que respecta a las ontologías es que estas carecen de la expresividad necesaria para representar completamente la semántica de las relaciones RCC [11]. No obstante, la mayor parte de los motores de inferencia, como RACER o Pellet, incluyen extensiones que permiten manejar relaciones RCC.

En nuestra arquitectura, hemos desarrollado un módulo especial para la gestión de datos y relaciones espaciales. Este módulo se ocupa de dos tareas: detección y actualización de propiedades espaciales de los objetos de la

escena y representación y razonamiento espacial en el modelo ontológico. El módulo consta de tres componentes: (i) una base de conocimiento integrada con el modelo de la escena; (ii) una implementación abstracta de RCC para interactuar con razonadores que ofrezcan esta funcionalidad; (iii) una estructura de datos adicional en la que se almacenan los datos geométricos de las entidades y que sirve para optimizar los cálculos numéricos.

Uno de los inconvenientes fundamentales de RCC (y del razonamiento espacial en general) es que sólo se contempla relaciones precisas entre los objetos. Esto significa que, aunque las regiones pueden ser imprecisas, las relaciones entre ellas deben cumplirse o no completamente. Esta característica es especialmente problemática cuando se utilizan sensores visuales, en cuyo caso es poco realista asumir que se pueden detectar con total precisión relaciones como, por ejemplo, EC (conectado externamente). Asimismo, tanto en la interpretación de la escena como en la presentación de información a los usuarios, es muy conveniente manejar otras relaciones espaciales imprecisas como 'cerca', 'lejos', 'en medio', etc.

En la literatura pueden encontrarse diferentes propuestas que definen extensiones difusas de RCC. Estas extensiones permiten, por ejemplo, que dos objetos se consideren conectados si son realmente adyacentes o si están muy cerca uno del otro [21]. De forma similar, otros autores han explorado la posibilidad de introducir este tipo de representación RCC difusa en lógicas de descripciones, aunque con varias limitaciones [24]. No obstante, actualmente no existe ninguna herramienta que permita introducir fácilmente estas capacidades en un sistema real, por lo que su uso efectivo en nuestro sistema constituye una línea de trabajo futuro.

5 LÍNEAS DE INVESTIGACIÓN Y TRABAJO FUTURO

En las secciones anteriores hemos mencionado algunas líneas de investigación interesantes para continuar el trabajo presentado en este artículo, con el objetivo de encontrar más sinergias entre las áreas de la lógica difusa y la fusión de información, y en particular, los sistemas de soporte a la decisión basados en señales de vídeo. Hemos mencionado, por ejemplo, la representación difusa de objetos y situaciones, el razonamiento con imprecisión e incertidumbre para interpretación de la escena o las relaciones espaciales imprecisas.

Además de estas, un área de trabajo futuro muy interesante es la aplicación de técnicas de aprendizaje automático para obtención de reglas para interpretación de situaciones. En la implementación actual del sistema, el experto en el dominio de aplicación debe definir las reglas abductivas que permiten obtener conocimiento de nivel de abstracción superior a partir de datos de nivel inferior. Esta tarea requiere un conocimiento notable del dominio de aplicación, así como un esfuerzo considerable en la caracterización de los escenarios que pueden ocurrir. En algunos casos, incluso, no será posible modelar las situaciones relevantes o de riesgo, y solamente será posible alertar al usuario cuando no sea posible etiquetar una acción como segura. Sería conveniente, por lo tanto, estudiar técnicas que permitan al sistema *aprender* las situaciones conflictivas, bien a partir de datos de entrada, bien a partir de las correcciones de los usuarios. Algunos avances aplicados a la fusión de bajo nivel han sido presentados en trabajos anteriores [19].

Asumiendo que las escenas pueden describirse mediante términos difusos, también resulta interesante estudiar cómo pueden agregarse diferentes valoraciones de riesgo de la escena. Esto permitiría ofrecer a los usuarios indicaciones más realistas y más fácilmente comprensibles sobre la escena actual.

6 CONCLUSIONES

En este trabajo hemos descrito un sistema para fusión de información procedente de fuentes de video. El sistema está basado en un modelo ontológico para la representación de los objetos y las relaciones de la escena. El objetivo final del sistema es ofrecer información relevante y contextualizada a usuarios sobre la situación actual y el riesgo o amenaza de dicha situación. El sistema utiliza técnicas de Lógica Difusa para mejorar su funcionamiento en dos componentes: (i) reglas heurísticas difusas para el seguimiento de los objetos móviles de la escena; y (ii), ontologías difusas para la representación imprecisa y razonamiento con escenas, en general, y con relaciones espaciales entre objetos, en particular. Actualmente, dichas extensiones se encuentran en fase de experimentación. Además de estas líneas de investigación, pueden identificarse otras direcciones de trabajo futuro muy interesantes, como el aprendizaje automático de reglas o la agregación de información de carácter lingüístico representada con ontologías difusas para evaluación de la amenazas.

Agradecimientos

Este trabajo de investigación ha sido financiado parcialmente por los proyectos: TIN2008-06742-C02-02/TSI, TEC2008-06732-C02-02/TEC, CONTEXTS (S2009/TIC-1485) y DPS 2008-07029-C02-02.

Referencias

[1] J.A. Besada, J. García, J. Portillo, J.M. Molina: Airport surface surveillance based on video images. *IEEE Transactions on Aerospace and Electronic Systems* 41, pp. 1075–1082, 2005.

[2] F. Bobillo, U. Straccia: fuzzyDL: An expressive Fuzzy Description Logic reasoner. En: *Proceedings of the 2008 International Conference on Fuzzy Systems (FUZZ-IEEE 2008)*, pp. 923–930, Hong Kong, China, 2008.

[3] F. Bobillo, M. Delgado, J. Gómez-Romero, U. Straccia: Fuzzy Description Logics under Gödel semantics. *International Journal of Approximate Reasoning* 50(3), pp. 494–514, 2009.

[4] P.C.G. Costa: *Bayesian Semantics for the Semantic Web*. Doctoral Thesis, George Mason University, VA, USA, 2005.

[5] J. García, M.A. Patricio, A. Berlanga, J.M. Molina: Fuzzy region assignment for visual tracking. *Soft Computing* 15(9), pp. 1–20, 2009.

[6] J. Gómez-Romero, M.A. Patricio, J. García, J.M. Molina: Context-based reasoning using ontologies to adapt visual tracking in surveillance. En: *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2009)*, pp. 226–231, Génova, Italia, 2009.

[7] J. Gómez-Romero, J. García, M. Kandefer, J. Llinas, J.M. Molina, M.A. Patricio, M. Prentice, S.C. Shapiro: Strategies and techniques for use and exploitation of contextual information in high-level fusion architectures. En: *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, Edimburgo (UK), 2010.

[8] J. Gómez-Romero, M.A. Patricio, J. García, J.M. Molina: Ontology-based context representation and reasoning for object tracking and scene interpretation in video. *Expert Systems with Applications* 38(6), pp. 7494–7510, 2011.

[9] J. Gómez-Romero, M.A. Serrano, M.A. Patricio, J. García, J.M. Molina: Context-based scene recognition from visual data in smart homes: an Information Fusion approach. *Personal and Ubiquitous Computing*, In Press.

[10] P. Hitzler et al.: OWL 2 Web Ontology Language primer. En: http://www.w3.org/TR/owl2-primer/ (accedido Oct. 2011).

[11] F. Hogenboom, B. Borgman, F. Frasincar, U. Kaymak: Spatial knowledge representation on the Semantic Web. En: *Proceedings of the IEEE 4th International Conference on Semantic Computing (ICSC 2010)*, pp. 252–259, Pittsburgh, PA, USA, 2010.

[12] P. Klinov: Pronto: A non-monotonic probabilistic Description Logic Reasoner. In: *Proceedings of the 5th European Semantic Web Conference (ESWC 2008)*, pp. 822–826, Tenerife, España, 2008.

[13] J. Llinas, C.L. Bowman, G. Rogova, A.N. Steinberg, E. Waltz, F. White: Revisiting the JDL data fusion model II. En: *Proceedings of the 7th international conference on Information Fusion (Fusion 2004)*, Estocolmo (Suecia), pp. 1218–1230, 2004.

[14] J. Llinas, D.L. Hall: Multisensor data fusion. En: M.E. Liggins, D.L. Hall, J. Llinas (Eds.), *Handbook of multisensor data fusion*. CRC Press, pp. 1–14, 2009.

[15] T. Lukasiewicz, U. Straccia: Managing uncertainty and vagueness in Description Logics for the Semantic Web. *Journal of Web Semantics* 6(4), pp. 291–308, 2008.

[16] J.M. Molina, J. García, J. Carbó, A. Berlanga, J. Portillo: Aplicaciones de la Lógica Borrosa en sistemas de vigilancia utilizando visión activa. En: *Actas del XII Congreso Español de Tecnologías y Lógica Fuzzy (ESTYLF2004)*, pp. 527-532, Jaén, España, 2004.

[17] J.C. Niebles, H. Wang, L. Fei-Fei: Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computer Vision* 79, pp. 299—318, 2008.

[18] J. Renz (Ed.): *Qualitative spatial reasoning with topological information*, Springer Verlag, Berlin, 2003.

[19] B. Rodríguez, O. Pérez, J. García, J.M. Molina: Application of machine learning techniques for acquiring new knowledge to simplify the association problem in image tracking. *Applied Artificial Intelligence*, 22(3), 2008.

[20] G. Rogova, P. Scott, C. Lollett, R. Mudiyanur: Reasoning about situations in the early post-disaster response environment. En: *Proceedings of the 9th International Conference on Information Fusion (Fusion 2006)*, Florencia, Italia, 2006.

[21] S. Schockaert, M. De Cock, E.E. Kerre: Spatial reasoning in a fuzzy Region Connection Calculus. *Artificial Intelligence* 173(2), pp. 258–298, 2009.

[22] M. Serrano, J. Gómez-Romero, M.A. Patricio, J. García, J.M. Molina: Topological properties in ontology-based applications. En: *Proceedings of the 11th International Conference on Intelligent Systems Design and Applications (ISDA 2011)*, Córdoba, España, 2011.

[23] A.N. Steinberg, C.L. Bowman: Rethinking the JDL data fusion levels. En: *Proceedings of the MSS national symposium on sensor and data fusion*, Columbia, SC (USA), 2004.

[24] U. Straccia: Towards spatial reasoning in fuzzy Description Logics. En: *Proceedings of the 2009 International Conference on Fuzzy Systems (FUZZ-IEEE 2009)*, pp. 512–517, Jeju Island, Korea, 2009.

[25] A. Yilmaz, O. Javed, M. Shah: Object tracking: A survey. *ACM Computing Surveys* 38(4), pp. 1–45, 2006.

IMPACTO DE LAS ETIQUETAS EN LA INTERPRETACIÓN DE LA ESCALA DE LIKERT

Nuria Martínez[1], Daniel Gómez[2], Javier Montero [1]

[1] *Facultad de Matemáticas, Universidad Complutense de Madrid, 28040, Madrid*

[2] *Escuela de Estadística, Universidad Complutense de Madrid, 28040, Madrid*

nuriamartinez@gmail.com, dagomez@estad.ucm.es, monty@mat.ucm.es

Resumen

En este trabajo se comparan las distribuciones de las respuestas dadas a un cuestionario formulado en dos versiones, una con etiquetas representando el grado de acuerdo, frente a otra con etiquetas expresando la frecuencia de realización de una acción. Se examina la relación entre la percepción de cada etiqueta lingüística y las etiquetas frecuentistas mediante números borrosos. Se analiza la equivalencia entre ambos métodos, estudiando en detalle la equidistancia en la percepción de las etiquetas extremas. Se evalúan además las diferencias en función de la importancia del tema sobre el que se interroga.

Palabras Clave: Lógica difusa, escala de Likert, frecuencia, etiqueta, proceso de decisión de compra, acciones sostenibles

1 INTRODUCCIÓN

En la investigación social es frecuentemente necesaria la medición de actitudes, opiniones o intensidades de un concepto. Para ello, es habitual el uso de diferentes escalas que generan niveles ordinales del atributo medido.

La escala de Likert es muy utilizada por su rápida construcción, aplicación e interpretación y por la suposición de que es posible aumentar la precisión de la respuesta incluyendo más categorías y suponiendo la escala continua [17]. Esta escala consiste en que para cada ítem el encuestado asigna una puntuación de acuerdo a su actitud respecto al ítem propuesto. La puntuación se selecciona de un rango de valores cerrado, a menudo de 1 a 5, aunque pueden utilizarse rangos más amplios (1 a 7, 1 a 10…etc).

La escala de Likert presenta algunos problemas [11] como la supuesta unidimensionalidad de las actitudes, la localización en un espacio continuo y la suposición de la equidistancia entre las alternativas expuestas.

Son varios los autores que han estudiado en los últimos años la influencia de la escala utilizada, del número de opciones y de la formulación de las respuestas incluidas en los cuestionarios [13], [17], planteando cuestiones abiertas como si el contenido del cuestionario tendrá influencia sobre las respuestas extremas [17].

En los cuestionarios de marketing es frecuente disponer de datos de muy diferente naturaleza, encontrando en el mismo cuestionario preguntas de tipo difuso, así como preguntas de carácter más frecuentista. Un ejemplo de pregunta de naturaleza difusa sería: *Expresar el grado de acuerdo con la siguiente frase: "Compro productos con ofertas especiales"*, mientras que una formulación frecuentista alternativa sería: *"¿Con qué frecuencia compra productos con ofertas especiales?"*.

Este trabajo plantea la cuestión de si las respuestas y conclusiones finales de un estudio varían dependiendo de la formulación de las opciones de respuesta. Para ello, se realizan dos versiones de un mismo cuestionario, formulando las preguntas de modo que la primera versión admita respuestas sobre el grado de acuerdo y la segunda versión sobre la frecuencia de realización.

La utilización de la lógica borrosa es amplia en muchas disciplinas, y en particular en los campos sociales. Concretamente se utiliza a menudo en estudios de evaluación o satisfacción, existiendo diversas investigaciones relacionadas con experimentos en línea con lo aquí presentado. Existen artículos describiendo experimentos tales como asignar de forma libre etiquetas lingüísticas y valores numéricos a las mismas para estudiar su relación [13],[16], presentar diferentes escalas lingüísticas y completar con la nota numérica para analizar la equivalencia y distancia entre ellas [11] u otros en los que se ofrecen escalas numéricas y verbales y se pide que se relacionen ambas [9].

Vemos por tanto que son muchos los autores que han abordado el problema sobre la dimensión numérica de variables lingüísticas. Nuestro objetivo es analizar la equivalencia, dentro de una encuesta construida sobre una escala subyacente de Likert, de una formulación de preguntas expresando el grado de acuerdo con una frase que

conlleve una acción y la formulación de la misma pregunta indicando la frecuencia con la que se realiza esta acción. Encontramos, por lo expuesto anteriormente, que es de gran interés profundizar en el tipo de relación existente entre etiquetas lingüísticas y etiquetas frecuentistas que a menudo parecen considerarse equivalentes, dada la mezcolanza que existe en la práctica entre ambas formulaciones.

En la sección 2 de este documento se describen el método y los materiales de esta investigación. En la sección 3 se presentan los resultados. En concreto, se recogerá la incertidumbre generada en torno a cada etiqueta lingüística mediante números borrosos. A continuación se analizará si ambos métodos son equivalentes o si existe algún tipo de tendencia o patrón en las respuestas. Se abordará también la cuestión encontrada en la literatura consultada [17] sobre si el tema interrogado puede ser un factor influyente en las diferencias de los resultados. Para ello, el cuestionario incluirá dos temas de distinta relevancia en la vida del encuestado. Finalmente, teniendo en cuenta los problemas mencionados sobre equidistancia, se estudiará un comportamiento encontrado en los gráficos por el que parece que la dispersión de las respuestas en torno al valor extremo "Muy de acuerdo" es mayor que la dada en torno al valor opuesto "Muy en desacuerdo". Se querrá detectar si estos conceptos lingüísticos corresponden a un rango de frecuencias diferente, no siendo por tanto equidistantes (ver Figura 1). Finalmente, en la sección 4 se exponen posibles líneas futuras de trabajo.

2 MATERIALES Y MÉTODO

El instrumento utilizado en este estudio es un cuestionario que interroga sobre dos temas. El primero consta de 13 preguntas (C1, C2… C13) relacionadas con las acciones llevadas a cabo durante el proceso de decisión de compra de un champú y el segundo se compone de 14 preguntas (S1, S2…S14) sobre la consideración de la sostenibilidad medioambiental en las acciones diarias del entrevistado.

El cuestionario se formula en dos versiones, atendiendo a la naturaleza de las respuestas:

- Cuestionario difuso: las respuestas a cada pregunta son definidas mediante etiquetas lingüísticas

D1. Muy de acuerdo
D2. Bastante de acuerdo
D3. Ni de acuerdo, ni en desacuerdo
D4. Bastante en desacuerdo
D5. Muy en desacuerdo

- Cuestionario frecuentista: las respuestas a cada pregunta son definidas mediante frecuencia de realización

F1. Alrededor del 100% de las veces
F2. Alrededor del 75% de las veces
F3. Alrededor del 50% de las veces
F4. Alrededor del 25% de las veces
F5. Alrededor del 0% de las veces

Previo al lanzamiento del campo, se llevó a cabo una fase de testeo de ambas versiones del cuestionario con el objeto de asegurar la correcta expresión y comprensión de cada pregunta, así como la equivalencia semántica de ambos cuestionarios. Con el análisis de los resultados de esta fase de test y los comentarios recibidos se redactaron los cuestionarios finales.

La primera parte del cuestionario se construye en torno a los factores de decisión de compra de un champú, conteniendo preguntas sobre las diferentes etapas del proceso habitual de compra de un producto (ver [7], [8], [14]). La segunda parte trata diferentes acciones sostenibles de los consumidores descritas en manuales de organizaciones públicas (ver [2], [10]) susceptibles de ser formuladas con etiquetas lingüísticas y con etiquetas frecuentistas. Se incluye una pregunta de filtro para asegurar que los encuestados han comprado champú en los últimos doce meses y tienen por tanto criterio para responder a la primera parte del cuestionario.

Se define un público objetivo que permita obtener una muestra suficiente para el análisis por lo que se decide llevar a cabo el estudio en el ámbito de la Universidad. La captación se realiza mediante contactos en colegios mayores, anuncios en diferentes facultades de la Universidad Complutense de Madrid y la publicación del estudio en la página de Facebook "Doctorado Matemáticas"

Se recoge una muestra que, tras eliminar las encuestas incompletas y depurar los datos, consta de 95 estudiantes universitarios que completaron ambas fases del estudio. Esta muestra se compone de un 58% de estudiantes con perfil de ciencias y un 42% con perfil de letras, habiendo un 17% de hombres y 83% de mujeres. Este último desbalanceo es probablemente debido a la pregunta de filtro sobre la compra de champú en los últimos doce meses.

La versión difusa del cuestionario se llevó a cabo mediante encuestas auto-administradas online desarrolladas con el software Surveymonkey durante el mes de Febrero de 2011. Dos semanas después se lanzó el cuestionario frecuentista, prolongando el campo hasta finales de Abril.

El análisis de este estudio requiere identificar las respuestas de un mismo entrevistado. Para ello, a cada individuo se le pidió que incluyese las iniciales de su nombre y primer apellido así como su fecha de nacimiento en ambos cuestionarios de forma que se pudiese generar un identificador único manteniendo el anonimato.

Para poder alcanzar a los mismos individuos en el segundo cuestionario, se pidió en la primera fase que incluyesen su dirección de correo electrónico. Se utilizó además la mencionada página de Facebook "Doctorado Matemáticas" para que aquellas personas que fuesen reacias a proporcionar este dato, pudiesen hacerse "fan" de la misma y ser alertadas de forma anónima en su perfil de Facebook cuando se lanzase el segundo cuestionario.

El análisis estadístico de los datos se llevó a cabo con Excel y con los paquetes estadísticos STATA y R.

3 RESULTADOS Y CONCLUSIONES

Se tienen por tanto 27 preguntas contestadas por 95 individuos aplicando dos métodos distintos (método difuso y método frecuentista).

Para cada pregunta se muestra a continuación un gráfico de burbujas con los datos de la tabla de contingencia asociada, teniendo en el eje horizontal las categorías difusas y en el vertical las frecuentistas:

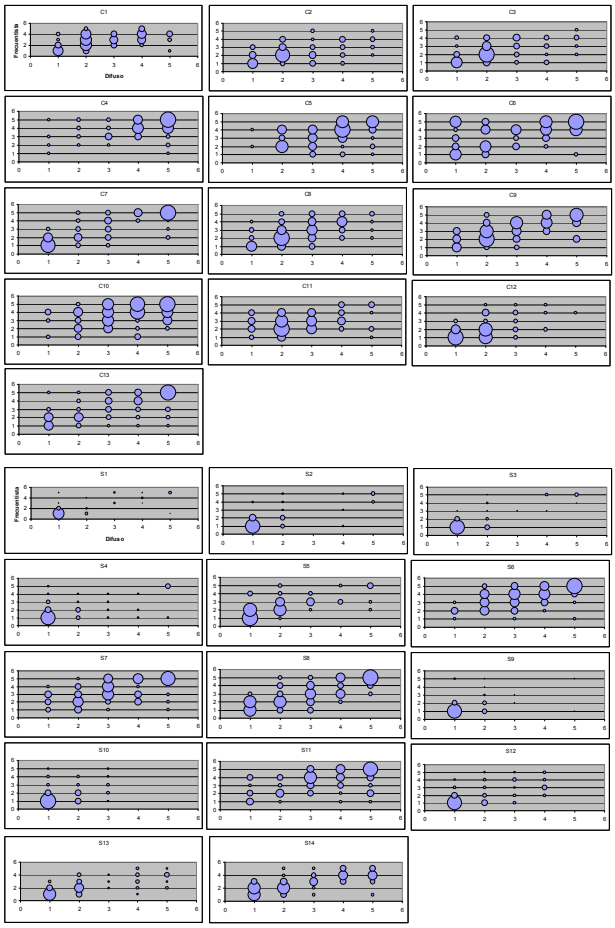


Figura 1: Gráfico de respuestas Difuso vs. Frecuentista, para cada pregunta

Se observa, como era de esperar, que en general las diagonales tienen el mayor número de casos, teniendo el 100% de ellos si ambos métodos fueran equivalentes. Nótese además que parece existir una mayor equivalencia en el segundo bloque de preguntas sobre sostenibilidad que en el primero. La figura anterior muestra también el comportamiento comentado previamente por el que a respuestas difusas "Muy de acuerdo" parece corresponderles una distribución frecuentista más amplia que la equivalente en respuestas "Muy en desacuerdo".

Para analizar las posibles diferencias en la percepción frecuentista de las etiquetas lingüísticas que definen el grado de acuerdo (D1, D2…D5, ver 2 MATERIALES Y MÉTODO), se realizó el cálculo de números borrosos triangulares que tratan de establecer esta relación. Un número borroso triangular es un subconjunto borroso A de X cuya función de pertenencia μ_A: U→[0,1] viene definida por:

$$\mu_A(u)= \begin{cases} \dfrac{(u-a)}{b-a} & \text{si } a \le u \le b \\ \dfrac{(u-c)}{b-c} & \text{si } b \le u \le c \\ 0 & \text{en otro caso} \end{cases}$$

El número borroso A puede ser por tanto representado mediante la tripleta A = (a, b, c) con a ≤ b ≤ c. En este estudio se ha tomado b = moda (X), a =b-2*s y c= b + 2*s, donde s es la desviación estándar.

Los resultados (ver Figura 2) muestran que las funciones de pertenencia son muy variadas, habiendo un 35% de ellos en los que las modas no se corresponden con la frecuencia esperada de la etiqueta asociada. El 22% muestra un desvío de la distribución frecuentista hacia la derecha del valor que le correspondería. Este comportamiento se encontrará también en el gráfico boxplot de diferencias entre las respuestas (ver Figura 3) y corrobora un patrón que se mencionará posteriormente sobre la equivalencia de los métodos. La percepción de la etiqueta "Ni de acuerdo, ni en desacuerdo" arroja, como era de esperar, el mayor porcentaje de casos de no coincidencia entre la moda de la distribución y la frecuencia "Alrededor del 50% de las veces", es decir, es donde existe una mayor incertidumbre en la equivalencia entre la etiqueta lingüística y la frecuencia de realización. Además, existen en general desviaciones en torno a la moda que varían hasta alcanzar gran amplitud, siendo ésta mayor en las preguntas relacionadas con el proceso de compra de un champú que en el tema de sostenibilidad.

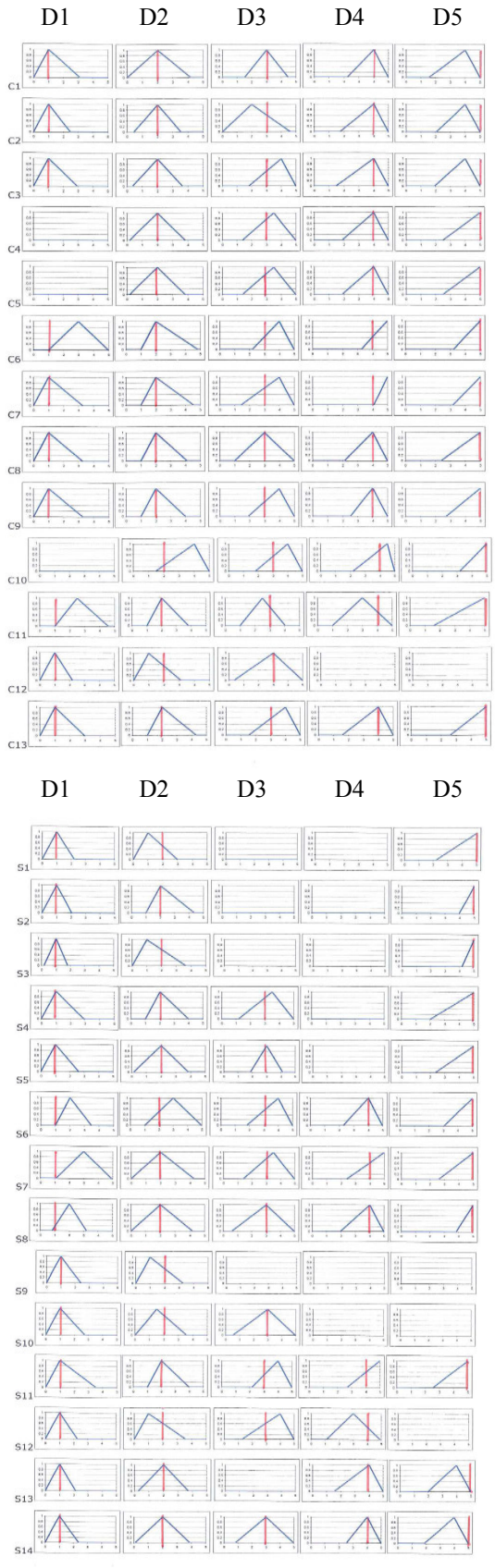


Figura 2: Números borrosos triangulares por pregunta para cada etiqueta lingüística

En cuanto a la equivalencia de los métodos, los análisis llevados a cabo no permiten concluir de forma general que exista evidencia para afirmar que los resultados con uno y otro método son significativamente distintos. La extensión de Stuart y Maxwell al test de McNemar para estudiar la homogeneidad marginal [4] permite detectar algunos ítems, 7 de 27, en los que existe evidencia para rechazar que las muestras marginales sean homogéneas:

CHAMPÚ		SOSTENIBILIDAD	
C1	**0,0084**	S1	0,5747
C2	0,6824	S2	0,0765
C3	**0,0376**	S3	0,1556
C4	0,5003	S4	0,4344
C5	0,6297	**S5**	**0,00001**
C6	**0,0169**	**S6**	**0,0388**
C7	0,0673	S7	0,0551
C8	0,3446	S8	0,4352
C9	0,1924	S9	0,6518
C10	0,2345	**S10**	**0,0469**
C11	0,4463	S11	0,1039
C12	0,0698	S12	0,3904
C13	0,4411	**S13**	**0,0410**
		S14	0,0902

Tabla 1: P-valores del contraste de homogeneidad marginal mostrando en rojo aquéllos inferiores a 0,05

Algunas de estas preguntas se muestran también relevantes en otros tests y análisis realizados coincidiendo todas ellas en que tienen una formulación lingüística más compleja que el resto.

Para cada pregunta se estudia el grado de asociación mediante la medida γ de Goodman y Kruskall ([1],[4]) así como el grado de acuerdo mediante el índice Kappa ([1],[3],[4]). En general se observa que existe asociación elevada en las respuestas dadas por cada uno de los métodos. El acuerdo encontrado no es excesivamente alto en muchas de las preguntas, pero, dadas las limitaciones y la sensibilidad del índice Kappa a la distribución de las marginales (ver, por ejemplo, [3], [15], [18]), y el uso meramente descriptivo hecho del índice, no se puede extraer ninguna conclusión directa.

Se aprecia que las preguntas dedicadas a sostenibilidad muestran mayor grado de asociación y acuerdo que las relacionadas con el proceso de compra de un champú.

El análisis del las diferencias entre las respuestas mediante el gráfico boxplot (ver Figura 3) muestra un patrón interesante: salvo valores atípicos, en 21 de las 27 preguntas sucede para al menos la mitad de los individuos que, o bien no hay diferencias (d_i-f_i = 0), o bien existe una tendencia por la que las personas declaran una frecuencia de realización de la acción inferior a la etiqueta lingüística seleccionada en el cuestionario difuso (d_i-f_i = -1 → d_i = f_i -1).

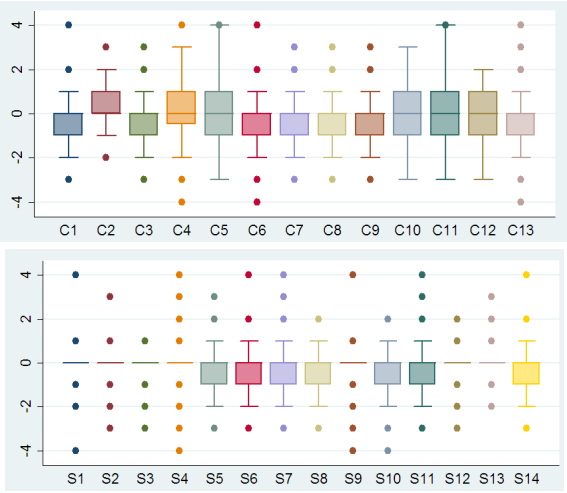


Figura 3: Boxplot de las diferencias de las respuestas dadas por cada individuo

Por ejemplo, una persona podría expresar que está "Bastante de acuerdo" con una pregunta en el cuestionario difuso, y declarar en el cuestionario frecuentista que lo lleva a cabo alrededor de la mitad de las veces (cuando la etiqueta equivalente en frecuentista sería "Alrededor del 75% de las veces"). Es decir, cuando no hay equivalencia directa entre los métodos, parece existir una tendencia a declarar un mayor acuerdo en la pregunta con etiquetas lingüísticas que en la pregunta formulada con etiquetas frecuentistas:

	di = fi - 1		
	di	fi	
Muy de acuerdo	1	2	Alrededor del 75% de las veces
Bastante de acuerdo	2	3	Alrededor del 50% de las veces
Ni de acuerdo, ni en desacuerdo	3	4	Alrededor del 25% de las veces
Bastante en desacuerdo	4	5	Alrededor del 0% de las veces
Muy en desacuerdo	5	-	

Tabla 2: Relación entre las respuestas de una pregunta para el individuo i-ésimo en el caso analizado

A la vista de los resultados obtenidos y los análisis realizados, entre otros el gráfico boxplot de diferencias y los niveles de asociación y acuerdo, podría concluirse que el tema sobre el que se interroga sí es un factor determinante en la posible equivalencia entre las distribuciones de las respuestas difusas o frecuentistas. Así, podría pensarse que temas con cierta relevancia social o simplemente "políticamente correctos", como es la realización de acciones sostenibles en el día a día, arrojan resultados similares independientemente de las etiquetas con las que se pregunte. En cambio, temas menos importantes en la vida del entrevistado, como son las acciones previas a la decisión de compra de un champú, sí parecen verse más influidos por las opciones de respuesta del cuestionario.

Para analizar la equidistancia de las respuestas frecuentistas en torno a las etiquetas lingüísticas "Muy de acuerdo" (D1) y "Muy en desacuerdo" (D5) se calcula el índice de entropía de la distribución frecuentista condicionada a los valores D1 ó D5 ([5],[12]). Se define la entropía de la v.a. X condicionada al valor $Y = y_j$ como:

$$H(X \mid Y=y_j) = -\sum_{i=1}^{I} p(x_i|y_j)\log_2 p(x_i|y_j)$$

Los resultados (ver Tabla 3) arrojan que existen varias preguntas para las que se confirma la hipótesis de una mayor entropía condicionada a D1 que la resultante condicionando a D5, no siendo éstas suficientes para concluir un resultado de forma general.

CHAMPU			SOSTENIBILIDAD		
Pregunta	Entropía condicionada a… D1	D5	Pregunta	Entropía condicionada a… D1	D5
C1	1,59	1,85	**S1**	**0,76**	**0,54**
C2	1,35	1,84	S2	0,84	0,95
C3	1,38	1,79	**S3**	**0,67**	**0,65**
			S4	**1,27**	**0,59**
			S5	**1,28**	**1,25**
C6	**2,07**	**1,19**	**S6**	**1,25**	**1,12**
C7	**1,20**	**0,90**	**S7**	**1,85**	**0,91**
C8	1,55	1,92	**S8**	**1,23**	**0,94**
C9	**1,55**	**1,38**			
C11	**1,95**	**1,84**	**S11**	**1,87**	**1,68**
C13	**1,63**	**1,12**	S13	0,97	1,66
			S14	1,40	1,67

Tabla 3: Resultados de la entropía condicionada para cada pregunta indicando en rojo aquellos casos en los que se cumple la hipótesis planteada

4 TRABAJOS FUTUROS

Para investigaciones futuras consideramos interesante tratar de llevar a cabo un análisis lingüístico de las preguntas para las que se han detectado diferencias significativas y comportamientos distintos en las respuestas. Esto permitiría extraer patrones acerca de cómo están estipuladas estas preguntas concretas y ver qué diferencias existen con respecto al resto.

Encontramos también de interés realizar un nuevo experimento similar teniendo en cuenta varios temas de diferente relevancia en la vida del encuestado, para estudiar si se repite el comportamiento encontrado en este estudio de a mayor relevancia (individual o social), mayor equivalencia entre los dos métodos.

Otra cuestión que no se ha analizado, pero que convendría tener en cuenta en futuras investigaciones, sería determinar la discrepancia natural en un cuestionario. Es decir, conocer el porcentaje de discrepancia que encontraríamos si repitiésemos dos veces un cuestionario idéntico a la misma muestra de individuos. Esto nos permitiría determinar cuánta de la discrepancia encontrada en el estudio se debe a la nueva formulación de las etiquetas de las respuestas y cuánta sería la discrepancia natural.

Un factor a analizar en el futuro es tratar de discernir el impacto de la percepción que tienen los entrevistados acerca del tipo de escala subyacente.

Finalmente, otro tema susceptible de estudio sería recoger si existen diferencias en las respuestas en función del perfil de estudios del entrevistado. Es decir, si universitarios de carreras de ciencias (tecnológicas, bio-sanitarias) o de letras (sociales o de humanidades) responden de manera significativamente distinta en un cuestionario de naturaleza difusa o frecuentista.

Agradecimientos

Este trabajo se ha realizado con financiación parcial del proyecto TIN2009-07901 del Plan Nacional de Investigación.

Referencias

[1] Análisis de datos cualitativos: tablas de contingencia, (2007). Sociedad española de bioquímica clínica y patología molecular (SEQC)

[2] Consumo sostenible, el equilibrio entre la acción humana y el planeta, (2009). *Guía de consumo responsable*. Consultado en Noviembre de 2010. Donostia-San Sebastián: Servicio Central de Publicaciones del Gobierno Vasco: http://www.ejgv.euskadi.net/r53-2291/es/contenidos/informacion/publi_didactico/es_1302/adjuntos/consumo_sostenible.pdf

[3] Abraira V. Errores en las mediciones y clasificaciones clínicas: Precisión y validez (n.d.). Hospital Universitario Ramón y Cajal. http://www.hrc.es/bioest/Intro_errores.html

[4] Aguilera del Pino, A.M. Análisis de tablas de contingencia bidimensionales. Abril 2005, http://www.ugr.es/~focana/dclasif/aaguilera.pdf

[5] Cadena Monroy L.A. (2006). Entropía: Información y Complejidad. *Senderos ambientales*. *VOL. 1 NO. 1, pp. 81 - 88*

[6] Gómez D., Biging G. y Montero J.(2008).Accuracy statistics for judging soft classification. *International Journal of Remote Sensing*, Vol 29, pp 693-709

[7] González Cobo M.A. y Carrero López E. (1997). *Manual de planificación de Medios*. Madrid: ESIC

[8] Howard J.A. (1993). El comportamiento del Consumidor en la estrategia de Marketing. Madrid : Diaz de Santos.

[9] Jaffe-Katz A., Budescu D.V., Wallsten T. (1989). Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty. *Memory&Cognition* Vol. 17 nº3, pp. 249-264

[10] Jóvenes por el cambio. Manual de educación para un consumo sostenible META (Asociación de los Medios, Ecología y Tecnología), División de Tecnología, Industria y Economía del PNUMA (Programa de Naciones Unidas para el Medio Ambiente) y la División para la Enseñanza Secundaria, Técnica y Profesional de la UNESCO. http://www.oei.es/decada/portadas/YXC_red.pdf

[11] Lalla M., Facchinetti G. y Mastroleo G. (2004). Ordinal Scales and Fuzzy Set Systems to Measure Agreement: An Application to the Evaluation of Teaching Activity, *Quality&Quantity 38: 577 – 601*, Kluwer Academic Publishers, Netherlands

[12] Larrañaga P. e Inza I. Teoría de la información estadística. Departamento de Ciencias de la Computación e Inteligencia Artificial. Universidad del País Vasco: http://www.sc.ehu.es/isg/

[13] Martínez García J.A. y Martínez Caro L. (2010). La medición de la satisfacción del consumidor de servicios deportivos a través de la lógica borrosa. *Revista de Psicología del Deporte* Vol. 19, nº 1, pp. 41-58

[14] Solé Moro M.L. (2003). *Los consumidores del siglo XXI*. Madrid: ESIC.

[15] Viera A.J., Garret J.M. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine* Vol.37, nº5, pp. 360 - 363

[16] Wallsten T.S., Budescu D.V. y Zwick R.(1993). Comparing the Calibration and Coherente of Numerical and Verbal Probability Judgments. Management Science Vol. 39, nº2, pp.176 - 190

[17] Watkins M.L. (1992). The implications of extreme response style (ERS) for cross-cultural anc domparative research in South Africa. *Journal of Industrial Psychology*, 18(1), pp. 13-19.

[18] Yimprayoon P. y Ruktamatakul S. (2009). Some Statistical Aspects of Measuring Agreement Based on a Modified Kappa. *Kasetsart Journal (Natural Science)*, nº43, pp. 389 - 397

EVALUACIÓN DE UN MÉTODO DE PONDERACIÓN DE ATRIBUTOS MULTIVALUADOS EN SISTEMAS DE RECOMENDACIÓN BASADOS EN CONTENIDO

Manuel J. Barranco[1], **Jorge Castro**[1], **Luis Martínez**[1]

[1] *Universidad de Jaén, Campus Las Lagunillas, 23071, Jaén*, {barranco,jcastro,martin}@ujaen.es

Resumen

Los sistemas de recomendación basados en contenido (SRBC) junto con los sistemas de filtrado colaborativo son los más utilizados en el ámbito del comercio electrónico. Los SRBC, basándose en información histórica, construyen un perfil para cada usuario, que se compara con los productos para recomendar los que mejor se ajusten al perfil. En este proceso es de gran interés saber qué atributos o características descriptivas de los productos resultan más importantes para cada usuario, de manera que el sistema puede dar mayor peso a dichos atributos. El método TF-IDF es uno de los más utilizados para la ponderación de los atributos, sin embargo cuando éstos toman valores en dominios multivaluados dicho método podría mejorarse. En esta contribución se propone un método de ponderación basado en la entropía y los coeficientes de correlación y de contingencia con el fin de mejorar el filtrado basado en contenido en entornos con atributos de múltiples valores.

Palabras Clave: Atributos multivaluados, Sistemas de Recomendación basados en contenido, Ponderación de atributos.

1 INTRODUCCIÓN

Los sistemas de recomendación son herramientas que ayudan a los usuarios en situaciones en que pueden sentirse desbordados por la información, por lo que han sido utilizados ampliamente en el ámbito del comercio electrónico [11]. En particular, los sistemas de recomendación basados en contenido (SRBC) [1, 8, 9], usados tradicionalmente, utilizan información disponible sobre las elecciones realizadas por el usuario en el pasado, para construir su perfil que representa sus preferencias o necesidades.

Las funciones básicas de un SRBC consisten en (i) actualizar el perfil de cada usuario, (ii) comparar el perfil con los productos disponibles y (iii) recomendar los productos que mejor encajan en el perfil. En este proceso, al comparar el perfil con los productos, se debería tener en cuenta que no todos los atributos son igualmente importantes. Normalmente, cuando un usuario elige un producto se está fijando en algunos rasgos o atributos del mismo que son importantes para él/ella e ignorando otros que no le resultan de interés. Esta consideración representa una ponderación implícita sobre el conjunto de atributos de manera que cada atributo tendrá un peso subjetivo y diferente para cada usuario.

Típicamente, un SRBC trabaja con análisis textual de modo que los atributos son palabras *relevantes* que aparecen en la descripción de los productos. De esta manera cada producto va a estar descrito por una tupla de unos y ceros que indica si una palabra-atributo aparece o no en la descripción textual del producto. Sin embargo, en un caso más general, los atributos pueden estar definidos en diferentes dominios de distintos tipo: numérico, linguístico o nominal, etc.

El propósito de esta contribución es presentar un método novedoso para obtener los pesos de los atributos usando las valoraciones implícitas, que pueden obtenerse de las elecciones realizadas por los usuarios en el pasado, asignando pesos a los atributos según la ponderación del usuario. Esta ponderación sobre el conjunto de atributos se basará en dos medidas, *correlación* (para atributos cuantitativos) y *contingencia* (para atributos cualitativos). Además de estas dos medidas, también se utilizará la entropía o cantidad de información de cada atributo. Cuanto mayor sea la entropía de un atributo, más información ofrece al sistema y por tanto mayor ponderación en el proceso de filtrado.

En la siguiente sección revisaremos algunos conceptos preliminares en los que se basa el trabajo realizado. Tras ello presentaremos el método propuesto y finalmente se hará una evaluación del mismo, terminando con unas conclusiones.

2 PRELIMINARES

Esta sección revisa los SRBC y los métodos más utilizados para ponderar características o atributos.

2.1 SISTEMAS DE RECOMENDACIÓN BASADOS EN CONTENIDO

En un sistema de recomendación basado en contenido [1, 8, 9], el punto de partida es un conjunto de productos $A = \{a_i, i = 1 \dots n\}$ que pueden ser recomendados y un conjunto de características o atributos que los describen $C = \left\{c_j, j = 1 \dots m\right\}$ definidos cada uno en un dominio D_j. De este modo, cada producto a_i queda descrito por un vector $V_i = \left\{v_j^i \in D_j, j = 1 \dots m\right\}$. El sistema almacena esta información en una base de datos mediante una tabla de doble entrada (ver Tabla 1).

Tabla 1: Datos de un SRBC

	c_1	$\dots$	c_j	$\dots$	c_m
a_1	v_1^1	$\dots$	v_j^1	$\dots$	v_m^1
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_n	v_1^n	$\dots$	v_j^n	$\dots$	v_m^n

Para cada usuario, u, existe un subconjunto $A_u = \{a_{x_{ui}} \in A\,,\ 1 \leq x_{ui} \leq n\ ,\ i = 1 \dots nu\ ,\ nu \leq n\}$ de nu productos que hayan sido elegidos por él/ella, y para cada producto, $a_{x_{ui}}$, se asocia una valoración de preferencia, $r_i^u \in D_u$, siendo D_u el dominio de valoraciones del usuario (ver Tabla 2). Usando la información del usuario, el SRBC obtiene un perfil de usuario P_u que representa sus preferencias para cada atributo, y un vector de pesos W_u que incluye los pesos de cada atributo de acuerdo a la relevancia en las necesidades o preferencias del usuario (última fila en Tabla 2):

- $P_u = \left\{p_j^u \in D_j, j = 1 \dots m\right\}$ es el perfil de usuario, es decir, los valores de cada atributo que mejor se ajustan a las preferencias del usuario. Pueden obtenerse de distintas maneras [1, 8, 9].
- $W_u = \left\{w_j^u, j = 1 \dots m, 0 \leq w_j^u \leq 1\right\}$ son los pesos que muestran la relevancia de cada atributo, de acuerdo a las necesidades del usuario.

Tabla 2: Datos de usuarios en un SRBC

	c_1	$\dots$	c_m	R_u
$a_{x_{u1}}$	$v_1^{x_{u1}}$	$\dots$	$v_m^{x_{u1}}$	r_1^u
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$a_{x_{u,nu}}$	$v_1^{x_{u,nu}}$	$\dots$	$v_m^{x_{u,nu}}$	r_{nu}^u
P_u	p_1^u	$\dots$	p_m^u	
W_u	w_1^u	$\dots$	w_m^u	

Un SRBC se divide habitualmente en dos subsistemas:

- Subsistema off-line: Básicamente se encarga del mantenimiento de la base de datos:
  1. Actualización de la base de datos de productos. Un grupo de expertos participarán en la incorporación de la información que describe los elementos del sistema.
  2. Actualización de los perfiles de usuario. El sistema actualiza los perfiles de usuario basándose en la información implícita obtenida en el proceso on-line, mediante la observación de las interacciones del usuario con los productos.
- Subsistema on-line: Asiste al usuario filtrando y recomendando los productos que mejor se adapten a su perfil:
  1. Cálculo de similitudes. Para cada producto el sistema calcula su similitud con el perfil de usuario.
  2. Recomendación. El sistema selecciona los productos con mayor grado de similitud y, a continuación, se recomiendan al usuario los N mejores.

2.2 PONDERACIÓN DE ATRIBUTOS EN SRBC

En la literatura encontramos distintos métodos de ponderación de atributos en SRBC que trabajan con palabras, es decir, los atributos son palabras clave que describen los productos [9, 13]. A continuación se describe brevemente el funcionamiento de tales sistemas.

Inicialmente se construye un perfil de usuario mediante el uso de las valoraciones implícitas obtenidas a partir de los productos adquiridos o puntuados por el usuario. Los valores de los atributos dependerán de la aparición de ciertas palabras o términos linguísticos en la descripciones de los productos. De este modo, dado un atributo definido por una palabra y dado un producto, el perfil del mismo puede tomar dos valores: 1 si dicha palabra aparece en la descripción o 0 en caso contrario.

En el proceso de filtrado, el SRBC busca los productos más adecuados para el usuario, comparando el perfil del usuario y las descripciones de los productos. Una mejora consiste en dar mayor peso a aquellos atributos que se consideren más relevantes para el usuario. El método TF-IDF (Term-Fequency - Inverse Document Frequency) [2, 13] calcula el grado de relevancia de cada atributo c_j para un determinado usuario mediante la siguiente expresión:

$$W(u, c_j) = FF(u, c_j) * IUF(c_j) \tag{1}$$

La relevancia del atributo c_j para el usuario u se obtiene multiplicando dos factores: (i) Una medida de la similitud intra-usuario, FF (Feature Frecuency o frecuencia de atributo), que indica la frecuencia del atributo c_j para el usuario u y (ii) una medida de la disimilitud inter-usuarios, IUF

(Inverse User Frequency o frecuencia inversa de usuario), que ofrece un mayor peso a los atributos distintivos, es decir, a los que menos se repiten en el conjunto de usuarios.

Comunmente, el factor $FF(u,c_j)$, se obtiene sumando el número de veces que el atributo c_j aparece en los productos que el usuario u ha valorado positivamente. El segundo factor, de acuerdo al esquema TF-IDF [2], se obtiene como $IUF(c_j) = log\frac{|U|}{UF(c_j)}$ siendo $UF(c_j)$ el número de usuarios que han valorado positivamente cualquier producto que posea el atributo c_j, y $|U|$ el número total de usuarios registrados en el sistema. Este método de ponderación resulta muy adecuado en SRBC que trabajen con atributos binarios: palabras que aparecen o no en las descripciones textuales de los productos. Sin embargo, en aquellos sistemas que tratan con descripciones más complejas, con atributos multivaluados, el enfoque anterior no resulta apropiado.

El problema de ponderación de atributos multivaluados ha sido tratado en áreas como las de recuperación de información y aprendizaje automático [6, 7]. Sin embargo, este problema ha sido abordado escasamente en el área de los sistemas de recomendación. Nuestra intención es ofrecer una propuesta que trate este asunto de manera satisfactoria en SRBC.

3 PONDERACIÓN DE ATRIBUTOS BASADA EN LA ENTROPÍA Y MEDIDAS DE DEPENDENCIA

Nuestro objetivo en esta contribución es aplicar y evaluar un método de ponderación de atributos en SRBC, basado en [3], que trate con atributos multivaluados. En esta propuesta se calcula un peso para cada atributo atendiendo a:

1. Disimilitud inter-usuario: se calcula qué atributos resultan más informativos y por tanto más relevantes en el proceso de filtrado. Para este cálculo se propone el uso de la entropía. Un atributo con mayor entropía (mayor cantidad de información) resulta más útil como elemento distintivo de los gustos o preferencias de los usuarios.
2. Similitud intra-usuario: se calcula un coeficiente para medir la correlación entre las valoraciones realizadas por el usuario en el pasado y los valores de los atributos en el conjunto de productos. Para llevar a cabo este cálculo usamos dos coeficientes, dependiendo de la naturaleza de los atributos (cuantitativa o cualitativa).

El método propuesto usa la estructura de datos mostrada en la sección anterior (ver Tablas 1 y 2). Para nuestra propuesta vamos a considerar dos familias de vectores:

- Para cada producto $a_{x_{ui}}$ elegido por el usuario, $V^u_{i.} = \left\{v^{x_{ui}}_j, j = 1\ldots m\right\}$ proporciona su descripción, es decir, los valores de cada atributo para dicho producto.
- Para cada atributo c_j, $V^u_{.j} = \left\{v^{x_{ui}}_j, i = 1\ldots nu\right\}$ ofrece los valores de dicho atributo para cada uno de los productos que el usuario ha valorado en el pasado, $a_{x_{ui}}$.

El proceso completo de ponderación de atributos basado en la entropía y medidas de dependencia consta de las siguientes fases:

1. Cálculo de la disimilitud inter-usuario. Para cada atributo c_j, se cálcula la entropía H_j o cantidad de información que dicho atributo puede ofrecer.
2. Cálculo de la similitud intra-usuario. Para cada atributo c_j, dado el usuario u podemos calcular un coeficiente de dependencia, DC_{uj}, entre las valoraciones obtenidas por parte del usuario, $R_u = \{r^u_i, i = 1\ldots nu\}$, y los valores del atributo en los productos valorados, $V^u_{.j} = \left\{v^{x_{ui}}_j, i = 1\ldots nu\right\}$.
3. Cálculo de pesos. Finalmente se obtienen los pesos de los atributos como resultado de multiplicar la entropía y el grado de dependencia.

En las siguientes secciones se exponen brevemente cada una de las fases del método, que de forma esquemática puede verse en la figura 1. Para un mayor detalle puede consultarse [3].

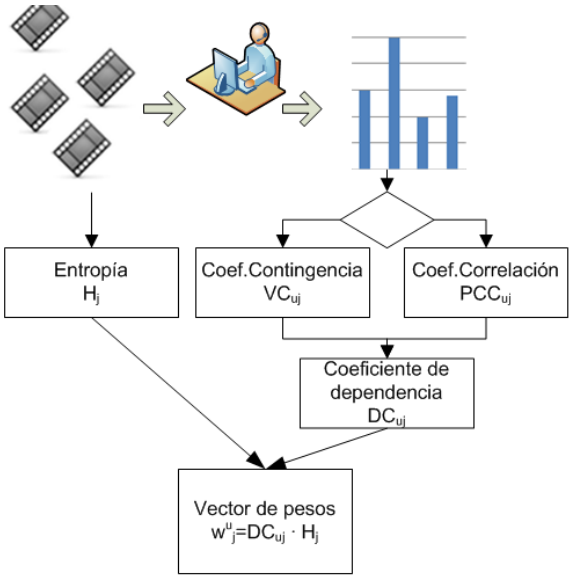


Figura 1: Ponderación de atributos basada en la entropía y medidas de dependencia

3.1 DISIMILITUD INTER-USUARIO

Para calcular la cantidad de información que proporciona cada atributo se propone el uso de la *entropía de la información* [5, 12]. Se define como la cantidad media de información, medida en bits, que contiene una variable aleatoria. Dada una variable aleatoria x, con una función de distribución de probabilidad $p(x)$, su entropía viene dada por:

$$H(x) = -\sum_i p(x_i)\, log_2\,(p(x_i)) \tag{2}$$

En el proceso de búsqueda de productos similares a un perfil de usuario dado, los atributos que tengan una entropía mayor van a resultar más interesantes y tendrán más peso en dicho proceso. Por ejemplo, en la tabla 3 tenemos dos atributos con distinta entropía, c_1 definido en el dominio $D_1 = \{A,B\}$ y c_2 en $D_2 = \{1,2,3,4,5,6\}$. Calculando las probabilidades de cada valor, dadas por su frecuencia relativa en el conjunto de datos conocidos, y aplicando la fórmula (2) tenemos que:

$$\begin{array}{l} H(c_1) = 1,\ dado\ que\ p(A) = p(B) = \frac{1}{2} \\ H(c_2) = 2.5,\ dado\ que\ p(1) = p(2) = p(3) = p(6) = \frac{1}{8} \\ \qquad\qquad y\ p(4) = p(5) = \frac{2}{8} \end{array}$$

Por tanto, el atributo c_2, con mayor entropía, está aportando más información al sistema y le daremos más peso en el proceso de filtrado.

Tabla 3: Ejemplo: dos atributos con distinta entropía

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
c_1	A	B	B	A	B	A	A	B
c_2	1	3	2	4	4	5	6	5

Así pues, para cada atributo c_j el sistema calcula la entropía H_j, y la entropía normalizada $H_j^* \in [0,1]$, como sigue:

$$\begin{array}{l} H_j = -\sum_{k_j}(f_{k_j}/n) Log_2(f_{k_j}/n) \\ H_j^* = \frac{H_j}{\sum_i H_i} \end{array} \tag{3}$$

siendo $\{k_j\}$ el conjunto de valores que el atributo c_j puede tomar, f_{k_j} la frecuencia del valor k_j en todo el conjunto de productos A y n el cardinal de A. Este cálculo considera log 0 = 0, de modo que los valores cuya frecuencia sea 0 no afecten el resultado.

3.2 SIMILITUD INTRA-USUARIO

Dado un conjunto de productos valorados por el usuario y dado un atributo, en esta fase se mide la dependencia entre dichas valoraciones realizadas por el usuario y los valores que toma dicho atributo en el conjunto de productos dado, dependiendo de su naturaleza. Si existe una dependencia entre estas variables, deduciremos que el atributo es importante para el usuario. Proponemos el uso de dos coeficientes de dependencia conocidos: el coeficiente de correlación de Pearson y el coeficiente de contingencia V de Cramer.

Coeficiente de correlación de Pearson [4]. Es un índice estadístico que mide la relación lineal entre dos variables dando una medida independiente de la escala empleada. Se obtiene dividiendo la covarianza por el producto de las desviaciones estándar de ambas variables. En nuestro caso lo aplicamos a las variables R_u y $V_{.j}^u$ obteniendo la siguiente fórmula:

$$PCC_{uj} = \frac{\sum_i r_i^u v_j^{x_{ui}} - \frac{\sum_i r_i^u \sum_i v_j^{x_{ui}}}{nu}}{\sqrt{\left(\sum_i (r_i^u)^2 - \frac{\left(\sum_i r_i^u\right)^2}{nu}\right)}\sqrt{\left(\sum_i \left(v_j^{x_{ui}}\right)^2 - \frac{\left(\sum_i v_j^{x_{ui}}\right)^2}{nu}\right)}} \tag{4}$$

Coeficiente V de Cramer [4]. Es uno de los ratios de contingencia más usados para medir la dependencia entre dos variables aleatorias, X e Y, donde al menos una de las dos es cualitativa. La fórmula para obtener dicho coeficiente sobre nuestros datos es la siguiente:

$$VC_{uj} = \sqrt{\frac{\sum_{k_u}\sum_{k_j} \frac{\left(f_{k_u,k_j} - \frac{f_{k_u} f_{k_j}}{n_u}\right)^2}{\frac{f_{k_u} f_{k_j}}{nu}}}{nu\ \ min\left(|D_u|, |D_j|\right)}} \tag{5}$$

donde k_u y k_j son índices en los conjuntos de datos R_u y $V_{.j}^u$ respectivamente, f_{k_u}, f_{k_j} son las frecuencias de los valores indexados por k_u y k_j respectivamente, y f_{k_u,k_j} es la frecuencia de la ocurrencia simultánea de los dos valores indexados por k_u y k_j.

Así pues, el coeficiente de dependencia, DC, entre las valoraciones realizadas por el usuario u sobre un conjunto de productos, y los valores del atributo j para cada uno de estos productos, viene dado por la siguiente expresión:

$$DC_{uj} = \begin{cases} |PCC_{uj}| & si\ c_j\ es\ cuantitativo \\ VC_{uj} & si\ c_j\ es\ cualitativo \end{cases}$$

3.3 CÁLCULO DE PESOS DE ATRIBUTOS

Una vez que los factores H_j^* y DC_{uj} han sido obtenidos, el sistema va a calcular el peso de cada atributo c_j como producto de ambos factores, de acuerdo a la fórmula (1).

$$w_j^u = DC_{uj} \cdot H_j^* \tag{6}$$

Para normalizar el vector de pesos $\{w_i\}$ debemos satisfacer la propiedad $\sum w_i = 1$, con lo que obtendremos:

$$W_u^* = \left\{ w_j^{*u} = \frac{w_j^u}{\sum_i w_i^u} \mid j = 1, \ldots, m \right\} \quad (7)$$

4 EVALUACIÓN

Para evaluar el método propuesto, se ha implementado y se han obtenido varias medidas usadas habitualmente en la evaluación de algoritmos de recuperación de información. Así mismo, dichas medidas se han aplicado a otras implementaciones de SRBC, con objeto de poder compararlas con nuestros resultados. Concretamente, las implementaciones realizadas han sido [1, 9]:

- Modelo booleano de SRBC sin ponderación de atributos usando la distancia euclídea como medida de similitud (CB-Euclídea)
- Modelo booleano de SRBC sin ponderación de atributos usando la función coseno como medida de similitud (CB-Coseno)
- Modelo booleano con ponderación de atributos TF-IDF.
- Modelo propuesto con ponderación de atributos multivaluados basado en la entropía y medidas de dependencia (PABED).

4.1 CASO DE ESTUDIO

El conjunto de datos utilizado para verificar la validez del método propuesto es el extraído del sistema *Movilens* (http://www.movielens.org) desarrollado por el grupo de investigación *GroupLens Research* de la Universidad de Minesota. Es un servicio libre que permite valorar películas para, a partir de dichas valoraciones, ofrecer recomendaciones a los usuarios siguiendo un modelo de filtrado colaborativo. En nuestro caso, solo hacemos uso de los datos para aplicarlos a un modelo de recomendación basado en contenido.

Para las pruebas descritas en este artículo, se han seleccionado los usuarios con 20 valoraciones o más, obteniendo así un conjunto de datos de 9464734 valoraciones hechas por 69878 usuarios sobre 9768 películas.

El conjunto de datos extraído consta de tuplas <usuario,película,valoración>, donde la valoración viene dada de 1 a 5, siendo 1 la peor puntuación y 5 la mejor. La información descriptiva de las películas la hemos obtenido de la base de datos *IMDB* (http://www.imdb.com) considerando los siguientes atributos: título, año, género, director y país.

Para evaluar la eficacia de las distintas implementaciones se han usado métricas habituales [10]: precisión, recall y f-medida.

$$precision = \frac{relevantes\ recomendadas}{recomendadas} \quad (8)$$

$$recall = \frac{relevantes\ recomendadas}{relevantes} \quad (9)$$

$$f - medida = (1 + \beta^2) * \frac{precision * recall}{\beta^2 * precision + recall} \quad (10)$$

Para el cálculo de la precisión y el recall, consideramos relevantes para un usuario las películas con una valoración de 4 ó 5. En el cálculo de la f-medida se ha utilizado $\beta = 1$, dando igual importancia a la precisión y al recall.

Para evaluar la eficiencia de las implementaciones a comparar, se ha usado el tiempo de ejecución medio. En este experimento, las ejecuciones se han realizadon en un servidor dedicado con una CPU de cuatro núcleos de 2.00GHz con 2MB de caché y 8GB de memoria RAM.

4.2 RESULTADOS

El experimento ha consistido en realizar cincuenta ejecuciones de cada implementación utilizando la técnica de validación cruzada (Cross Validation) con 5 particiones. Asimismo, se han probado diferentes tamaños de la lista de items recomendados (K=1, 5, 10, 50, 100) obteniendo diferentes resultados. Tal y como puede observarse en las Figuras 2, 3 y 4, el método propuesto (PABED) supera considerablemente al método básico, sin ponderación (CB-Euclidea y CB-Coseno), en las métricas de precisión, recall y f-medida. En cuanto a la comparación con el método TF-IDF, se obtienen resultados muy similares en las métricas mencionadas (incluso se observa una ligera mejora para valores altos de K) pero en la métrica de eficiencia (tiempo de ejecución) la mejora conseguida es mucho más acentuada (ver Figura 5).

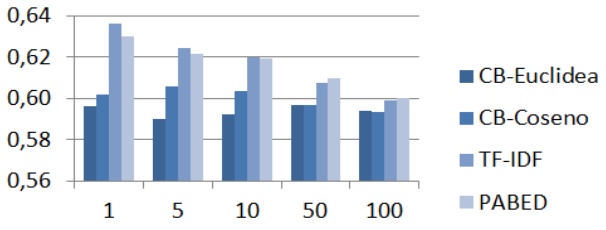


Figura 2: Precisión de cada sistema en función de k.

5 CONCLUSIONES

Los métodos de ponderación de atributos mejoran considerablemente los resultados en los sistemas de recomendación basados en contenido. TF-IDF es el método de ponderación más utilizado cuyo funcionamiento es bueno con

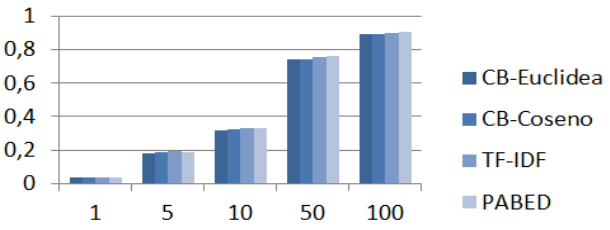


Figura 3: Recall de cada sistema en función de k.

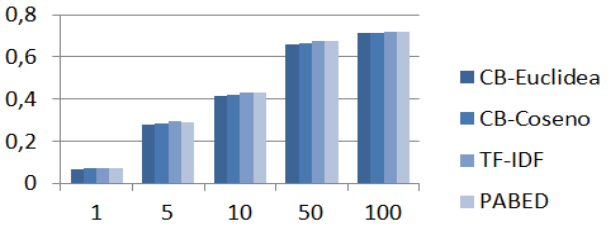


Figura 4: F1-medida de cada sistema en función de k.

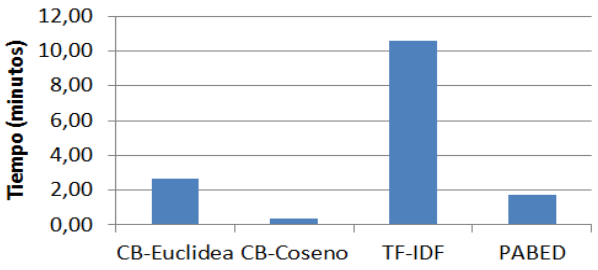


Figura 5: Tiempo de ejecución de cada sistema.

atributos booleanos. No obstante, cuando los atributos toman múltiples valores, dicho método puede mejorarse. En esta contribución hemos propuesto y evaluado un nuevo método para calcular pesos de atributos en sistemas de recomendación basados en contenido con atributos multivaluados que pueden ser tanto cuantitativos como cualitativos. El nuevo método está basado en dos factores: similitud intra-usuario y disimilitud inter-usuario. El primer factor se calcula con los coeficientes de correlación de Pearson, para atributos cuantitativos y de contingencia V de Cramer, para los cualitativos. Y para el segundo factor usamos la entropía que mide la cantidad de información aportada por cada atributo. El método ha sido evaluado obteniendo resultados satisfactorios tanto en eficacia como en eficiencia.

Agradecimientos

Esta contribución está parcialmente financiada por los proyectos de investigación TIN2009-08286, P08-TIC-03598, AGR-6581 y los fondos FEDER.

Referencias

[1] G. Adomavicius, A. Tuzhilin: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Trans. on Knowledge and Data Engineering, vol 17, No. 6, June, pp.734-749, (2005)

[2] Akiko Aizawa: An information-theoretic perspective of TF-IDF measures. Information Processing and Managemente, 39:45-65 (2003)

[3] Manuel J. Barranco, Luis Martínez: A Method for Weighting Multi-valued Features in Content-Based Filtering. IEA/AIE 2010, Part III, LNAI 6098, Springer-Verlag, Berlin, pp. 409-418 (2010)

[4] Y.M.M.Bishop, S.E. Fienberg, P.W. Holland: Discrete Multivariate Analysis: Theory and Practice. The MIT Press, England (1995)

[5] T.M. Cover, J.A. Thomas: Elements of Information Theory. John Wiley & Sons, Inc. (1991)

[6] Tzung-Pei Hong, Jyh-Bin Chen: Finding relevant attributes and membership functions. Fuzzy Sets and Systems 103: 389-404 (1999)

[7] George H. John, Ron Kohavi, Karl Pfleger: Irrelevant features and the subset selection problem. Machine Learning: Proc. of the 11th int. conf., 121-129, Morgan Kaufmann Publishers, San Francisco, CA. (1994)

[8] L. Martínez, L.G. Pérez, M.J. Barranco: A Multigranular Linguistic Content-Based Recommendation Model. International Journal of Intelligent Systems, 22:5, pp.419-434 (2007)

[9] Michael J. Pazzani, Daniel Billsus: Content-Based Recommendation Systems, The Adaptive Web. Lecture Notes in Computer Science, Springer-Verlag, Vol. 4321, pp. 325-341 (2007)

[10] Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Application of dimensionality reduction in recommender systems-a case study. In ACM WebKDD Workshop (2000)

[11] J. Ben Schafer, Joseph A. Konstan, John Riedl: E-Commerce Recommendation Applications. Data Mining and Knowledge Discovery, 5, pp. 115-153, (2001)

[12] C.E. Shannon: A mathematical theory of communication. The Bell System Technical Journal, 27:379-423,623-656 (1948)

[13] P. Symeonidis, A. Nanopoulos, Y. Manolopoulos: Feature-weighted user model form recommender systems. Lecture Notes in Computer Science, Springer-Verlag, 4511:97-106 (2007)

APRENDIZAJE EVOLUTIVO DE UN SISTEMA BASADO EN REGLAS DIFUSAS PARA UN SISTEMA DE CONTROL DE ENTORNO POR MEDIO DE PUNTERO LÁSER

Francisco Chávez[1], **Francisco Fernández**[1], **Rafael Alcalá**[2], **Jesús Alcalá**[2], **Francisco Herrera**[2]

[1]*Dept. De Sistemas Informáticos y Telemáticos, Universidad de Extremdadura. España*, {fchavez, fcofdez}@unex.es
[2]*Dept. de Ciencias de la Computación e Inteligencia Artificial, CITIC-UGR, Universidad de Granada. España*, {alcala, jalcala, herrera}@decsai.ugr.es

Resumen

Estudios recientes en hogares inteligentes han propuesto métodos para utilizar un puntero láser que interactúa con los dispositivos eléctricos y electrónicos del hogar, lo que representa un dispositivo fácil de usar y barato para el control de un hogar. Sin embargo, la detección del punto láser originado por el dispositivo es actualmente un problema abierto, ya que se pretende utilizar cámaras estándar sin costosos filtros, lo que nos obliga a utilizar algoritmos especializados de búsqueda en una imagen.

En este artículo proponemos un método híbrido, que combina una técnica clásica utilizada en los procesos de detección de imágenes (*Template Matching*) junto con un aprendizaje evolutivo de los sistemas basados en reglas difusas para la detección de puntos láser en entornos domésticos reales. Este trabajo analiza la eficacia del sistema propuesto, que mejora la tasa de éxito en las imágenes sin punto láser junto con una notable disminución de falsos positivos, al compararlo con sistemas previos utilizados para la detección del punto láser.

Palabras Clave: Sistemas de Interacción, Sistemas de control domóticos, Detección de puntos láser, Sistemas Basados en Reglas Difusas, Aprendizaje Evolutivo.

1 Introducción.

Hoy en día, una vivienda puede ser controlada por diferentes dispositivos, como dispositivos especiales de control en el interior, o por medio de teléfonos móviles, ordenadores, usando *Internet* si los usuarios están al aire libre. Recientes investigaciones en los hogares inteligentes proporcionan grandes soluciones para el control de dispositivos para el hogar [1].

Por otro lado, el trabajo presentado en [9] hace uso de un robot para ayudar a personas con discapacidad a la hora de coger objetos. En este trabajo, los autores también proponen una interesante forma de señalar los objetos deseados por medio de un puntero láser. El robot debe ser capaz de detectar el punto láser sobre un objeto para poder cogerlo posteriormente. Otro uso de los punteros láser ha sido como elemento indicador para el control de grandes pantallas [10, 15]. Esta idea ha sido adaptada a los sitemas de control de entorno que presentamos en este trabajo. El objetivo principal en este tipo de sistemas es la detección del punto láser de forma eficaz y barata sin utilizar costosos equipos para ello.

Una alternativa es el análisis de imágenes utilizando cámaras estándar [2], donde un conjunto de algoritmos son capaces de detectar la presencia de un punto láser. Sin embargo, estos algoritmos presentan falsos positivos, lo que provocaría el envío de una orden incorrecta al sistema de control domótico, obteniendo así situaciones potencialmente peligrosas e incontroladas. Por esta razón, en [3] se presenta un nuevo enfoque de hibridación de una técnica clásica con el uso de un Sistema de Basado en Reglas Difusas (SBRD) diseñado a partir del conocimiento de expertos [14] para tratar de mejorar la tasa de éxito en las imágenes sin punto láser. Además, este SBRD inicial fue ajustado mediante un Algoritmo Genético (AG) en el trabajo presentado en [4], con el fin de mejorar la tasa de éxito en las imágenes sin punto láser.

En este trabajo se presenta un nuevo enfoque para detectar el punto láser con respecto al presentado en [3] y [4] por medio de un aprendizaje evolutivo del SBRD para el sistema de detección de puntos láser. Consiste en un algoritmo de aprendizaje integrado que ajusta la Base de Datos (BD) basándose en el uso de la representación de 2-tuplas lingüísticas [7] con el objetivo de obtener la mejor Base de Reglas (BR). Los resultados obtenidos mediante el método propuesto utilizan un modelo de validación cruzada para un

conjunto grande de ejemplos, 990 imágenes en total, que simulan una situación real del sistema de control, donde recogen las diferentes situaciones de luz y distancia que se puede dar en un hogar convencional. Estos resultados muestran una mejor tasa de éxito general con respecto a las técnicas anteriores. Pero cabe destacar la ventaja más importante obtenida con el modelo, la práctica eliminación de los falsos positivos.

Este trabajo se divide como sigue: En la sección 2, se describe el sistema de control de entorno junto con los algoritmos utilizados para la detección del punto láser. La sección 3 presenta el aprendizaje evolutivo del SBRD utilizado en este trabajo. Los resultados obtenidos con la técnica clásica, y los enfoques híbridos basados en la derivación del SBRD a partir del conocimiento experto, el ajuste automático y el aprendizaje automático se describen y analizan mediante la comparación de ellos en la sección 4. Finalmente, en la sección 5 se exponen las conclusiones del trabajo.

2 Descripción del sistema y trabajos previos.

Como se ha mencionado, el sistema de control de entorno por medio de un apuntador láser permite a sus usuarios indicar qué dispositivo desean controlar. El sistema está dividido en tres secciones (ver figura 1).

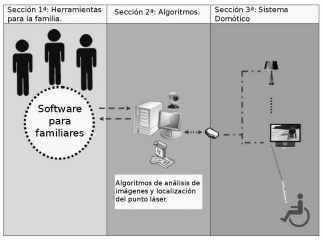


Figura 1: Sistema de Control de Entorno.

- Sección 1. La herramienta de configuración permite a los usuarios seleccionar los dispositivos del hogar que serán controlados por el sistema. Con esta herramienta, se marcarán las diferentes zonas activas para cada dispositivo, utilizando las imágenes enviadas por una cámara de vídeo estándar.
- Sección 2. En esta sección, el sistema utiliza diferentes técnicas para analizar la imagen enviada por la cámara de video y localizar la posición del punto láser.
- Sección 3. La arquitectura KNX/EIB [11] permite controlar los dispositivos del hogar una vez que el punto láser se encuentra y éste está sobre una zona activa. El sistema envía al sistema domótico las órdenes necesarias para activar/desactivar el dispositivo seleccionado.

La principal contribución de este trabajo se centra en la sección 2, con el fin de mejorar la capacidad de detección de punto de láser. En la figura 2, puede verse un ejemplo de las imágenes con y sin punto láser.

Figura 2: Imagen con punto láser (izq). Imagen sin punto láser (der).

Los algoritmos utilizados en trabajos anteriores para la detección del punto láser han sido los siguientes:

1. Umbralización Dinámica (UD)
2. Template Matching (TM)
3. Template Matching + Umbralización Dinámica ($TM+UD$)
4. Template Matching + SBRD diseñado por un experto ($TM+SBRD_{experto}$)
5. Template Matching + SBRD ajustado automáticamente por un AG ($TM+SBRD_{ajustado-AG}$).

El sistema de control de entorno basado en el puntero láser que se presenta en [2] utilizó tres algoritmos clásicos diferentes, UD, TM y TM + UD. En [3] se propone un modelo híbrido inicial mediante la combinación de TM con un SBRD obtenidos a partir de la experiencia de expertos, donde posteriormete se ajustan las funciones de pertenencia del mismo a través de un AG [4].

3 Template Matching más sistema de aprendizaje evolutivo para detección del punto láser.

En esta sección, se presenta un nuevo enfoque del problema presentado. Éste se basa en el uso de una técnica más avanzada para un aprendizaje completo del SBRD utilizado para detectar el punto láser de la imagen obtenida por TM. Por lo tanto, en este trabajo se propone la técnica de TM en combinación con el aprendizaje de la Base de Conocimiento (BC) (BD+BR) completa.

Con el fin de aplicar eficazmente este nuevo enfoque a nuestro problema concreto, consideramos ocho variables de entrada: las variables utilizadas en el SBRD inicial. Para mayor detalle el lector puede consultar el trabajo presentado en [3], donde se describe por completo la extracción

de las variables. Como complemento al conjunto incial de variables, se extraen tres nuevas variables que representan el color especial de láser (rojo, verde y azul). El sistema utiliza una variable de salida para denotar si la información analizada corresponde o no a un punto láser. A este nuevo enfoque del SBRD inicial, utilizando el aprendizaje de reglas y al AG lo llamaremos $SBRD_{apr-AG}$, que en combinación con la técnica de TM da paso a un nuevo sistema híbrido, $TM + SBRD_{apr-AG}$.

De esta manera, se presenta un algoritmo de aprendizaje de la BC que puede administrar eficientemente el espacio de búsqueda. El esquema de aprendizaje se presenta en la siguiente subsección, donde se introduce un modelo de representación de la BD, lo que disminuye el espacio de búsqueda con respecto a las representaciones clásicas de las funciones de pertenencia triangulares. Finalmente, se presenta el AG para el aprendizaje del SBRD.

3.1 Esquema del SBRD basado en aprendizaje.

El aprendizaje de la BC a través del aprendizaje de la BD, es un enfoque reciente que implica un espacio de búsqueda más sencillo que el aprendizaje clásico de la BC. Consiste en la obtención de la BD y por separado una BR, basado en un aprendizaje [5, 12] (véase la figura 3). Esta forma de trabajar nos permite conocer el contexto más adecuado para cada partición difusa, que es necesaria en diferentes contextos de aplicación y modelos difusos de extracción de reglas.

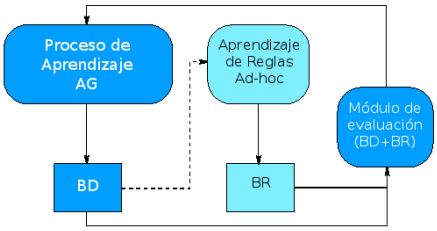


Figura 3: Esquema de aprendizaje de la Base de Conocimiento.

El esquema de aprendizaje utilizado en este trabajo está compuesto por:

- Proceso para aprender la BD, que permite definir:
  - El número de etiquetas por cada variable lingüística.
  - El desplazamiento lateral de cada variable.

  Se utilizan funciones de pertenencia triangulares debido a su simplicidad.
- Se utiliza un método *Ad-hoc* basado en datos para obtener la BR [13]. Este método se ejecuta para cada definición de la BD generada por el AG utilizado, y por lo tanto, permite que el proceso de aprendizaje híbrido propuesto obtenga finalmente la definición de la BC (BD y BR) por medio de la acción cooperativa de ambos métodos.

3.2 Representación del modelo lingüístico de 2-Tuplas.

El modelo de representación lingüístico de 2-Tuplas introducido en [7] permite un desplazamiento lateral del soporte de una etiqueta y mantiene la interpretabilidad asociada al SBRD lingüistico obtenido. Este modelo incluye un nuevo enfoque basado en el concepto de translación simbólica, que permite desplazar lateralmente el soporte de las etiquetas considerando un único parámetro (pequeños desplazamientos a la izquierda/derecha de las funciones de pertenencia originales), con el objetivo de obtener etiquetas que contengan un conjunto de ejemplos con mejor cubrimiento (búsqueda de precisión) conservando su forma original, y reducir el espacio de búsqueda del proceso de ajuste respecto al ajuste clásico, que normalmente considera 3 parámetros en caso de funciones de pertenencia triangulares.

La translación simbólica de una etiqueta es un número (α) dentro del intervalo [-0.5, 0.5), representando este intervalo el dominio de la etiqueta cuando ésta se mueve entre sus dos etiquetas laterales adyacentes (Ver figura 4). Consideremos un conjunto de etiquetas $S(S = \{s_1, s_2, \ldots, s_n\})$ representando una pertición difusa. Formalmente, para representar la translación simbólica de una etiqueta en S tenemos el par,

$$(s_i, \alpha_i), \quad s_i \in S, \ \alpha_i \in [0.5, -0.5).$$

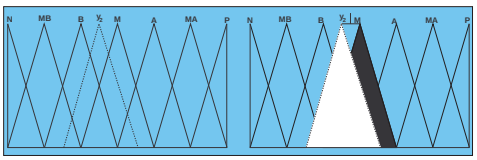


Figura 4: Desplazamiento lateral de las etiquetas lingüísticas.

La figura 4 muestra el desplazamiento lateral de las diferentes etiquetas, siendo $y2$ una nueva etiqueta que representa a la antigua M desplazada lateralmente, localizándose entre B y M.

3.3 Algoritmo Genético utilizado.

En este trabajo se utilizará CHC [6] para diseñar el algoritmo de aprendizaje. CHC es un AG que presenta un equilibrio entre exploración y explotación, usado para problemas de espacio de búsqueda complejos. Considerando este enfoque, el proceso de aprendizaje de la BD debe ser definido

como: la granularidad de la particiones lingüísticas y los desplazamientos laterales de las funciones de pertenencia.

3.3.1 Codificación de la Base de Datos.

La *granularidad* y los *parámetros de traslación* se representan por $C = C_1 + C_2$, donde:

- C_1 (Número de etiquetas) : Vector de tamaño N (siendo N el número de variables del sistema). El conjunto de etiquetas se encuentra entre $\{3,\ldots,9\}$:

$$C_1 = (L^1,\ldots,L^N) \ .$$

- C_2 (Desplazameinto lateral): Vector de números reales de tamaño $N*9$ (N variables con un máximo de 9 etiquetas por variable). Si un cromosoma no tiene el número máximo de etiquetas en una variable, el espacio reservado sobrante es ignorado en el proceso de evaluación. Por ello, la parte C_2 tiene la siguiente estructura:

$$C_2 = (\alpha_1^1,\ldots,\alpha_{L^1}^1,\ldots,\alpha_1^N,\ldots,\alpha_{L^N}^N)$$

3.3.2 Evaluación del cromosoma.

Para evaluar un determinado cromosoma se aplicará el método de Wang y Mendel [13] para obtener la BR. Una vez obtenida la BC (BD + BR) completa se trata de minimizar la función:

$$F_C = w_1 \cdot FE + w_2 \cdot NR,$$

donde:

- NR es el número de reglas obtenidas, penalizando un exeso de reglas.
- $w_1 = 1$.
- w_2 es un factor del Error Cuarático Medio (ECM) y NR obtenido por el modelo, que se obtendría fijando el número máximo de etiquetas a 9 sin considerar los parámetros de desplazamiento.

$$w_2 = \alpha \cdot \frac{ECM_{max-etq}}{NR_{max-etq}} \ ,$$

  siendo α un porcentaje fijado por un experto que determina el equilibrio entre precisión y complejidad. Se utilizan valores mayores a 1.0 para modelos con pocas reglas y valores menores que 1.0 para modelos con gran precisión. En este problema utilizamos $\alpha = 1.15$.

- FE es la Función de Evaluación utilizada para analizar el número de falsos negativos (FN) y falsos positivos (FP) obtenidos con la BC generada,

$$FE = \frac{|FN|}{|D|} + 10 \cdot \frac{|FP|}{|D|}$$

  donde $|FN|$ es el número de falsos negativos, $|FP|$ número de falsos positivos y $|D|$ tamaño del conjunto de datos.

  Para cada ejemplo, el SBRD debe generar un valor en el intervalo [0, 1]. Si ese valor es mayor a 0.5, el ejemplo se considera como punto láser, en caso contrario como imagen sin punto láser, nos podemos encontrar las siguientes situaciones:

  - FN: Si el ejemplo es clasificado como imagen sin láser pero realmente es una imagen con láser.
  - FP: Si el ejemplo es clasificado como imagen con láser pero realmente es una imagen sin láser.
  - Acierto: El ejemplo está bien clasificado.

4 Resultados.

Para evaluar la eficacia de la propuesta presentada en este trabajo, hemos considerado el sistema de control de entorno por medio de un apuntador láser presentado en [2]. A su vez, para demostrar la eficacia y mejora en los resultados del sistema presentado en [4] junto con el actual, se ha ampliado la base de ejemplos de 105 imágenes usadas en [4] a un conjunto de 990 imágenes. Este nuevo conjunto ha sido dividido aleatoriamente en 5 subconjuntos, con el 20 % de las imágenes para cada uno de ellos. Una vez creado los subconjuntos se ha utilizado la técnica de *5-fold cross-validation* para obtener los resultados de los diferentes sistemas, utilizando en cada iteración del método un total de 792 imágenes de entrenamiento y 198 de test.

La tabla 1 presenta una descripción de los métodos estudiados en este trabajo.

Los parámetros utilizados por los diferentes algoritmos $TM + SBRD_{experto}$ y $TM + SBRD_{ajustado-AG}$ fueron seleccionados de acuerdo con las recomendaciones de los autores utilizadas en [3] y [4]. Por último, el número de evaluaciones utilizadas por el AG en el proceso de ajuste para el algoritmo $TM + SBRD_{ajustado-AG}$ fué de 50.000. Los valores de entrada de los parámetros utilizados en este trabajo son los siguientes:

- Evaluaciones = 50.000
- Tamaño de la población = 50
- Bits por gen = 30
- Factor α = 1.15

La tabla 2 presenta los resultados medios del algoritmo $TM + UD$ utilizando el nuevo conjutno de 990 imágenes con una configuración inicial similar a la presentada en [2]. Igualmente, se muestran los resultados medios de los sistemas $TM + SBRD_{experto}$ y $TM + SBRD_{ajustado-AG}$ con

Tabla 1: Técnicas utilizadas en los experimentos.

Ref	Método	Año	Descripción
[2]	$TM+UD$	2008	Template Matching más Umbralización Dinámica
[3]	$TM+SBRD_{experto}$	2010	Template Matching más SBRD diseñado por experto
[4]	$TM+SBRD_{ajustado-AG}$	2010	Template Matching más SBRD diseñado por experto y ajustado por AG
Nuestra propuesta	$TM+SBRD_{apr-AG}$		Template Matching más SBRD basado en aprendizaje

los parámetros utilizados en los trabajos [3] y [4] respectivamente. Podemos observar que las técnicas clásicas como $TM+UD$ tienen mejores resultados que los algoritmos que utilizan un SBRD como $TM+SBRD_{experto}$ y $TM+SBRD_{ajustado-AG}$ utilizando el nuevo y más completo conjunto de imágenes. Ésto se debe a que los sistemas clásicos trabajan con datos heterogéneos tales como brillo y luminosidad de la imagen, sin embargo, el SBRD trabaja con datos con la forma y apariencia de la imagen, tales como similitud a un círculo, entre otros. Este tipo de datos es muy útil cuando se trabaja con algoritmos basados en aprendizaje. Por esta razón, se han combinado técnicas clásicas TM junto con algoritmos de aprendizaje, obteniendo una nueva hibridación llamada $TM+SBRD_{apr-AG}$. Los resultados de este nuevo algoritmo son los mejores que se obtienen para la serie de experimentos presentados en este trabajo, como puede verse en la tabla 3.

Podemos observar que el porcentaje de aciertos en imágenes con punto láser es del 70,50 % y en imágenes sin punto láser es de 98,76 %, muy cercano al porcentaje óptimo del sistema que sería el 100 %.

5 Conclusiones.

En este trabajo se presenta un nuevo modelo híbrido que combina técnicas clásicas como *Template Matching* junto con aprendizaje evolutivo de SBRD para la detección de puntos láser en un sistema de control de entorno. Gracias a este tipo de sistemas, junto con un sistema de control domótico de viviendas, se obtiene un sistema de interacción amigable y fácil de utilizar, así como barato.

El trabajo recoge un estudio completo donde se compara la eficacia de algoritmos que utilizan técnicas clásicas presentados en trabajos anteriores, con los nuevos algoritmos híbridos que combinan técnicas clásicas y aprendizaje evolutivo de SBRDs. Para este estudio se considera un conjunto de datos de 990 imágenes que simula una situación real del sistema de control, donde recogen las diferentes situaciones de luz y distancia que se puede dar en un hogar convencional. Podemos ver que el algoritmo propuesto obtiene mejores resultados que los anteriores, pero el dato más significativo es que este nuevo sistema obtiene una tasa de aciertos del 98,76 % en imágenes sin punto láser, reduciendo de esta forma los falsos positivos, tan perjudiciales para el sistema, hasta el porcentaje ideal del 100 %.

El nuevo enfoque presentado en este trabajo es particularmente útil para actuar a distancia sobre los dispositivos de una vivienda estándar. Por otra parte, creemos que los resultados muestran una utilidad añadida en la industria, tal como el uso de dispositivos en entornos donde no son fácilmente accesibles, ya que con el puntero láser podrán ser controlados a distancia.

Agradeciminetos.

Este trabajo ha sido subvencionado por el proyecto TIN2008-06681-C06-01 del Ministerio de Educación y Ciencia y la Consejería de Economía-Comercio e Innovación de la Junta de Extremadura, bajo el proyecto GRU09105.

Referencias

[1] Chan M., Estève D., Escriba C. and Campo E., A review of smart homes-present state and future challenges, *Computer Methods and Programs in Biomedicine*, vol. 91, no. 1, pp. 55–81, 2008.

[2] Chávez F., Fernández F., Olague G. and Llano J., An independent and non-intrusive laser pointer environment control device system, *Proceedings of the 5th international conference on Pervasive services (ICPS'08)*. Sorrento, Italy: ACM, pp. 37–46, 2008.

[3] Chávez F., Fernández F., Alcalá R., Alcalá-Fdez j., Olague G. and Herrera F., Hybrid laser pointer detection algorithm based on template matching and fuzzy rule-based systems for domotic control in real home environments, *Applied Intelligence*, pp. 1–17, 10.1007/s10489-010-0268-6, 2010.

[4] Chávez F., Fernández F., Alcalá-Fdez J., Alcalá R., Herrera F. and Olague G., Genetic tuning of a laser pointer environment control device system for handicapped people with fuzzy systems, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2010)*, pp. 1–8, 2010.

[5] Cordón O., Herrera F., Magdalena L. and Villar P., A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base, *Information Sciences*, vol. 136, pp. 85–107, 2001.

Tabla 2: Resultados de $TM+UD$

Resultados medios de $TM+UD$				
	Entrenamiento		**Test**	
	Media	Desviación Estándar	Media	Desviación Estándar
Aciertos Generales	83,48 %	0,01206	83,54 %	0,04705
Tasa de acierto con Punto Láser	67,86 %	0,01575	68,13 %	0,05705
Tasa de acierto sin Punto Láser	98,32 %	0,00481	98,36 %	0,01647
Resultados medios de $TM+SBRD_{experto}$				
	Entrenamiento		**Test**	
	Media	Desviación Estándar	Media	Desviación Estándar
Aciertos Generales	78,59 %	0,01229	78,59 %	0,04917
Tasa de acierto con Punto Láser	66,21 %	0,01845	66,52 %	0,06770
Tasa de acierto sin Punto Láser	90,35 %	0,00821	90,30 %	0,03226
Resultados medios de $TM+SBRD_{ajustado-AG}$				
	Entrenamiento		**Test**	
	Media	Desviación Estándar	Media	Desviación Estándar
Aciertos Generales	75,80 %	1,81965	74,97 %	5,03560
Tasa de acierto con Punto Láser	51,44 %	3,59527	50,51 %	6,84627
Tasa de acierto sin Punto Láser	98,91 %	0,33398	98,35 %	1,34530

Tabla 3: Resultados de $TM+SBRD_{apr-AG}$

	Entrenamiento		**Test**	
	Media	Desviación Estándar	Media	Desviación Estándar
Aciertos Generales	87,27 %	0,78990	84,92 %	2,90425
Tasa de acierto con Punto Láser	73,87 %	1,28823	70,50 %	4,48146
Tasa de acierto sin Punto Láser	100 %	0,00000	98,76 %	1,14364

[6] Eshelman L., The chc adaptive search algorithm: How to have safe serach when engaging in nontraditional genetic recombination, *Foundations of genetic Algorithms*, vol. 1, pp. 265–283, 1991.

[7] Herrera F. and Martínez L., A 2-tuple fuzzy linguistic representation model for computing with words, *IEEE Transactions on Fuzzy Systems*, vol. 8, pp. 746–752, 1999.

[8] Herrera F., Lozano M. and Sánchez A.M., A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study, *International Journal of Intelligent Systems*, vol. 18, pp. 309–338, 2003.

[9] Kemp C. et al., A point-and-click interface for the real world: laser designation of objects for mobile manipulation, *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction (HRI'08)*, pp. 241–248, 2008.

[10] Kim N. W., Lee S. J., Lee B. G. and Lee J. J., Vision based laser pointer interaction for flexible screens, *Human-Computer Interaction. Interaction Platforms and Techniques, Lecture Notes in Computer Science*, vol. 4551, pp. 845–853, 2007.

[11] KNX/EIB, *http://www.knx.org*.

[12] Teng Y. W. and Wang W. J., Constructing a user-friendly ga-based fuzzy system directly from numerical data, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 5, pp. 2060–2070, 2004.

[13] Wang L. X. and Mendel J. M., Generating fuzzy rules by learning from examples, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414–1427, 1992.

[14] Zadeh L., Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 28–44, 1973.

[15] Zhang L., Shi Y., and Chen J., Drag and drop by laser pointer: Seamless interaction with multiple large displays, *Ubiquitous Intelligence and Computing, Lecture Notes in Computer Science*, vol. 4159, pp. 12–20, 2006.

AJUSTE GENÉTICO LATERAL DE LAS ETIQUETAS LINGÜÍSTICAS EN DESCUBRIMIENTO DE SUBGRUPOS

C.J. Carmona, **P. González**, **M.J. Gacto**, **M.J. del Jesus**

Departamento de Informática. Universidad de Jaén. Campus las Lagunillas, s/n. 23071 Jaén (Jaén),
{ccarmona,pglez,mgacto,mjjesus}@ujaen.es

Resumen

El descubrimiento de subgrupos es una técnica de minería de datos cuyo objetivo es la obtención de conocimiento descriptivo mediante aprendizaje supervisado. Este conocimiento permite descubrir relaciones atípicas y relevantes en los datos para una variable de interés.
En este trabajo se presenta un método de ajuste, aplicado en una etapa de post procesamiento, para mejorar la calidad de las reglas obtenidas previamente por cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa. Este nuevo método consiste en la modificación o ajuste de las etiquetas lingüísticas de un conjunto de reglas mediante un desplazamiento lateral de las mismas, basado en la representación de 2-tuplas lingüísticas. Como muestran los resultados obtenidos, este nuevo enfoque consigue mejoras significativas en la mayoría de las medidas de calidad para descubrimiento de subgrupos.

Palabras Clave: Descubrimiento de subgrupos, Sistemas difusos evolutivos, NMEEF-SD, Ajuste genético lateral.

1 INTRODUCCIÓN

El descubrimiento de subgrupos es una técnica de minería de datos descriptiva que se resuelve mediante aprendizaje supervisado; es decir, su principal objetivo consiste en obtener conocimiento que describa el problema para una variable objetivo. Las principales propiedades que un algoritmo de descubrimiento de subgrupos debe aportar al problema es la obtención de relaciones parciales o genéricas con una alta precisión y novedad. Además, los subgrupos deben cubrir a la mayor parte de ejemplos de la variable objetivo.

A pesar de que en la bibliografía existe un amplio abanico de algoritmos para realizar esta tarea, únicamente los algoritmos SDIGA [9], MESDIF [8] y NMEEF-SD [6] utilizan una representación del conocimiento basada en lógica difusa [23]. En concreto, estos algoritmos basados en sistemas difusos evolutivos [14] están orientados a la búsqueda de modelos y reglas con alta interpretabilidad que obtengan valores adecuados en las medidas de calidad utilizadas para descubrimiento de subgrupos.

La búsqueda de alta interpretabilidad puede limitar la precisión en los modelos obtenidos, sobre todo en problemas del mundo real, ya que se promueve en mayor medida la búsqueda de reglas muy generales. Es por ello que en determinados problemas puede ser necesario paliar esta pérdida de precisión aplicando modelos de ajuste para optimizar la precisión sin alterar demasiado la interpretabilidad de los resultados.

En este trabajo, se presenta el modelo de post-procesamiento de Ajuste GEnético Lateral para Descubrimiento de Subgrupos (AGEL-SD) que se incluye dentro de los algoritmos evolutivos de ajuste de sistemas difusos. Tiene como objetivo mejorar el rendimiento de cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa mediante un ajuste basado en la representación de 2-tuplas lingüísticas [16]. En este trabajo se analiza su comportamiento sobre las reglas obtenidas por NMEEF-SD [6]. Tal y como se puede observar en la experimentación llevada a cabo, la utilización de AGEL-SD sobre las reglas obtenidas por el algoritmo NMEEF-SD permite mejorar significativamente la precisión de las reglas difusas de descubrimiento de subgrupos.

2 DESCUBRIMIENTO DE SUBGRUPOS

El concepto de descubrimiento de subgrupos fue introducido inicialmente por Kloesgen [17] y Wrobel [21], y definido más formalmente por Siebes [20] pero empleando el nombre de *Data Surveying* para descubrimiento de subgrupos interesantes. El concepto de descubrimiento de subgrupos se puede definir como [22]:

En descubrimiento de subgrupos, asumimos una población de individuos dada (objetos, clientes, ...) y una propiedad de estos individuos en la que estemos interesados. La tarea del descubrimiento de subgrupos es entonces descubrir los subgrupos de la población que son estadísticamente "más interesantes", es decir, individuos que sean tan grandes como sea posible y tenga una distribución estadística los más atípica posible, con respecto a una propiedad de interés.

El descubrimiento de subgrupos intenta buscar relaciones entre diferentes propiedades o variables de un conjunto con respecto a una variable objetivo. Debido a que el descubrimiento de subgrupos está centrado en la extracción de relaciones con características interesantes, no es necesario obtener relaciones completas sino que suele ser suficiente con relaciones parciales. Estas relaciones son descritas en forma de reglas individuales.

Así, una regla (R), que consiste de una descripción de un subgrupo inducido, puede ser definida formalmente como [13, 18]:

$$R : Cond \rightarrow VarObj$$

donde $VarObj$ es el valor de la variable de interés o variable objetivo para la tarea de descubrimiento de subgrupos (puede aparecer además en la bibliografía específica como *Clase*), y *Cond* es comúnmente una conjunción de funciones (pares atributo-valor) que es capaz de describir una distribución estadística inusual con respecto a la variable objetivo.

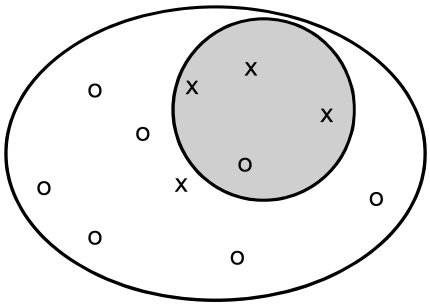


Figura 1: Representación de una regla de descubrimiento de subgrupos con respecto al valor (x) de la variable objetivo

En la figura 1 se representa una regla de descubrimiento de subgrupos, donde se pueden observar dos valores para la variable objetivo ($VarObj = \{x,o\}$). En la figura se representa un subgrupo correspondiente al valor x, donde el subgrupo o regla cubre un alto número de valores de la variable objetivo mediante una descripción sencilla (en la figura representada como un círculo). Sin embargo, el grupo no cubre todos los ejemplos para el valor x de la variable objetivo, incluso no todos los ejemplos son cubiertos correctamente, pero la forma de la función es simple y muy interpretable.

2.1 SISTEMAS DIFUSOS EVOLUTIVOS PARA DESCUBRIMIENTO DE SUBGRUPOS

La lógica difusa es una herramienta que permite representar conceptos de forma lingüística, lo que aporta expresividad y facilidad de comprensión para el usuario. Por tanto, esta herramienta se sitúa en el nivel de representación y se puede utilizar en combinación con otras técnicas. Así, la combinación de lógica difusa y algoritmos evolutivos para la inducción de conocimiento de tipo descriptivo, da lugar a los sistemas difusos evolutivos [14].

Un sistema evolutivo difuso [7] es básicamente un sistema difuso incrementado por un proceso de aprendizaje basado en computación evolutiva. Ésta última incluye a los algoritmos genéticos, programación genética, estrategias evolutivas, entre otros algoritmos evolutivos [10]. Por regla general, este tipo de sistemas consideran la estructura de un modelo en la forma de reglas difusas y se denominan sistemas basados en reglas difusas, los cuáles han demostrado su habilidad con respecto a diferentes problemas como control, modelado, clasificación o minería de datos en un amplio número de aplicaciones. Una de las principales ventajas de los sistemas basados en reglas difusas es la obtención de una representación comprensible del conocimiento extraído siendo además una herramienta útil para el procesamiento de las variables continuas.

Según la optimización que se desee realizar de la base de conocimiento, los sistemas difusos evolutivos se pueden agrupar en torno a dos enfoques [14], aunque es difícil hacer una clara distinción entre ellos, ya que establecer una frontera entre ambos puede llegar a ser tan difícil como definir los conceptos de aprendizaje en si mismos:

- Algoritmos de ajuste genético que aplican un proceso de ajuste para mejorar el rendimiento del sistema basado en reglas difusas sin realizar ninguna modificación sobre la base de reglas, sólo sobre la base de datos y/o el sistema de inferencia.

- Algoritmos de aprendizaje de la base del conocimiento que realizan el aprendizaje sobre la base de reglas, sobre la base de datos, o bien sobre todos los componentes del sistema basado en reglas difusas simultáneamente. Se puede incluir en el aprendizaje un motor de inferencia adaptativo.

2.2 AGEL-SD: Algoritmo de ajuste genético lateral de las etiquetas lingüísticas en descubrimiento de subgrupos

El modelo AGEL-SD es un algoritmo evolutivo de ajuste de sistemas difusos. Tiene como objetivo mejorar el rendimiento de un algoritmo de descubrimiento de subgrupos mediante el enfoque de ajuste basado en la representación de 2-tuplas lingüísticas. Este nuevo modelo puede ser aplicado sobre cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa. La metodología aplicada consiste en la optimización de la definición previa de la base de datos una vez que la base de reglas ha sido obtenida [2]. El ajuste introduce una variación lateral de las funciones de pertenencia que mejora su interacción global, con el principal objetivo de inducir una mejor cooperación entre las reglas [7, 14]. De esta forma, el objetivo del proceso de ajuste no es encontrar funciones de pertenencia específicas de una forma independiente, sino encontrar la mejor configuración global de las funciones de pertenencia y para ello se intenta no afectar excesivamente la interpretabilidad de las mismas.

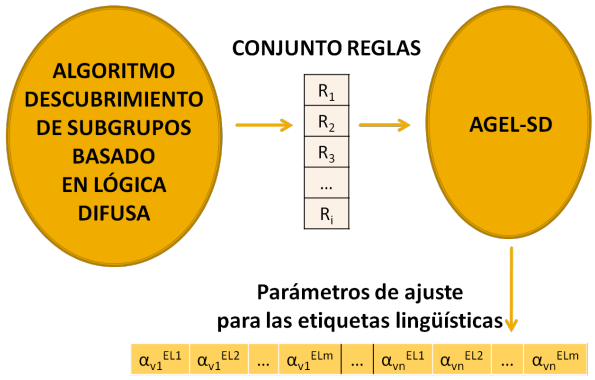


Figura 2: Diagrama de funcionamiento del algoritmo evolutivo para ajuste lateral AGEL-SD

El funcionamiento del algoritmo AGEL-SD se puede observar gráficamente en la Fig. 2. Así, un algoritmo de descubrimiento de subgrupos obtiene un conjunto de reglas (R_1, R_2, ..., R_i), las cuáles están formadas por un conjunto de variables (v_1, v_2, ..., v_n) y cada una de ellas por un número de etiquetas lingüísticas (EL_1, EL_2, ..., EL_m). El objetivo del algoritmo AGEL-SD es la obtención de un individuo formado por un conjunto de ajustes (p) asociados a cada una de las variables continuas codificadas mediante etiquetas lingüísticas obtenidas por el algoritmo de descubrimiento de subgrupos previamente.

En este enfoque de ajuste se emplea un modelo de representación de reglas basado en la representación de 2-tuplas lingüísticas [16]. Esta representación permite un desplazamiento lateral de las etiquetas considerando solo un parámetro (un desplazamiento lateral hacia la izquierda/derecha de la función de pertenencia original). Esta representación implica una simplificación del espacio de búsqueda. Además, este proceso de contextualización de las funciones de pertenencia permite conseguir un mejor grado de cobertura manteniendo las formas originales, obteniendo mejoras en la precisión del modelo sin pérdidas en la interpretabilidad de las etiquetas lingüísticas. En la literatura especializada, la representación de las 2-tuplas se ha utilizado para abordar diferentes problemas como por ejemplo clasificación [2, 1]. Además, dentro de la bibliografía se pueden encontrar diversas propuestas que emplean evolutivos basadas en la representación de 2-tuplas lingüísticas para resolver el problema [3, 12].

El movimiento simbólico de un término lingüístico es un valor dentro del intervalo $[-0.5, 0.5)$, que representa el dominio de una etiqueta cuando se mueve entre sus dos etiquetas laterales, tal y como se puede observar en la figura 3. Siendo S un conjunto de etiquetas para una partición difusa, formalmente se puede observar el par $(s_i, \alpha_i), s_i \in S, \alpha_i \in [-0.5, 0.5)$, donde s_i es la etiqueta y α_i su desplazamiento.

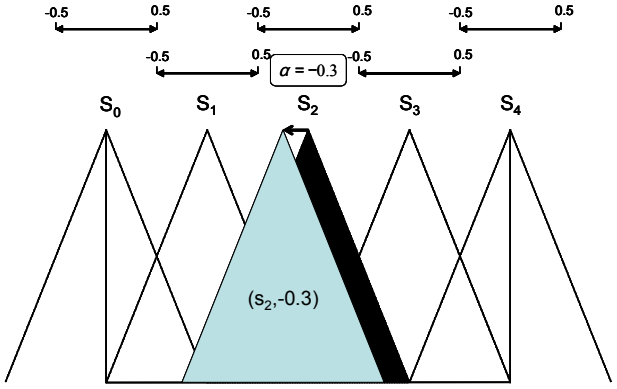


Figura 3: Ajuste lateral de una etiqueta lingüística

El ajuste lateral se realiza mediante un ajuste global de la semántica. De esta forma, se aplica a nivel de la partición lingüística, por tanto el par (valor, etiqueta) toma el mismo valor de ajuste en todas las reglas donde se emplee esta etiqueta. Por ejemplo, para una variable x con una valor de ajuste $(0.35, Alto)$, presentará el mismo valor para aquellas reglas en las que la variable x tenga la etiqueta $Alto$. De esta forma se mantiene una alta interpretabilidad dentro del modelo.

El algoritmo evolutivo AGEL-SD realiza un proceso de búsqueda manteniendo un buen equilibrio entre diversidad y convergencia, constituyendo una buena elección en problemas con espacios de búsqueda complejos. Debido a que es necesario la obtención de una única configuración de ajustes para las etiquetas lingüísticas del conjunto de reglas, se emplea el algoritmo CHC [11] que obtiene el mejor individuo del proceso evolutivo. Entre los operadores de este algoritmo evolutivo destacan:

- *Mecanismo de selección elitista*. El algoritmo parte de

$$f(r_i,C) = \frac{(SENS_{(r_i,C)} * \omega_1) + (ATIP_{(r_i,C)} * \omega_2) + (CNFD_{(r_i,C)} * \omega_3)}{\omega_1 + \omega_2 + \omega_3} \quad (1)$$

una población M y genera sus descendientes. Ambas poblaciones se unen para seleccionar los M mejores individuos a formar parte de la población de la siguiente generación.

- *Mecanismo de prevención de incesto*. Este operador se considera al realizar el cruce. Se combinan dos padres cuando la distancia entre ambos (considerando una métrica adecuada) dividido por dos está por encima de un umbral predeterminado L.

- *Enfoque de re-inicialización*. Se emplea en sustitución del operador de mutación y se presenta en [11]. Una vez que se establece el umbral predeterminado L como la máxima distancia posible entre dos individuos dividido por cuatro, L se decrementa por uno cuando no existen nuevos individuos en la población en una generación. Cuando L está por debajo de cero el algoritmo re-inicializa la población. En este caso, todos los cromosomas son generados de forma aleatoria dentro del intervalo $[$-0.5,0.5$)$. Donde la mejor solución global encontrada se incluye en la población para incrementar la convergencia del algoritmo.

La evaluación del cromosoma se realiza mediante una función de agregación (f) que considera diferentes medidas de calidad de descubrimiento de subgrupos. Las medidas seleccionadas como objetivos son la sensibilidad (conocida en inglés como *Sensitivity*), la atipicidad (conocida como *Unusualness*) y la confianza difusa (conocida como *Confidence*). Para cada una de las medidas de calidad se utiliza un peso asociado. Dependiendo de los pesos utilizados se le da una mayor importancia a unas características u otras del descubrimiento de subgrupos. En esta experimentación se ha utilizado esta combinación de pesos: ω_1 =0.2, ω_2 =0.4 y ω_3 =0.4, para las medidas de sensibilidad (SENS), atipicidad (ATIP) y confianza difusa (CNFD). La función de adaptación de los cromosomas se puede observar en la ecuación 1, y las medidas de calidad utilizadas en [15].

Puesto que AGEL-SD es un algoritmo de ajuste global de la semántica en la base de reglas, hay que considerar que para cada una de las reglas que tenemos dentro de la base de reglas se obtiene un valor de la función de adaptación. De ahí, que el valor de la función de adaptación (F) para cada individuo sea la media aritmética de la adaptación en el conjunto de reglas extraídas (R), tal y como se puede observar en la ecuación 2.

$$F(R,C) = \frac{\sum_{i=1}^{|R|} f(r_i,C)}{i} \quad (2)$$

3 ESTUDIO EXPERIMENTAL

El objetivo de este estudio experimental es analizar el rendimiento del uso del ajuste basado en el enfoque lingüístico de las 2-tuplas para un algoritmo de descubrimiento subgrupos, en concreto, se utiliza el algoritmo NMEEF-SD [6] en conjuntos de datos con variables continuas. La experimentación se ha realizado con conjuntos de datos del repositorio KEEL [4, 5].

3.1 MARCO EXPERIMENTAL

Los conjuntos de datos utilizados en la experimentación se muestran en la tabla 1, en la que se especifica para cada base de datos, el número de variables discretas (n_{vD}), número de variables continuas (n_{vC}) y el número de ejemplos (n_s).

Tabla 1: Propiedades del conjunto de datos utilizados del repositorio KEEL

Nombre	$\boldsymbol{n_{vD}}$	$\boldsymbol{n_{vC}}$	$\boldsymbol{n_s}$
Balance	0	4	625
Haberman	0	3	306
Iris	0	4	150
Phoneme	0	5	5404
Pima	0	8	768

Los parámetros utilizados tanto por el algoritmo NMEEF-SD, como por el algoritmo de ajuste AGEL-SD, se pueden observar en la tabla 2. Debido a la naturaleza no determinista del algoritmo NMEEF-SD, el algoritmo se ejecuta cinco veces para cada conjunto de datos con una 10-validación cruzada para cada conjunto de datos. Los resultados finales mostrados son por tanto, la media de los resultados obtenidos para cada conjunto de datos en las diferentes ejecuciones; es decir, la media de los 50 experimentos por conjunto.

Tabla 2: Especificación de los parámetros para los algoritmos evolutivos empleados en la experimentación

Algoritmo	***Parámetros***
NMEEF-SD	Tamaño de la población=50, Evaluaciones=10000, Probabilidad de cruce=0.6, Probabilidad de mutación=0.1, Re-inicialización basada en cobertura (50 % de sesgo) y Confianza mínima=0.6.
AGEL-SD	Tamaño de la población=50, Evaluaciones=10000, α =0.5

Tabla 3: Resultados obtenidos por el algoritmo NMEEF-SD y el algoritmo NMEEF-SD + AGEL-SD

Et	*Algoritmo*	*RELE*	*ATIP*	*PREC*	*SENS*	*CNFN*	*CNFD*	
3	NMEEF-SD	2.310	**0.077**	0.712	0.862	0.757	0.778	
	NMEEF-SD + AGEL-SD	**5.276**	0.071	**0.734**	**0.927**	**0.779**	**0.802**	$\star$
5	NMEEF-SD	4.366	**0.075**	0.736	0.785	0.799	0.809	
	NMEEF-SD + AGEL-SD	**5.182**	0.073	**0.743**	**0.806**	**0.806**	**0.820**	$\star$

Para el estudio del ajuste genético lateral mediante representación lingüística de 2-tuplas en el algoritmo NMEEF-SD, se han empleado dos niveles de granularidad: 3 y 5 etiquetas lingüísticas para mostrar cuál es la mejora del ajuste genético lateral con respecto a distintas granularidades en un algoritmo de descubrimiento de subgrupos.

3.2 RESULTADOS OBTENIDOS

A continuación, se realiza un estudio completo sobre diferentes medidas de calidad para el algoritmo de descubrimiento de subgrupos NMEEF-SD y el mismo algoritmo con ajuste lateral (NMEEF-SD + AGEL-SD). Para cada medida de calidad se muestra la media obtenida por los algoritmos en los conjuntos de datos mencionados anteriormente, es decir, las medidas de calidad mostradas en la tabla 3 son la media para el conjunto de reglas, donde se pueden observar las medidas de relevancia (conocidad en inglés como *Significance* y con las siglas *RELE*), atipicidad (*ATIP*), precisión (*PREC*) [19], sensibilidad (*SENS*), confianza nítida (*CNFN*) y confianza difusa (*CNFD*). Además, se indican tanto el número de etiquetas lingüísticas (*Et*) como el nombre del *Algoritmo*. En [15] se puede observar una revisión sobre las diferentes medidas de calidad para descubrimiento de subgrupos.

La media de los resultados obtenidos por los algoritmos se puede observar en la tabla 3, donde el algoritmo NMEF-SD + AGEL-SD obtiene los valores medios más altos en la mayoría de las medidas de calidad, lo que sugiere un buen comportamiento de este algoritmo. Además, un aspecto importante a considerar, es la mejora obtenida por el algoritmo NMEEF-SD + AGEL-SD con respecto a la relación entre sensibilidad y confianza, es decir, la representación de 2-tuplas lingüísticas es capaz de incrementar el valor de ambas medidas de calidad a la vez, cuando lo normal es que la mejora en una de ellas degrade la otra.

Los principales resultados aportados por AGEL-SD se resumen a continuación:

- No se modifican sensiblemente el nivel de interpretabilidad, pues el algoritmo trabaja con el mismo conjunto de reglas sin realizar modificación sobre ellas a nivel individual, pues este enfoque es un enfoque de ajuste global.
- Los resultados obtenidos por NMEEF-SD + AGEL-SD sobre los obtenidos por NMEEF-SD, proporcionan una mayor precisión. De esta forma, los expertos pueden obtener descripciones de los problemas de una forma más exacta.
- Los niveles de sensibilidad también aumentan, por lo que la relación sensibilidad-confianza mejora sustancialmente, pues permite una mejora en la precisión del modelo sin degradar la sensibilidad del mismo.
- En el resto de medidas de calidad, salvo en atipicidad, también se observa una mejora de los valores tras la aplicación del algoritmo AGEL-SD.
- Se permite la utilización de diversas medidas de calidad dentro de la función de adaptación del algoritmo para así ajustarse mejor a las necesidades del experto.

4 CONCLUSIONES

La propuesta consiste en un algoritmo evolutivo de ajuste genético lateral mediante la representación de 2-tuplas lingüísticas, AGEL-SD. Este algoritmo se puede aplicar a cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa, ya que realiza el ajuste sobre las etiquetas lingüísticas.

El modelo presentado está basado en un enfoque genético mono-objetivo, y realiza un ajuste global de la semántica, aplicado a nivel de la partición lingüística, es decir el par (valor, etiqueta) toma el mismo valor de ajuste en todas las reglas donde se emplee esta etiqueta.

Para demostrar la calidad de este nuevo enfoque, se ha presentado un estudio de AGEL-SD aplicado sobre los resultados del algoritmo NMEEF-SD, que es uno de los algoritmos más representativos del descubrimiento de subgrupos. En este estudio se demuestra la mejora en la calidad de los resultados con respecto a componentes de precisión, mejorando además en los resultados de sensibilidad.

Agradecimientos

Esta trabajo ha sido subvencionado por el Ministerio de Educación, Política Social y Deportes bajo el proyecto TIN-2008-06681-C06-02, Fondos FEDER, y el por el Plan de Investigación de Andalucía bajo el proyecto TIC-3928, Fondos FEDER.

Referencias

[1] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto, and F. Herrera. Rule base reduction and genetic tuning of fuzzy sys-

tems based on the linguistic 3-tuples representation. *Soft Computing*, 11(5):401–419, 2007.

[2] R. Alcalá, J. Alcalá-Fdez, and F. Herrera. A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. *IEEE Transactions on Fuzzy Systems*, 15(4):616–635, 2007.

[3] J. Alcalá-Fdez, R. Alcalá, M. J. Gacto, and F. Herrera. Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms. *Fuzzy Sets and Systems*, 160(7):905–921, 2009.

[4] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. *Journal of Multiple-Valued Logic and Soft Computing*, 17(2-3):255–287, 2011.

[5] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, and F. Herrera. KEEL: A Software Tool to Assess Evolutionary Algorithms for Data Mining Problems. *Soft Computing*, 13(3):307–318, 2009.

[6] C. J. Carmona, P. González, M. J. del Jesus, and F. Herrera. NMEEF-SD: Non-dominated Multi-objective Evolutionary algorithm for Extracting Fuzzy rules in Subgroup Discovery. *IEEE Transactions on Fuzzy Systems*, 18(5):958–970, 2010.

[7] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*. World Scientific, 2001.

[8] M. J. del Jesus, P. González, and F. Herrera. Multiobjective Genetic Algorithm for Extracting Subgroup Discovery Fuzzy Rules. In *Proceedings of the IEEE Symposium on Computational Intelligence in Multicriteria Decision Making*, pages 50–57. IEEE Press, 2007.

[9] M. J. del Jesus, P. González, F. Herrera, and M. Mesonero. Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A case study in marketing. *IEEE Transactions on Fuzzy Systems*, 15(4):578–592, 2007.

[10] A. E. Eiben and J. E. Smith. *Introduction to evolutionary computation*. Springer, 2003.

[11] L. J. Eshelman and J. D. Schaffer. Preventing premature convergence in genetic algorithms by preventing incest. In *Proceedings of the 4th International Conference on Genetic Algorithms*, pages 115–122, 1991.

[12] A. Fernández, M. J. del Jesus, and F. Herrera. On the 2-Tuples Based Genetic Tuning Performance for Fuzzy Rule Based Classification Systems in Imbalanced Data-SetsInformation Sciences 180:8 (2010) 1268-1291. *Information Sciences*, 180(8):1268–1291, 2010.

[13] D. Gamberger and N. Lavrac. Expert-Guided Subgroup Discovery: Methodology and Application. *Journal Artificial Intelligence Research*, 17:501–527, 2002.

[14] F. Herrera. Genetic fuzzy systems: taxomony, current research trends and prospects. *Evolutionary Intelligence*, 1:27–46, 2008.

[15] F. Herrera, C. J. Carmona, P. González, and M. J. del Jesus. An overview on Subgroup Discovery: Foundations and Applications. *Knowledge and Information Systems*, In press, 2011.

[16] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[17] W. Kloesgen. Explora: A Multipattern and Multistrategy Discovery Assistant. In *Advances in Knowledge Discovery and Data Mining*, pages 249–271. American Association for Artificial Intelligence, 1996.

[18] N. Lavrac, B. Cestnik, D. Gamberger, and P. A. Flach. Decision Support Through Subgroup Discovery: Three Case Studies and the Lessons Learned. *Machine Learning*, 57(1-2):115–143, 2004.

[19] C. Romero, P. González, S. Ventura, M. J. del Jesus, and F. Herrera. Evolutionary algorithm for subgroup discovery in e-learning: A practical application using Moodle data. *Expert Systems with Applications*, 36:1632–1644, 2009.

[20] A. Siebes. Data Surveying: Foundations of an Inductive Query Language. In *Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining*, pages 269–274. AAAI Press, 1995.

[21] S. Wrobel. An Algorithm for Multi-relational Discovery of Subgroups. In *Proceedings of the 1st European Symposium on Principles of Data Mining and Knowledge Discovery*, volume 1263 of *LNAI*, pages 78–87. Springer, 1997.

[22] S. Wrobel. *Inductive logic programming for knowledge discovery in databases*, chapter Relational Data Mining, pages 74–101. Springer, 2001.

[23] L. A. Zadeh. Soft Computing and Fuzzy Logic. *IEEE Software*, 11(6):48–56, 1994.

UN ESTUDIO EXPERIMENTAL DEL USO DE DOMINIOS CON INTENSIFICACIONES

David García, **Antonio González**, **Enrique Leyva**, **Raúl Pérez**

Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, 18071-Granada (Spain),
{dgarcia,a.gonzalez,eleyvam,raul_perez}@decsai.ugr.es

Resumen

Desarrollar algoritmos de aprendizaje de reglas difusas que obtengan un buen balance entre precisión e interpretabilidad es uno de los campos de investigación dentro del aprendizaje de sistemas difusos. En este trabajo se presenta una propuesta que permite a un algoritmo de clasificación elegir entre cuatro posibles configuraciones para construir el dominio asociado a sus atributos. Estas cuatro configuraciones incluyen la distribución uniforme y tres distribuciones que intensifican la capacidad de discriminación de los atributos en función de la distribución de los datos sobre dicho atributo. Una medida de información determina la configuración más apropiada para cada atributo. Los resultados experimentales muestran que la propuesta presentada mantiene la capacidad de predicción y reduce el número de reglas.

Palabras Clave: Aprendizaje inductivo, Reglas difusas, Balance precisión/interpretabilidad.

1. INTRODUCCIÓN

Uno de los puntos de debate entre los investigadores del campo del aprendizaje de reglas difusas es determinar el grado de interpretabilidad del conocimiento que se extrae, así como la relación entre la interpretabilidad y la capacidad de predicción de dicho conocimiento. En la bibliografía podemos encontrar multitud de trabajos que tratan sobre este problema y que presentan diversas propuestas para mantener un buen balance entre precisión e interpretabilidad. Algunos de ellos, como [12], establecen posibles configuraciones alternativas (pero fijadas a priori) sobre los dominios, que se asignan a los atributos dejando que el algoritmo de aprendizaje decida cuales son las apropiadas para describir el conocimiento. Otros definen modificadores semánticos [8], que son aprendidos por el algoritmo e incluidos en las reglas, y permiten alterar ligeramente la semántica de las etiquetas. Algunos otros definen métodos de ajuste [3, 11], concebidos como un proceso posterior al aprendizaje, que modifican la semántica de las etiquetas una vez aprendida la base de reglas.

En la mayoría de las situaciones y fundamentalmente cuando no está disponible la información de un experto, se suele asociar a las variables un dominio con distribución uniforme de etiquetas (Figura 1). Esta distribución mantiene ciertas propiedades que, consensuado por la mayoría de los investigadores en este campo, dotan al conocimiento que se extrae de un cierto grado de interpretabilidad, aunque podemos entender que esta distribución juega el papel de una distribución de posibilidad a priori sobre un fenomeno no conocido.

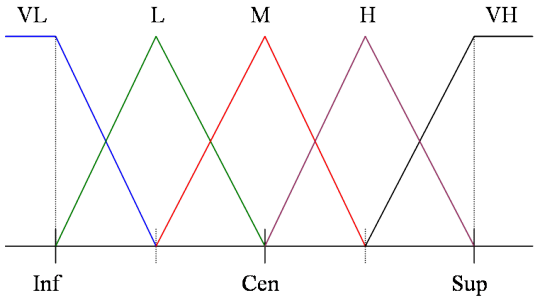


Figura 1: Distribución uniforme con cinco etiquetas.

Sin embargo, es posible realizar una distribución semejante a la uniforme, teniendo en cuenta la información que nos aportan los ejemplos en el caso del aprendizaje inductivo. Siguiendo esta idea de aprovechar la información de los datos proponemos un método de asignación de dominios a variables que elige entre cuatro configuraciones posibles, la uniforme y tres que aumentan la sensibilidad sobre la zona inferior, central o superior del intervalo de definición. Este proceso de asignación de dominios es previo al proceso de aprendizaje y es guiado por una medida de información.

Los resultados experimentales muestran que este método

de asignación permite mantener la capacidad de predicción y reducir el número de reglas del conocimiento que se extrae.

En la siguiente sección se define la propuesta para la asociación de los dominios difusos a los atributos en un problema de clasificación. En la sección 3 se describe el algoritmo de aprendizaje NSLV-F que será utilizado en la parte experimental que se presenta en la sección 4. Por último, en la sección 5 se muestran las conclusiones de este trabajo.

2. PATRONES DE DISCRETIZACIÓN DE DOMINIOS DIFUSOS

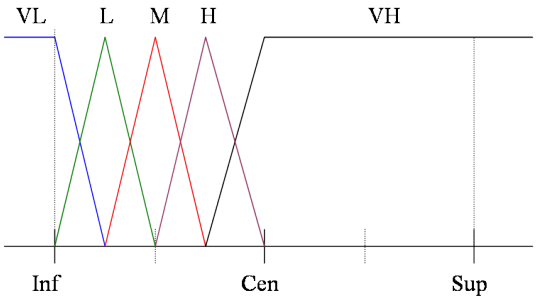
El objetivo del trabajo es presentar un modelo de distribución de etiquetas difusas para definir el dominio de un atributo que mantenga las propiedades de interpretabilidad de la distribución uniforme, pero que permita un cierto ajuste sobre la distribución de los datos sobre dicho atributo, para mejorar su capacidad de discriminación y por tanto, mejorar su capacidad de predicción.

En la propuesta se definen tres configuraciones fijas que se presentan como alternativa a la distribución uniforme y que tratan de reflejar el desplazamiento de la masa significativa de los datos sobre el dominio de definición. Estas tres configuraciones reflejan la necesidad de dotar de mayor sensibilidad (y por tanto, de mayor número de etiquetas) en ciertas zonas del intervalo de definición de las variables. En la Figura 2 se muestran estas configuraciones que tratan de dotar de mayor sensibilidad en la parte inferior, en la parte central y en la parte superior del intervalo de definición, respectivamente.

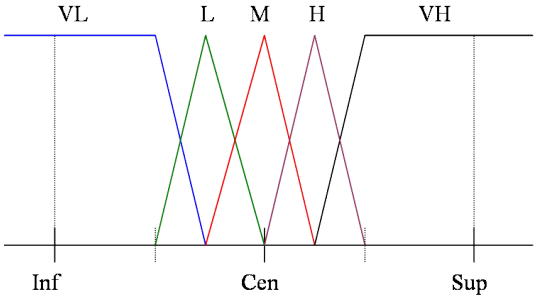
En estas distribuciones, al igual que en la uniforme (Figura 1), todas las etiquetas se cortan a nivel **0.5** y se mantiene el orden de las mismas. La construcción es simple, el soporte de las etiquetas centrales del dominio (en las tres distribuciones) es justo la mitad del soporte de las etiquetas centrales en la distribución uniforme y por consiguiente, definen justo la mitad del intervalo de definición. Las etiquetas de los extremos del dominio amplian o disminuyen su soporte para completar el intervalo de definición.

Así, lo que proponemos es permitir que un algoritmo de aprendizaje pueda elegir a priori entre la distribución uniforme o alguna de estas distribuciones presentadas para definir el dominio asociado a las variables involucradas en el problema de aprendizaje. La idea sería seleccionar aquella distribución que dota a la variable de mayor capacidad de discriminación. Para conocer la capacidad de discriminación de cada una de las distribuciones consideradas utilizamos una medida de información. Esta medida se detalla en [13, 16]:

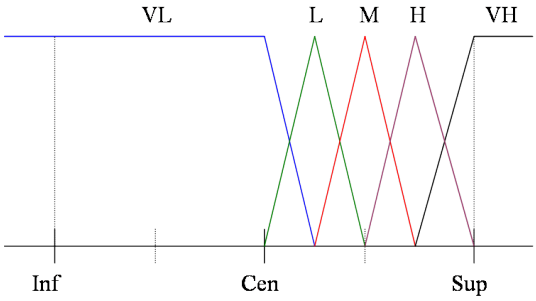
$$\rho(X,Y) = \frac{I(X,Y)}{H(X,Y)} \qquad (1)$$



(a) Intesificación inferior



(b) Intensificación central



(c) Intensificación superior

Figura 2: Ejemplo de las distintas intensificaciones sobre un domino compuesto por cinco etiquetas difusas.

donde

$$I(X,Y) = \sum_x \sum_y -p(x,y)\ log_2(\frac{p(x)p(y)}{p(x,y)}) \qquad (2)$$

y H(X,Y) es la entropía de Shannon sobre dos variables definida como:

$$H(X,Y) = \sum_x \sum_y -p(x,y)\ log_2\ p(x,y). \qquad (3)$$

La medida ρ estima la dependencia entre las variables X e Y de la siguiente forma: valores de $\rho(X,Y)$ cercanos a cero indican un alto grado de independencia entre las dos variables, mientras que valores cercanos a uno indican un alto grado de dependencia funcional entre ellas.

Así, el proceso de elección de la distribución a asociar a los atributos del problema consiste en asignar a cada atributo

X_i la distribución entre las cuatro presentadas (Uniforme o con intensificaciones Inferior, Central o Superior) que le ofrezca el mayor valor de $\rho(X_i, Y)$, siendo Y la variable de clasificación.

3. EL ALGORITMO DE APRENDIZAJE NSLV-F

NSLV-F [5] es un algoritmo de aprendizaje de reglas difusas que usa la estrategia de recubrimiento secuencial [15] y esta basado en el algoritmo NSLV [10]. La estrategia de recubrimiento secuencial descompone el problema de aprender un conjunto de reglas a un problema más simple que consiste en aprender una regla cada vez hasta completar el conjunto de reglas. El proceso fundamental dentro de esta estrategia es un algoritmo de búsqueda que es el encargado en encontrar en cada iteración la siguiente regla a incluir. En NSLV-F es un algoritmo genético (AG) el que realiza esta tarea. El AG toma como entrada la variable de clasificación, el conjunto completo de atributos y un conjunto de ejemplos y ofrece como salida la mejor regla encontrada que satisface los criterios de búsqueda. En NSLV-F, el objetivo es encontrar reglas simples que satisfagan los siguientes criterios:

> **O1**: La regla describe correctamente un gran número de ejemplos y describe pocos incorrectamente.
>
> **O2**: La regla usa un número reducido de variables en su antecedente.
>
> **O3**: El antecedente de la regla tiene asignaciones comprensibles entre variables y valores.

El problema de seleccionar la mejor regla es un problema multiobjetivo, ya que la solución debe satisfacer tres criterios distintos. Para resolver este problema, NSLV-F usa una función de evaluación lexicográfica (LEF), es decir, considera que los objetivos no tienen el mismo nivel de importancia, y los ordena usando un criterio de preferencia. En NSLV-F, ese orden coincide exactamente con el orden en el cual los objetivos se han descrito anteriormente.

Así, el criterio más importante (**O1**) consiste en encontrar la regla que cubre más ejemplos positivos y menos ejemplos negativos. Este criterio, en los algoritmos de aprendizaje clásico está relacionado con dos condiciones bien conocidas denominadas condiciones de consistencia y completitud [14]. Estas dos condiciones fueron extendidas para trabajar con reglas difusas en [6]. Estas extensiones establecen el grado en el cual una regla satisface cada una de estas condiciones. En NSLV-F el primer criterio de valoración de una regla es el producto de los grados de satisfación de las condiciones extendidas de consistencia y completitud.

Los otros dos objetivos (**O2** y **O3**) están relacionados con la simplicidad de la regla medida desde dos puntos de vista. Mientras el segundo objetivo da preferencia a reglas con el menor número de variables en el antecedente de la regla, el tercero da preferencia a asignaciones comprensibles en el antecedente de la regla.

En NSLV-F, cada individuo en la población de búsqueda representa una regla DNF completa, así que cada variable puede tomar como valor una disyunción de términos linguísticos de su dominio. A diferencia de NSLV, NSLV-F permite incluir nuevos atributos construidos a partir de los atributos iniciales del problema. El proceso de selección del conjunto de los nuevos atributos es semejante al utilizado por los métodos filtro de los algoritmos de selección de variables. En dicho proceso filtro, se asocia una unidad de medida a las variables iniciales y toma un conjunto de funciones para combinar estas variables. Cada una de las funciones del conjunto establece restricciones sobre qué variables se pueden combinar a través de ella. Una medida de información se utiliza para determinar la aportación de la combinación en relación a la que aportan los atributos iniciales. El conjunto de combinaciones que aportan al menos la información de los atributos iniciales se seleccionan y se almacenan en un *catálogo de funciones* que incluye las nuevas variables a considerar en el proceso de aprendizaje.

El algoritmo genénico representa cada regla DNF mediante cuatro subcromosomas o niveles (Fig. 3):

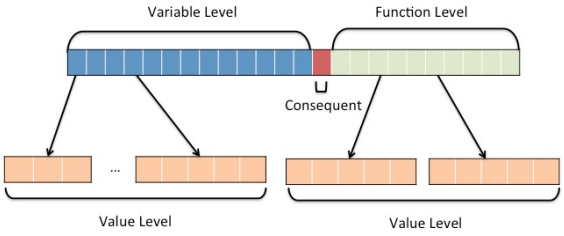


Figura 3: Estructura de un individuo en la población genética

- *Nivel de consecuente*. Codifica el valor de la variable de clasificación de la regla y está compuesto por un único gen representado mediante un valor entero.

  En este nivel, se considera como único operador genético la mutación uniforme. La población inicial se genera con el mismo número de individuos para cada una de las clases que están involucradas en el problema.

- *Nivel de variable*. Utiliza codificación real y contiene un gen por cada uno de los atributos iniciales. Cada gen representa el grado de relevancia inicial de cada atributo con respecto a la clase. Además, contiene un gen especial que se interpreta como el umbral de activación, de manera que los genes asociados a los atributos que tengan un valor menor a este umbral no son

considerados en el antecedente de la regla. Por consiguiente, el uso de este nivel permite desarrollar un proceso de selección de variables inmerso [9].

El nivel usa dos operadores genéticos, el cruce sobre dos puntos y la mutación uniforme sobre codificación real. En la primera población, el gen asociado a cada atributo toma como valor la relevancia inicial de dicho atributo sobre el conjunto de entrenamiento condicionada al valor de la clase. El valor umbral se genera aleatoriamente entre la medida de información mínima y máxima.

- *Nivel de valor*. Esta compuesto de la secuencia de asignaciones de los atributos iniciales, donde cada asignación variable/valor se representa mediante una secuencia binaria (donde el uno representa la presencia del valor en la asignación y el cero la no presencia). El nivel completo está compuesto por la concatenación de las cadenas binarias de todos los atributos. La asignación a una determinada variable aparecerá en la descripción de la regla si su valor asociado en el nivel de variable permite considerar a esta variable como relevante para la regla.

  El nivel usa codificación binaria y dos operadores genéticos, el cruce sobre dos puntos y la mutación uniforme. La generación de la población inicial de este nivel es la siguiente: conocido el valor del *Nivel de consecuente* (ya que dicho nivel se genera inicialmente antes que este), se toma un ejemplo del conjunto de entrenamiento para dicha clase. A partir de él, se obtiene la regla más específica que mejor lo cubre. Se toma el antecedente de dicha regla para codificar al individuo. Este proceso se repite para todos los individuos de la población.

- *Nivel de función*. Usa codificación entera y juega un papel semejante al *Nivel de variable* pero aplicado en este caso a los nuevos atributos, es decir, a los atributos que contiene *el catálogo de funciones*. Por tanto, determina qué nuevos atributos formarán parte del antecedente de la regla teniendo en cuenta que el valor cero indica que no hay variable a considerar y que un valor distinto de cero hace referencia al nuevo atributo que ocupa dicha posición dentro del catálogo.

  Inicialmente, se considera un número máximo de nuevos atributos. Este nivel usa dos operadores genéticos, el cruce sobre dos puntos y la mutación uniforme. En la población inicial no se considera ninguna variable nueva, es decir, todo el nivel se inicializa a cero.

NSLV-F utiliza un algoritmo genético estacionario y divide la población en subpoblaciones donde cada subpoblación está asociada a una de las clases que se tienen que aprender y contiene los mejores individuos para dicha clase. Usa el proceso clásico de selección de un algoritmo genético estacionario pero con una modificación: los individuos que son eliminados de la población se seleccionan entre los peores individuos de las subpoblaciones que mantienen un número de individuos superior al número mínimo de individuos que ha de mantenerse durante toda la evolución. Este proceso se realiza para mantener siempre individuos que representan a todas las clases que se deben aprender y asegurar que dicho número es suficiente para permitir a dicha subpoblación evolucionar hacia las mejores soluciones. El objetivo de las subpoblaciones es mantener alta diversidad en la población genética permitiendo controlar la convergencia prematura del proceso de búsqueda.

4. ESTUDIO EXPERIMENTAL

En esta sección comparamos los resultados obtenidos con NSLV-F considerando una discretización uniforme con cinco etiquetas sobre todos sus atributos (con excepción de los nominales) con los resultados de este mismo algoritmo pero considerando para la discretización de los dominios de sus atributos el que aporta mayor información entre la distribución uniforme y las tres intensificaciones propuestas. A este segundo algoritmo lo llamaremos NSLV-F2, y al igual que NSLV-F utilizará cinco etiquetas.

Para este estudio experimental hemos considerado un total de 28 bases de datos, 27 son del UCI Repository of Machine Learning Databases and Domain Theories [1] y la base de datos Tao [2]. En la Tabla 1 se describen estas bases de datos donde cada fila representa el nombre y la abreviatura de la base de datos, el número de ejemplos (#E), variables (#V), variables continuas (#C), variables nominales (#N), clases (#Cl) y si tiene o no valores perdidos (M).

Tabla 1: Bases de Datos utilizadas en la experimentación.

Database	#E	#V	#C	#N	#Cl	M
Annealing (ann)	898	38	6	32	5	Si
Automobile (aut)	205	25	15	10	6	Si
Balance (bal)	625	11	4	7	3	No
Bupa (bpa)	345	6	6	0	2	No
Pittsburgh Bridges (brd)	106	11	4	7	7	No
Car Evaluation (car)	1728	6	0	6	2	No
Horse Colic (col)	368	22	7	15	2	Si
Contraceptive Method (cmc)	1473	9	2	7	3	No
Credit Approval (crx)	690	15	6	9	2	Si
Dermatology (drm)	366	34	33	1	7	Si
Glass (gls)	214	9	9	0	6	No
Heart-C (h-c)	303	13	6	7	2	Si
Heart Disease (h-h)	296	13	6	7	2	Si
Hepatitis (hpt)	155	19	6	13	2	Si
Heart-S (h-s)	270	13	6	7	2	Si
Iris (irs)	150	4	4	0	3	No
Pima (pim)	768	8	7	1	2	No
Sonar (son)	208	60	60	0	2	No
Tao (tao)	888	2	2	0	2	No
Thyroid (thy)	215	5	5	0	5	No
Blood Transfusion (trf)	748	4	4	0	2	No
Vehicle Silhouettes (veh)	946	18	18	0	4	No
Congresional Voting (vot)	435	16	0	16	2	Si
Wisconsing Original (wbdc)	699	10	10	0	2	Si
Wisconsing Diagnostic (wdbc)	569	30	30	0	2	No
Wine (wne)	178	13	13	0	3	No
Wisconsing Prognostic (wpbc)	198	33	33	0	2	Si
Zoo (zoo)	101	17	0	17	7	No

Ambos algoritmos utilizan el mismo tratamiento sobre los valores perdidos [7] y usan el mismo conjunto de funciones

para generar los nuevos atributos. Las funciones consideradas son las siguientes:

- SUM (X_i,X_j), definida como $X_i + X_j$

  Restricción: $Unit(X_i) \neq 0$ y $Unit(X_j) \neq 0$ and $Unit(X_i) = Unit(X_j)$

- SUBT (X_i,X_j), definida como $X_i - X_j$.

  Restricción: $Unit(X_i) \neq 0$ y $Unit(X_j) \neq 0$ and $Unit(X_i) = Unit(X_j)$

- PRODUCT (X_i,X_j), definida como X_i * X_j

  Restricción: $Unit(X_i) \neq 0$ y $Unit(X_j) \neq 0$

- DIV (X_i, X_j), definida como $\frac{X_i}{X_j}$

  Restricción: $Unit(X_i) \neq 0$ y $Unit(X_j) \neq 0$

indicando $Unit(X)$ la unidad de medida asociada a la variable X.

Las funciones suma y resta sólo combinan variables con la misma unidad de medida, mientras que el producto y la división permite combinar variables con diferentes unidades de medida.

Los resultados han sido obtenidos usando validación cruzada de 10 subconjuntos usando las mismas particiones para ambos algoritmos. Además, el número de individuos de las poblaciones genéticas es de 100, el cardinal máximo del catálogo de funciones es de 30. La experimentación se ha enfocado a estudiar dos parámetros: la capacidad de predicción sobre el conjunto de test (Test) y el número de reglas (Reglas). Para evaluar los resultados hemos utilizado el test de Bonferroni-Dunn [4]. Este test establece que dos clasificadores tienen diferencias, si la diferencia entre ambos excede una *distancia crítica*. Esta distancia crítica depende del número de bases de datos, del número de clasificadores a comparar y un valor crítico q_α, con error α. En este caso, tomamos $\alpha = 0{,}05$ y $q_\alpha = 1{,}960$. La distancia crítica para 28 bases de datos y dos clasificadores es de **0.37**.

Los resultados obtenidos se muestran en la Tabla 2. La primera columna indica las bases de datos consideradas. La segunda y la cuarta muestran los resultados obtenidos sobre el acierto medio en los conjuntos de test y el número medio de reglas para el algorithmo NSLV-F. Las columnas tercera y quinta los resultados de los parámetros anteriores para el algoritmo NSLV-F2. Las tres últimas filas muestran los valores estimados para el test de Bonferroni-Dunn así como el valor medio de cada columna.

El estudio muestra que NSLV-F2 presenta una mejor capacidad de predicción sobre el conjunto de test (algo más de un punto porcentual en media) aunque el test estadístico refleja que no hay diferencias significativas entre ambos algoritmos. En relación al otro parámetro estudiado, si que hay diferencias significativas en favor de NSLV-F2, obteniendose una reducción en media de más de seis reglas.

Tabla 2: Resultados en Test y Reglas obtenidos por NSLV-F y NSLV-F2

	Test		Reglas	
	NSLV-F	NSLV-F2	NSLV-F	NSLV-F2
ann	97.9(2)	98.2(1)	9.9(2)	6.5(1)
aut	77.5(1.5)	77.5(1.5)	23.9(2)	20.7(1)
bal	63.6(2)	86.9(1)	72(2)	19.7(1)
bpa	66.9(2)	69.5(1)	13.7(1)	15.6(2)
brd	54.8(2)	65.1(1)	22.1(2)	20.3(1)
car	98.1(1)	91.4(2)	34.6(2)	19(1)
cmc	48(2)	48.4(1)	145.3(2)	126(1)
col	81.2(2)	81.4(1)	11.3(2)	7.6(1)
crx	82(2)	84.7(1)	28.2(2)	12.6(1)
drm	92.6(2)	94.7(1)	14.4(2)	10.3(1)
gls	70.1(1)	67.8(2)	20.5(1)	20.6(2)
h-c	77.1(2)	77.5(1)	23.5(2)	12.2(1)
h-h	63.2(1)	59.4(2)	35.8(2)	33.1(1)
hpt	78(2)	83.2(1)	11.4(2)	7.2(1)
h-s	75.9(2)	78.1(1)	21.2(2)	11.5(1)
irs	96.6(1)	96(2)	5.1(2)	4.8(1)
pim	74.4(1)	73.9(2)	26.7(1)	27.7(2)
son	78.3(2)	78.7(1)	14.6(2)	9.3(1)
tao	92.1(1)	89.2(2)	7.1(1)	8.3(2)
thy	94.9(1)	93.5(2)	6.2(2)	5.7(1)
trf	76.4(1)	75.1(2)	4.7(1)	4.9(2)
veh	73.5(2)	74.5(1)	59.2(2)	45.4(1)
vot	94.4(2)	95.8(1)	6.6(2)	6.4(1)
wbcd	95.2(2)	95.4(1)	12(2)	7.8(1)
wdbc	95.2(1)	94.9(2)	8.8(2)	8.1(1)
wne	96(1)	95.4(2)	6.5(2)	4.8(1)
wpbc	72.3(1)	70.5(2)	11.4(2)	7.2(1)
zoo	94.1(2)	97(1)	8.4(2)	7.8(1)
Rnk	1.58929	1.41071	1.82143	1.17857
Pos	2	1	2	1
Media	80.725	81.9179	23.7536	17.5393

Estos resultados ponen de manifiesto que si entendemos que la construcción de los dominios construidos a partir de las intensificaciones descritas en esta propuesta se consideran interpretables, las bases de conocimiento incluyen un número menor de reglas, lo que favorece aún más la interpretabilidad del conocimiento, manteniendose la capacidad predictiva de dicho conocimiento.

5. CONCLUSIONES

En este trabajo se ha presentado una propuesta para asociar a atributos continuos dominios difusos guiado por una medida de información. El sistema elige entre cuatro posibles configuraciones de discretización: la distribución uniforme y tres intensificaciones (izquierda, centro y derecha) para adaptarse a la distribución de los datos y tratando de mantener la interpretabilidad de la discretización uniforme.

Los resultados experimentales muestran que esta flexibilidad restringuida en la elección de los dominios permite mantener la capacidad de predicción del conocimiento extraido y reducir la complejidad de dicho conocimiento en número de reglas.

Agradecimientos

Este trabajo está parcialmente financiado por el proyecto de investigación P09-TIC-04813.

Referencias

[1] A. Asuncion, D.J. Newman, "UCI Machine Learning Repository" Irvine, CA: University of California, School of Information and Computer Science, 2007. Available: http://www.ics.uci.edu/ mlearn/MLRepository.html.

[2] E. Bernardó-Mansilla, X. Llorá, J. Garrel (2002) XCS and Gale: "A comparative study of two learning classifier", *In: P. Lanzi, W. Stolzman, S. Wilson (eds) Advances in Learning Classifier Systems, LNCS* **2321**, Springer, Heidelberg, pp. 115–132.

[3] J. Casillas, O. Cordon, M.J. del Jesus, F. Herrera, "Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction", *Fuzzy Systems, IEEE Transactions on*, **13(1)** (2005) 13- 29.

[4] O. Dunn,"Multiple Comparisons among means", *Journal of the American Stadistical Association*, **56 (293)** (1961) 52-64.

[5] D. García, A. González, R. Pérez, "A two-step approach of feature construction for a genetic learning algorithm", *IEEE International Conference on Fuzzy Systems*, Taipei, Taiwan, 27-30 June, 2011 pp. 1255–1262.

[6] A. González, R. Pérez ,"Completeness and consistency conditions for learning fuzzy rules". *Fuzzy Sets and Systems*, **96** (1998) 37–51.

[7] A. González, R. Pérez, "SLAVE: to genetic learning system based on an iterative aproach", *IEEE Transaction on Fuzzy System*, **27 (2)** (1999) 176–191.

[8] A. González, R. Pérez, "A Study About the Inclusion of Linguistic Hedges in a Fuzzy Rule Learning Algorithm", *International Journal of Uncertainty Fuzziness and KnowledgeBased Systems*, **7** (1999) 257-266.

[9] A. González, R. Pérez, "An analysis of the scalability of an embedded feature selection model for classification problems". *Proceedings of the Information Processing and Management of Uncertainty on Knowledge-Based Systems IPMU 2006*, Granada 1996 pp. 1949–1956.

[10] A. González, R. Pérez, "Improving the genetic algorithm of SLAVE", *Mathware & Soft Computing*, **16** (2009) 59–70.

[11] F. Herrera, M. Lozano, and J.L. Verdegay, "Tuning fuzzy controllers by genetic algorithm", Int J. Approx. Reason., **12**, (1995) 299-315.

[12] Ishibuchi, H. et al, "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **29 (5)** (1999) 601-618.

[13] S. Kullback, *Information Theory and Statistics*, Gloucester, Mass, 1978.

[14] R.S. Michalski, "Understanding the nature of learning," Machine Learning: An artificial intelligence approach (Vol **II**). San Mateo, CA: Morgan Kaufmann, 1986.

[15] T. Mitchell, *Machine Learning*, Ed. MacGraw-Hill, 1997.

[16] L. Wehenkel, "On uncertainty measures used for decision tree induction", *Proceedings of the Information Processing and Management of Uncertainty on Knowledge-Based Systems IPMU 2006*, Granada 1996 pp. 413–418.

RECONOCIMIENTO DEL HABLA UTILIZANDO CONJUNTOS DIFUSOS INTERVALO-VALORADOS AJUSTADOS CON TÉCNICAS EVOLUTIVAS

Juan Cerrón, **Raúl Orduna**, **Mikel Galar**, **Edurne Barrenechea**

Dpto. Automática y Computación, Universidad Pública de Navarra,
cerron.60794@e.unavarra.es,{raul.orduna,mikel.galar,edurne.barrenechea}@unavarra.es

Resumen

En este trabajo presentamos un sistema de detección de fonemas basado en técnicas difusas (conjuntos difusos intervalo-valorados) y ajuste de las funciones de pertenencia, asociadas a las bandas del espectrograma de voz, con algoritmos genéticos. El análisis experimental muestra que el modelo obtiene resultados de alta precisión y que se comporta de forma robusta en presencia de ruido.

Palabras Clave: Reconocimiento automático del habla, Conjuntos difusos intervalo-valorados, Algoritmos genéticos.

1 Introducción

El Reconocimiento Automático del Habla (RAH) es un área de la Inteligencia Artificial que está en auge en los últimos años. En la literatura especializada existen diferentes aproximaciones como por ejemplo, los Modelos Ocultos de Márkov [5], las Redes Neuronales [1, 8], el Análisis de Componentes Independientes [9] o la Lógica Difusa [6], entre otras. Sin embargo, todavía no ha sido posible maximizar dos de los criterios más importantes como son la precisión y la velocidad. Adicionalmente existe una diferencia cualitativa entre los resultados obtenidos con sistemas inteligentes y los obtenidos por humanos.

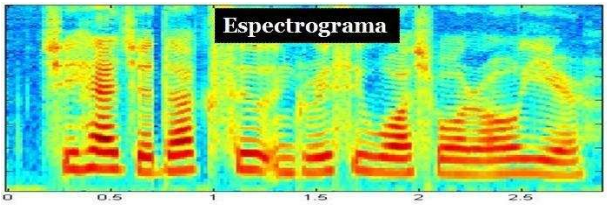
La Lógica Difusa es una de las mejores herramientas para modelar la falta de conocimiento preciso de los humanos [10]. La teoría de los conjuntos difusos y sus extensiones (como por ejemplo, los conjuntos difusos intervalo-valorados y los conjuntos intuicionistas difusos de Atanassov) han sido utilizas con éxito en procesamiento de imagen [2, 3, 4]. En este trabajo interpretamos un espectrograma, es decir, la representación de un sonido, como un tipo de imagen con ciertas restricciones (una de las dimensiones de la imagen corresponde a una magnitud temporal en el mundo físico). Por tanto, consideramos que los buenos resultados obtenidos en procesamineto de imagen con la teoría difusa también pueden obtenerse en el RAH. En [6] utilizan los conjuntos difusos para el RAH, nuestro objetivo es, basándonos en dicho trabajo, utilizar los conjuntos difusos intervalo-valorados para representar los fonemas y comparar los resultados obtenidos con ambas aproximaciones. Más aún, vamos a intentar reconocer muestras de fonemas en medios ruidosos.

2 Preliminares

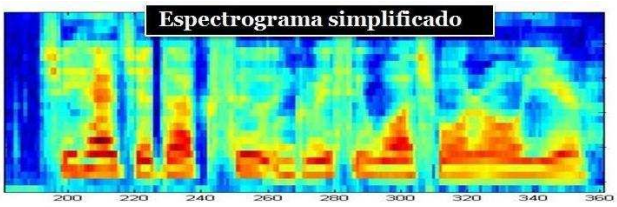
A continuación, recordamos algunos conceptos que son necesarios para el desarrollo de este trabajo.

Un espectrograma es la representación de un sonido, es decir, es la representación de la amplitud del sonido para diferentes frecuencias en intervalos de tiempo. Por tanto, un espectrograma proporciona una representación visual de fonemas, palabras y frases (ver Fig. 1). Podemos observar en la Fig. 1(a) un escenario que contiene mucha información, por este motivo son realizados normalmente dos procesos de simplificación. El primero es una adaptación a la percepción humana de las frecuencias del sonido y el segundo, es una agrupación de los valores del eje vertical en bandas horizontales. En la Fig. 1(b) observamos el resultado de ambos procesos de simplificación.

En la realización del primer proceso es utilizada la *escala de Mel* que tiene como objetivo situar de forma lineal aquellos tonos que son percibidos por un observador humano como equidistantes. Dicha escala ajusta las frecuencias, dadas en Hz, a partir de datos subjetivos y, por tanto, se adapta mejor a la percepción humana de los sonidos. Además, el ser humano no interpreta un espectrograma con precisión (se centra en frecuencias entre 4 y 6 KHz) y tampoco necesita saber cuál es la amplitud exacta de una onda (con saber si es alta o baja es suficiente). De igual manera sucede con la duración de un sonido, solo se guarda información de su duración aproximada (larga, corta, muy larga, etc). Por todos estos motivos, consideramos apropiado la utilización de la teoría difusa y en concreto, los conjuntos



(a) Espectrograma inicial



(b) Espectrograma reducido

Figura 1: Reducción del espectrograma

difusos intervalo-valorados.

Para el segundo proceso de simplicación, agrupamos el eje vertical en 25 bandas horizontales (veáse [6]). Los valores de cada banda son representados por la media de un porcentaje N de valores mínimos (ver Eq. 1).

$$MinN(V) = \frac{\sum_{1}^{round(\frac{N}{100}\times M)} Ordenar(V)_i}{(\frac{N}{100}\times M)}, donde \qquad (1)$$

V es un vector con M números y Ordenar(V) es la función que ordena V de menor a mayor.

Estructuramos un fonema en 25 bandas horizontales. Cada banda está compuesta por dos etiquetas diferentes que representan rangos de amplitud de onda en una escala cromática (de menor a mayor: negro, azul, magenta, rojo, amarillo, verde, cian y blanco). Además, el fonema completo requiere una etiqueta adicional para representar la longitud de un intervalo de tiempo (muy corto, corto, medio, largo y muy largo).

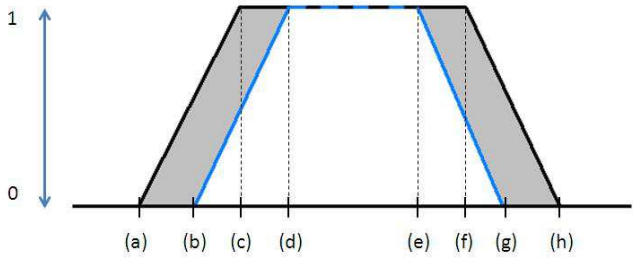
Denominamos *bloque* al intervalo de tiempo que define un fonema en el espectrograma simplificado. En la Tabla 1 mostramos la definición del fonema 'EH' (en inglés es el primer fonema de la palabra *have*).

En este trabajo representamos las etiquetas asociadas a la longitud y a las amplitudes de cada fonema mediante conjuntos difusos intervalo-valorados. Es conocido que la bondad de los resultados obtenidos con la teoría de conjuntos difusos depende de la elección de la función de pertencia asociada a dichos conjuntos. Los conjuntos difusos intervalo-valorados nos permiten modelizar la incertidumbre de la elección de la función de pertenencia, utilizando la longitud de los intervalos asociados a cada elemento del conjunto.

Fonema:	/EH/	
Banda 1:	Negro	Azul
Banda 2:	Negro	Azul
Banda 3:	Azul	Magenta
Banda 4:	Rojo	Amarillo
⋮		
Banda 25:	Negro	Azul
Longitud:	Muy largo	

Tabla 1: Definición del fonema 'EH'

Tanto las bandas como la longitud del fonema (ver Tabla 1), están definidas mediante funciones de pertenencia difusas intervalo-valoradas (dibujamos un ejemplo de función intervalar en la Fig. 2).



(a) inicialMin (b) inicialMax
(c) primerMaximoMin (d) primerMaximoMax
(e) ultimoMaximoMin (f) ultimoMaximoMax
(g) finalMin (h) finalMax

Figura 2: Representación intervalar de una etiqueta

En [6] los autores utilizan conjuntos difusos cuyas funciones de pertenencia están inicializadas según mostramos en la Tabla 2. En este trabajo partimos de dichas funciones puntuales para inicializar las funciones intervalo-valoradas y construimos intervalos que deben satisfacer la siguiente condición:

$$Inicial \leq PrimerMax \leq UltimoMax \leq Final \qquad (2)$$

donde
Inicial= $[inicialMin, inicialMax]$,
PrimerMax= $[primerMaximoMin, primerMaximoMax]$,
UltimoMax= $[ultimoMaximoMin, ultimoMaximoMax]$,
Final= $[finalMin, finalMax]$

Evidentemente, la inicialización de las funciones de pertenencia que realizamos en este trabajo satisfacen la condición (2).

3 Algoritmo de reconocimiento de fonemas

El proceso de reconocimiento consiste en extraer la secuencia de fonemas que componen una grabación de voz. Re-

alizamos el reconocimiento de cada fonema comparando cada bloque de la grabación con nuestra representación de los fonemas. De esta forma obtenemos, para cada bloque, un valor de coincidencia de todos nuestros fonemas. Finalmente asignamos al bloque el fonema cuyo valor de coincidencia es el más alto. Los tres pasos que detallan este proceso son los siguientes.

3.1 Comparación entre bloque y fonemas

Es necesario recordar que el espectrograma simplificado tiene 25 bandas. El fonema está compuesto por dos etiquetas para cada banda que hacen referencia a las funciones de pertenencia definidas. Calculamos la pertenencia, a las etiquetas asociadas a cada banda del fonema, de los valores de cada banda del bloque. De esta forma obtenemos como resultado dos intervalos para cada punto único del bloque.

3.2 Elección del intervalo representante de la intensidad de un fonema

En este momento tenemos dos intervalos como resultado de cada banda. Para cada banda elegimos el máximo de los dos intervalos como representante de la misma, calculándolo como el intervalo definido entre el máximo de los extremos inferiores y el máximo de los extremos superiores. Una vez obtenido un representante por banda, elegimos el mínimo de entre todos los representantes de las bandas como representante de la intensidad del fonema de forma análoga, como el mínimo de los extremos inferiores y el mínimo de los extremos superiores.

Entendemos que si el intervalo obtenido es alto, la probabilidad de que las intensidades del bloque correspondan a las intensidades asociadas al fonema es alta.

3.3 Elección del valor de pertenencia

Nuestro objetivo es obtener un único valor que represente la pertenencia de un fonema a un bloque. Sin embargo tenemos el intervalo que representa la pertenencia de las intensidades de un bloque a un fonema. Es necesario incorporar la información sobre la longitud del bloque, comparándola con la longitud esperada para cada fonema (banda 26 en Tabla 1). Para ello calculamos la pertenencia de la longitud real del bloque a la función de longitud que caracteriza al fonema, obteniendo un nuevo intervalo. Incorporamos esta información mediante una t-norma intervalo-valorada t-representable construida a partir de la t-norma producto: [minRepresentante $\times$ minLontigud, maxRepresentante $\times$ maxLongitud].

Por último realizamos una agregación de los valores que definen el intervalo resultante para obtener un único valor, interno al intervalo, que lo representa. En nuestro caso escogemos la media aritmética como función de agregación.

	Inicial	**PrimerMax**	**ÚltimoMax**	**Final**
Negro	0	0	0.05	0.35
Azul	0	0.05	0.22	0.5
Magenta	0	0.23	0.35	0.63
Rojo	0.05	0.36	0.5	0.8
Amarillo	0.23	0.51	0.63	0.93
Verde	0.36	0.63	0.8	0.99
Cian	0.51	0.81	0.93	0.99
Blanco	0.63	0.94	0.99	0.99
Muy corto	0.02	0.02	0.03	0.05
Corto	0.02	0.02	0.06	0.10
Media	0.02	0.06	0.12	0.17
Larga	0.08	0.12	0.20	0.30
Muy larga	0.12	0.18	0.69	0.99

Tabla 2: Inicialización de las funciones intervalares de pertenencia

4 Optimización de la definición de fonemas utilizando algoritmos genéticos

Los fonemas que forman parte de una lengua pueden definirse de diferentes formas. Sea cual sea el tipo de definición utilizada muchos comparten una estructura y definición similares. Optimizando la definición de los fonemas pretendemos que las pequeñas diferencias que caracterizan un fonema y lo distinguen de otro se ajusten de manera automática.

Además un fonema se puede pronunciar de muchas formas según la persona o el tono de su voz. Por ello también intentamos generalizar la definición de cada fonema para que se pueda reconocer en distintos hablantes.

En este trabajo utilizamos algoritmos genéticos para la optimización de las funciones de pertenencia intervalo-valoradas y la definición de los fonemas a partir de las mismas.

	Inicial	PrimerMax	ÚltimoMax	Final
Negro	[0, 0]	[0, 0]	[0.1, 0.15]	[0.33, 0.34]
Azul	[0, 0]	[0.04, 0.04]	[0.38, 0.38]	[0.53, 0.53]
Magenta	[0, 0.03]	[0.22, 0.24]	[0.45, 0.63]	[0.73, 0.8]
Rojo	[0.03, 0.03]	[0.27, 0.32]	[0.73, 0.73]	[0.94, 0.94]
Amarillo	[0.14, 0.16]	[0.42, 0.5]	[0.78, 0.8]	[0.94, 0.96]
Verde	[0.36, 0.4]	[0.6, 0.6]	[0.91, 0.91]	[0.99, 0.99]
Cian	[0.44, 0.48]	[0.77, 0.8]	[0.95, 0.96]	[0.99, 0.99]
Blanco	[0.6, 0.68]	[0.93, 0.96]	[0.97, 0.99]	[0.99, 0.99]
Muy corto	[0, 0.01]	[0, 0.02]	[0.03, 0.03]	[0.06, 0.07]
Corto	[0, 0.02]	[0.02, 0.03]	[0.06, 0.08]	[0.07, 0.14]
Media	[0, 0.01]	[0.01, 0.03]	[0.10, 0.11]	[0.16, 0.23]
Larga	[0, 0.09]	[0.11, 0.13]	[0.21, 0.21]	[0.30, 0.32]
Muy larga	[0.14, 0.17]	[0.14, 0.18]	[0.65, 0.69]	[0.97, 0.98]

Tabla 3: Funciones intervalares de pertenencia entrenadas

Para cada individuo es necesario almacenar la descripción de fonema que utiliza y las funciones de pertenencia asociadas a esta descripción. La definición de las funciones de pertenencia y la descripción de los fonemas puede variar de

Definición de colores				
Negro	Inicial	PrimerMax	ÚltimoMax	Final
Azul	Inicial	PrimerMax	ÚltimoMax	Final
Magenta	Inicial	PrimerMax	ÚltimoMax	Final
Rojo	Inicial	PrimerMax	ÚltimoMax	Final
Amarillo	Inicial	PrimerMax	ÚltimoMax	Final
Verde	Inicial	PrimerMax	ÚltimoMax	Final
Cian	Inicial	PrimerMax	ÚltimoMax	Final
Blanco	Inicial	PrimerMax	ÚltimoMax	Final
Definición de longitudes				
Muy corto	Inicial	PrimerMax	ÚltimoMax	Final
Corto	Inicial	PrimerMax	ÚltimoMax	Final
Media	Inicial	PrimerMax	ÚltimoMax	Final
Larga	Inicial	PrimerMax	ÚltimoMax	Final
Muy larga	Inicial	PrimerMax	ÚltimoMax	Final
Definición de fonemas				

Fonema 1		
Banda 1:	Color 1	Color 2
Banda 2:	Color 1	Color 2
Banda 3:	Color 1	Color 2
Banda 4:	Color 1	Color 2
⋮		
Banda 25:	Color 1	Color 2
Banda 26:	Longitud	
⋮		
Fonema 2		
Banda 1:	Color 1	Color 2
Banda 2:	Color 1	Color 2
Banda 3:	Color 1	Color 2
Banda 4:	Color 1	Color 2
⋮		
Banda 25:	Color 1	Color 2
Banda 26:	Longitud	

Tabla 4: Genoma de un individuo

un individuo a otro. Véase un ejemplo de la composición de un genotipo en la Tabla 4.

En este trabajo creamos intervalos a partir de la inicialización mostrada en la Tabla 2. Utilizamos cada valor como límite inferior y superior del intervalo asociado al mismo. Sin embargo, inicializamos la descripción de los fonemas de cada individuo de manera aleatoria.

Una vez que establecemos cuál es la composición del genoma es necesario detallar las funciones de cruce y mutación utilizadas, ya que estas funciones están adaptadas a la estructura del mismo.

Para cada función de pertenencia definida en el genotipo realizamos el cruce entre dos individuos escogiendo, aleatoriamente y con probabilidad uniforme, dicha función de uno de los progenitores para crear el primer descendiente. Utilizamos la función de pertenencia del otro progenitor para crear el segundo descendiente.

De manera análoga, completamos el cruce de los individuos escogiendo descripciones completas de cada fonema para crear ambos descendientes.

Llevamos a cabo la mutación sobre los siguientes elementos de un individuo: un extremo del intervalo de una función de pertenencia o una etiqueta de la descripción de un fonema.

En las funciones de pertenencia sumamos o restamos 0.01 a uno de los puntos de un intervalo de forma aleatoria. Si el resultado de la mutación incumple la ecuación (2) reordenamos los valores para generar nuevos intervalos que la satisfagan.

En la descripción de un fonema cambiamos una de las 51 etiquetas (25 bandas $\times$ 2 colores + 1 longitud) por la etiqueta inmediatamente anterior o posterior. Mostramos el orden de las etiquetas en la Tabla 2, agrupando por un lado los colores y por otro las longitudes. Por ejemplo la etiqueta 'Negro', puede mutar a 'Azul' o a 'Blanco'. Nótese que el orden lo definimos de forma cíclica en los extremos.

Realizamos un aprendizaje supervisado en el que partimos de fonemas localizados y clasificados. En la Tabla 3 mostramos un ejemplo de los intervalos de las funciones de pertenencia tras uno de los entrenamientos realizados.

La evaluación de los resultados la realizamos aplicando a cada bloque de la muestra el proceso de reconocimiento, utilizando la descripción de los fonemas y funciones que caracterizan a cada individuo de la población. Después comparamos los resultados obtenidos (*lista* de fonemas con sus valores de pertenencia) con el resultado conocido (fonema que correspone acada bloque) para evaluar cada individuo.

La evaluación final de cada individuo la realizamos mediante el siguiente algoritmo:

1) Buscar en Lista_i, asociada al Bloque_i, la posición del fonema dado correspondiente a ese bloque, para todo $i \leq$ Número de fonemas de la muestra;
2) Multiplicar el valor de coincidencia calculado en la sección 3.3 por el valor de ajuste correspondiente a su posición en la Lista_i (ver Tabla 5);
3) Sumar todos los productos.

Así conseguimos no sólo tener en cuenta el valor de pertenencia del fonema obtenido, sino la posición en la que aparece. Como la función de evaluación representa la cantidad y posición de fonemas acertados, la selección de los individuos se realiza con el objetivo de maximizar este valor.

Posicion en Lista	Valor de ajuste
1	1
2	0.5
3	0.4
4	0.4
5	0.3
6	0.3
7	0.3
8	0.15
9	0.15
10	0.15
11	0.15
>11	0

Tabla 5: Valores de ajuste asociados a la posición en la que se encuentra el fonema correcto

5 Resultados Experimentales

En primer lugar fijamos los parámetros específicos que usamos en nuestro algoritmo genético. Posteriormente, realizamos el entrenamiento de los parámetros para aprender la mejor descripción de los fonemas y las mejores funciones de pertenencia intervalo-valoradas que los describen. Finalmente, evaluamos los resultados obtenidos con una muestra de sonido a la que añadimos diferentes niveles de ruido.

5.1 Parametros del ajuste genético

En este apartado detallamos todos los factores que hemos tenido en cuenta para realizar el algoritmo genético.

Población. Realizamos pruebas con 40, 100 y 200 individuos.

Selección. Realizamos la selección de individuos para el cruce mediante un torneo de tres individuos. El proceso de selección termina cuando obtenemos una población final con el doble de individuos que la población inicial.

Reemplazo generacional. Optamos por un reemplazo elitista en el que conservamos para la siguiente generación la mitad de individuos de la población doblada que tienen un mayor valor de evaluación.

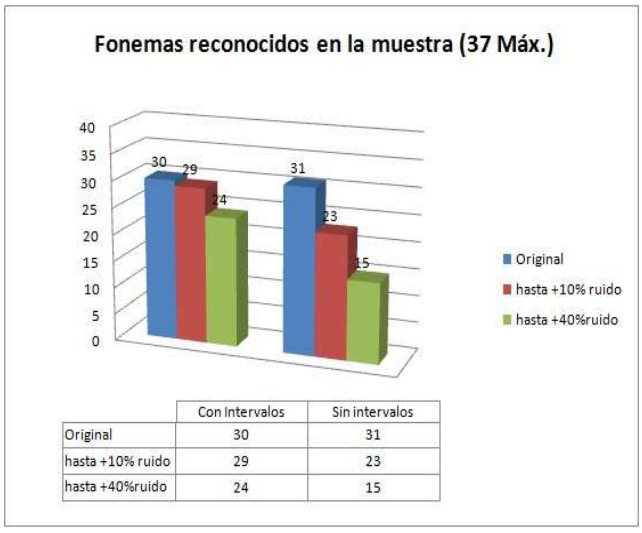
Criterio de parada. Fijamos un número determinado de iteraciones. Una iteración comprende todos los procesos realizados para obtener una nueva generación. Entrenamos con diferentes parámetros, en concreto con 400, 800 y 5000 iteraciones.

5.2 Evaluación del aprendizaje

Como conjunto de entrenamiento tomamos la frase "She had her dark suit in greasy wash water all year". La grabación, así como un análisis fonético de la misma, puede encontrarse en la base de datos (Linguistic Data Consortium) además de suministrarse a modo de ejemplo en la página web *http://www.ldc.upenn.edu/*. Obtenemos los mejores resultados con una población de 200 individuos y un criterio de parada establecido en 5000 iteraciones.

Generamos conjuntos de test con muestras calculadas al añadir ruido blanco a la muestra original. Este ruido consiste en aumentar o disminuir aleatoriamente un porcentaje a cada valor de la muestra original. El conjunto de test lo generamos a partir del conjunto de entrenamiento: sin añadir ruido, añadiendo un ruido de un 10% y añadiendo un ruido de un 40%. Finalmente probamos el proceso de reconocimiento sobre el conjunto de test que hemos generado con los fonemas obtenidos en la fase de entrenamiento. Contamos como fonemas acertados aquellos que han sido clasificados correctamente en el primer puesto de la lista utilizando los fonemas y funciones de pertenencia obtenidos del entrenamiento.

Mostramos los resultados sobre la bondad de los fonemas que se han reconocido en la muestra original y en las muestras con ruido de dos formas diferentes. El caso "Sin intervalos" corresponde con el implementado en [6]. En la Figura 3 mostramos el número total de fonemas reconocidos en cada una de las muestras, mientras que en la Tabla 6 detallamos si cada fonema concreto ha sido detectado correctamente en cada una de las muestras utilizadas.



	Con Intervalos	Sin intervalos
Original	30	31
hasta +10% ruido	29	23
hasta +40%ruido	24	15

Figura 3: Representa el número de fonemas detectados correctamente en cada una de las muestras

6 Conclusiones y futuros trabajos

Hemos probado que el uso de conjuntos difusos intervalo-valorados nos permite reconocer, con una alta tasa de acierto, una muestra con mucha variabilidad. Como apreciamos en los resultados, el número de fonemas encontrados en muestras con ruido al aplicar nuestro algoritmo es bastante

													Fonemas de la muestra																										
1	2	3	4	5	6	7	8	6	9	10	11	12	13	14	15	16	17	18	19	3	13	3	20	21	2	22	20	21	23	24	21	25	26	8	24
h	sh	ix	hv	eh	dcl	jh	ih	dcl	d	ah	kcl	k	s	ux	q	en	gcl	g	r	ix	s	ix	w	ao	sh	epi	w	ao	dx	axr	ao	l	y	ih	axr
																			Encontrados en Original																	
1	2	3	4	5	6	7	20	6	9	21	11	12	6	20	15	16	17	18	19	3	6	3	20	21	2	22	20	21	23	24	21	25	26	20	24
100	100	93	100	56	100	100	59	100	100	36	100	100	100	76	100	100	100	100	100	76	100	100	100	79	100	100	100	99	100	100	99	100	100	92	100
																			Encontrados en +10% ruido																	
1	2	3	4	5	6	7	25	6	9	25	11	12	2	20	15	16	17	18	19	3	11	3	20	21	2	22	25	21	23	24	21	25	26	3	24
100	100	70	86	54	100	100	59	100	87	33	100	100	99	73	100	100	100	100	94	75	99	100	98	58	100	98	100	96	100	100	99	100	100	85	100
																			Encontrados en +40% ruido																	
1	2	3	4	5	6	6	25	6	9	25	11	12	2	25	15	16	17	18	19	25	6	3	25	21	11	22	19	21	23	3	21	25	26	20	24
100	100	65	75	47	100	98	58	100	69	25	100	100	98	65	100	100	97	98	90	59	99	100	98	51	100	97	94	99	99	96	93	99	96	87	100

Tabla 6: Fonemas encontrados en cada muestra con sus respectivos procentajes de coincidencia. La frase de muestra es *She had her dark suit in greasy wash water all year*.

elevado. Si comparamos nuestro método con el método desarrollado en [6], que usaba conjuntos difusos, podemos concluir que nuestro método es mucho más robusto ante la presencia de ruido (Fig. 3).

Aunque el tiempo de entrenamiento es más lento porque trabajamos con intervalos, este hecho no afecta a la fase de reconocimento. Por lo tanto es un mecanismo apropiado para ser usado en tiempo real.

En futuros trabajos pretendemos incorporar diferentes técnicas a la hora de elegir el intervalo representante de una banda, ya que el máximo y el mínimo pueden no resultar muy precisos en algunos casos. Además, queremos comprobar cómo reconoce el algoritmo fonemas pronunciados por diferentes individuos.

Referencias

[1] M. R. Berthold, A time delay radial basis function network for phoneme recognition. En: *Proceedings of the IEEE International Conference on Neural Networks* 7, pp. 4470–4473, 1994.

[2] H. Bustince, E. Barrenechea, M. Pagola, J.Fernandez: Interval-valued fuzzy sets constructed from matrices: Application to edge detection, *Fuzzy Sets and systems* 160, pp. 1819–1840, 2009.

[3] H. Bustince, E. Barrenechea, M. Pagola, J.Fernandez: Contrast of a fuzzy relation, *Information Sciences* 180, pp. 1326–1344, 2010.

[4] H. Bustince, M. Pagola, E. Barrenechea, J. Fernandez, P. Melo-Pinto, P. Couto, H.R. Tizhoosh, J. Montero: Ignorance functions. An application to the calculation of the threshold in prostate ultrasound images, *Fuzzy Sets and systems* 161, pp. 20–36, 2010.

[5] S.E. Golowich, D.X. Sun: A support vector/Hidden Markov model approach to phoneme recognition. En: *Proceedings of the ASA Statistical Computing Section*, pp. 125–130, 1998.

[6] R. Halavati, S.B. Shouraki, S.H. Zadeh: Recognition of human speech phonemes using a novel fuzzy approach, *Applied Soft Computing* 7, pp. 828–839, 2007.

[7] R. Halavati, S.B. Shouraki, M. Eshraghi, M. Alemzadeh, P. Ziaie: A Novel Fuzzy Approach to Speech Recognition. En: *Proceedings of the 4th International Conference on HIS*, Kitakyushu (Japón), pp. 340–345, 2004.

[8] N.K. Kasabov, R. Kozma, M.J. Watts: Phoneme-based speech recognition via fuzzy neural networks modeling and learning,*Information Sciences* 110, pp. 61–79, 1998.

[9] O.W. Kwona, T.W. Lee: Phoneme recognition using ICA-based feature extraction and transformation. En: *Signal Process*, 84 pp. 1005–1019, 2004.

[10] L.A. Zadeh: Fuzzy sets. En: *Information Control* 8, pp. 338–353, 1965.

UNA PROPUESTA COOPERATIVA ENTRE CONJUNTOS DIFUSOS INTERVALO-VALORADOS Y AJUSTE EVOLUTIVO PARA MEJORAR EL RENDIMIENTO DE ÁRBOLES DE DECISIÓN DIFUSOS

José Antonio Sanz[1], **Alberto Fernandez**[2], **Humberto Bustince**[1], **Francisco Herrera**[3]

[1]*Departamento de Automatica y Computacion, Universidad Publica de Navarra, 31006 España*, {joseantonio.sanz,bustince}@unavarra.es

[2]*Departamento de Ciencias de la Computación, Universidad de Jaén, 23071 España*, alberto.fernandez@ujaen.es

[3]*Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, 18071 España*, herrera@decsai.ugr.es

Resumen

La definición de las funciones de pertenencia tiene un rol fundamental en el éxito de los sistemas de clasificación difusos en general y de los árboles de decisión difusos en particular. Sin embargo, la posible falta de información existente a la hora de definirlas, causa un cierto grado de ignorancia que dificulta dicho proceso.

En este trabajo, proponemos modelar este grado de ignorancia mediante conjuntos difusos intervalo-valorados y, de este modo, mejorar el rendimiento de los árboles de decisión difusos al afrontar problemas de clasificación. Para ello, definimos una nueva metodología, compuesta por cuatro etapas, en la que generemos e integramos los conjuntos difusos intervalo-valorados en el algoritmo ID3 difuso.

En el estudio experimental, demostramos la calidad de nuestra propuesta comparando sus resultados con los obtenidos por el algoritmo ID3 difuso y los obtenidos por dos propuestas de aprendizaje evolutivo de árboles de decisión difusos.

Palabras Clave: Árboles de decisión difusos, conjuntos difusos intervalo-valorados, funciones de ignorancia, ajuste evolutivo, clasificación.

1 Introducción

Los Árboles de Decisión Difusos (ADDs) son una solución apropiada para afrontar problemas de clasificación [2] puesto que combinan el manejo de la incertidumbre aportado por la lógica difusa y la interpretabilidad de los árboles de decisión. Entre el extenso número de técnicas propuestas en la literatura especializada para diseñar ADDs, el algoritmo ID3 difuso se utiliza ampliamente ya que es un método popular y eficiente para construirlos [13].

Uno de los factores más importantes en el éxito de los sistemas difusos, como los ADDs, es realizar una buena definición de las funciones de pertenencia asociadas a cada etiqueta lingüística que utilice el sistema. En numerosas ocasiones puede haber una falta de conocimiento o información para realizarla, causando la existencia de un cierto grado de ignorancia asociado a su definición. La teoría de los Conjuntos Difusos Intervalo-Valorados (CDIVs) [10] ha demostrado ser adecuada para modelar este grado de ignorancia [3]. Concretamente, en nuestros trabajos anteriores hemos comprobado que los CDIVs son útiles para mejorar el rendimiento de los sistemas de clasificación basados en reglas difusas [11, 12].

En este trabajo, presentamos una nueva metodología para mejorar el rendimiento de ADDs. Esta propuesta está compuesta por cuatro etapas: 1) generamos un ADD inicial utilizando el algoritmo ID3 difuso [15]; 2) modelamos sus etiquetas lingüísticas mediante CDIVs, para lo cual utilizamos un nuevo método parametrizado de construcción de CDIVs basado en funciones de ignorancia débiles [12]; 3) extendemos el método de razonamiento difuso para tener en cuenta el grado de ignorancia en todo el proceso de inferencia; 4) definimos una propuesta de ajuste evolutivo para calcular el mejor grado de ignorancia que representa cada CDIV, puesto que la falta de conocimiento o información puede variar en función de cada función de pertenencia. Debemos destacar que en todo momento mantenemos la estructura del ADD inicial y por tanto su interpretabiliad permance intacta.

El objetivo del trabajo es demostrar la bondad de nuestra propuesta. Para ello, comparamos el rendimiento de nuestro nuevo método con el obtenido por el algoritmo ID3 difuso [15] y el de dos propuestas de generación de ADDs que utilizan algoritmos evolutivos [9, 4]. Analizamos los resultados de los métodos utilizando 20 conjuntos de datos numéricos seleccionados del repositorio KEEL [1] y realizamos un estudio estadístico exhaustivo [7, 8] para avalar nuestras conclusiones.

El trabajo esta organizado de la siguiente forma: en la

Sección 2 recordamos varios conceptos preliminares sobre CDIVs y describimos el método de construcción de CDIVs utilizando funciones de ignorancia débiles. En la Sección 3 describimos la extensión del método de razonamiento difuso para poder trabajar con CDIVs y la propuesta de ajuste evolutivo. A continuación, presentamos el análisis de los resultados obtenidos en la Sección 4 y finalizamos el trabajo con las conclusiones en la Sección 5.

2 Etiquetas lingüísticas modeladas mediante conjuntos difusos intervalo-valorados

En esta sección recordamos varios conceptos teóricos que son necesarios para entender el resto del trabajo. A continuación, definimos el método de construcción de CDIVs en el que utilizamos funciones de ignorancia débiles. Finalmente, presentamos nuestro método para modelar las etiquetas lingüísticas con CDIVs.

2.1 Preliminares

Denotamos con $L([0,1])$ al conjunto de todos los subintervalos cerrados del intervalo cerrado $[0,1]$, es decir: $L([0,1]) = \{\mathbf{x} = [\underline{x},\overline{x}] | (\underline{x},\overline{x}) \in [0,1]^2 \text{ y } \underline{x} \leq \overline{x}\}$.

Para ordenar los intervalos, utilizamos la siguiente relación de orden total de intervalos definida por Xu y Yager [14]. Dados dos intervalos $[a,b]$ y $[c,d]$, $[a,b] \leq [c,d]$ si y solo si $a+b < c+d$ o $a+b = c+d$ y $d-c \leq b-a$.

Utilizando esta relación en $L([0,1])$ es fácil comprobar que $0_L = [0,0]$ y $1_L = [1,1]$ son el menor y el mayor elemento de $L([0,1])$ respectivamente.

Definición 1 *Un conjunto difuso intervalo-valorado (CDIV) A en el universo $U \neq \emptyset$ es una función $A_{IV} : U \to L([0,1])$, tal que el grado de pertenencia de $u \in U$ está dado por $A_{IV}(u) = [\underline{A}(u), \overline{A}(u)] \in L([0,1])$.*

Denotamos con W a la amplitud del intervalo bajo consideración: $W(A_{IV}(u_i)) = \underline{A}(u_i) - \overline{A}(u_i)$. Esta amplitud del intervalo puede interpretarse como una representación del grado de ignorancia asociado a la definición de la función de pertenencia. Independientemente de la fuente de la ignorancia, puede ser cuantificada por medio del concepto de función de ignorancia débil [12]. En este trabajo, utilizamos $g(x) = 2 \cdot \text{mín}(x, 1-x)$ como función de ignorancia débil.

Además, utilizamos $\mathbf{T}_{T_a,T_b}$ para denotar a las t-normas intervalo-valoradas t-representables. Este concepto es la extensión de las t-normas en CDIVs [5].

2.2 Método de construcción de conjuntos difusos intervalo-valorados

Nuestro objetivo en esta sección es construir un CDIV a partir de un conjunto difuso. Para ello, definimos una función G parametrizada por δ y γ que satisface un determinado conjunto de propiedades. Estas propiedades permiten que la amplitud del intervalo construido sea proporcional al grado de ignorancia y que el grado de pertenencia inicial esté dentro del intervalo construido.

Proposición 1 *Dados $\delta, \gamma \in [0,1]$ con $\delta \geq \gamma \geq \delta \cdot x$. La función*

$$\begin{aligned} &G : [0,1]^4 \to L([0,1]) \text{ dada por} \\ &G(x,y,\delta,\gamma) = [x \cdot (1 - \delta \cdot y), x \cdot (1 - \delta \cdot y) + \gamma \cdot y)] \end{aligned} \tag{1}$$

satisface las siguientes propiedades:

1. $x \in G(x,y,\delta,\gamma)$;
2. $W(G(x,y,\delta,\gamma)) = \gamma \cdot y$;
3. *Si* $x = 0$, *entonces* $G(0,y,\delta,\gamma) = [0, \gamma \cdot y]$;
4. *Si* $y = 0$, *entonces* $G(x,0,\delta,\gamma) = [x,x]$;
5. *Si* $\delta = \gamma$, *entonces:*

$$\begin{aligned} \underline{G}(x,y,\delta,\delta) + \overline{G}(1-x,y,\delta,\delta) = 1 \\ \overline{G}(x,y,\delta,\delta) + \underline{G}(1-x,y,\delta,\delta) = 1. \end{aligned}$$

De acuerdo a la proposición anterior, definimos el Teorema 1 como el nuevo método de construcción de CDIVs.

Teorema 1 *Sea μ_A un conjunto difuso. Si para cada $u_i \in U$ tomamos $g(\mu_A(u_i))$, $\delta(u_i)$, $\gamma(u_i) \in [0,1]$, entonces el conjunto*

$$\begin{aligned} &A_{IV} = \{(u_i, A_{IV}(u_i)) | u_i \in U\} \text{ donde} \\ &A_{IV}(u_i) = G(\mu_A(u_i), g(\mu_A(u_i)), \delta(u_i), \gamma(u_i)) \end{aligned} \tag{2}$$

es un CDIV en U.

2.3 Modelado de las etiquetas lingüísticas mediante conjuntos difusos intervalo-valorados

Tal y como hemos descrito en la sección anterior, para construir un CDIV necesitamos un conjunto difuso de partida. Para ello, creamos las funciones de pertenencia asociadas a las etiquetas lingüísticas de cada variable realizando un particionado lineal del espacio de entrada. La línea discontinua de la Figura 1 muestra un ejemplo de función de pertenencia, cuya ecuación es la siguiente:

$$\mu_A(u_i) = \begin{cases} 0, & \text{si } u_i \leq a, \\ \frac{2}{b-a}(u_i - a), & \text{si } a \leq u_i \leq \frac{a+b}{2} \text{ y } b \neq a, \\ \frac{2}{a-b}(u_i - b), & \text{si } \frac{a+b}{2} \leq u_i \leq b \text{ y } b \neq a, \\ 0, & \text{si } b \leq u_i, \end{cases} \tag{3}$$

Posteriormente, aplicamos el Teorema 1 a cada una de las funciones de pertenencia que utilice el sistema. Para la

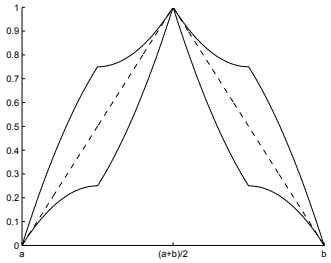


Figura 1: CDIV inicial.

construcción inicial de todos los CDIVs consideramos el nivel medio de grado de ignorancia. Por tanto, inicializamos $\delta(u_i) = \gamma(u_i) = 0.5$ para todo $u_i \in U$ puesto que, de acuerdo con el Teorema 1, el valor mínimo de ambos parámetros es 0 y el máximo 1. De este modo, el CDIV construido a partir de la función de pertenencia de la Ecuación 3 es el mostrado por la línea sólida en la Figura 1.

3 Algoritmo ID3 difuso con conjuntos difusos intervalo-valorados

En esta sección, describimos nuestro ADD con CDIVs. Nuestra nueva propuesta está compuesta por 4 etapas. La primera de ellas consiste en generar el ADD inicial. Para ello, utilizamos el algoritmo ID3 difuso propuesto por Yuan y Shaw en [15] utilizando como conjuntos difusos los generados al realizar un particionado lineal del espacio de entrada y cuyas funciones de pertenencia asociadas están descritas por la Ecuación 3.

En la segunda etapa, modelamos las etiquetas lingüísticas, del ADD inicial generado en el paso anterior, mediante CDIVs aplicando el método de construcción presentado en la Sección 2.3. De este modo, aportamos al ADD la capacidad de manejar el grado de ignorancia asociado a la definición de sus funciones de pertenencia y mantenemos su interpretabilidad, puesto que tanto el número de reglas como los antecedentes de las mismas son los mismos que los del ADD generado por el algoritmo ID3 difuso.

En el resto de esta sección describimos en detalle las dos últimas etapas de nuestro método: la extensión del método de razonamiento difuso (Sec. 3.1) y la propuesta de ajuste evolutivo (Sec. 3.2).

3.1 Método de razonamiento difuso intervalo-valorado

El modelado de las etiquetas lingüísticas por medio de CDIVs implica principalmente dos modificaciones. Las frecuencias relativas de las clases en cada hoja deben ser elementos de $L([0,1])$. Para conseguirlo, utilizamos la aritmética de intervalos definida por Descrijiver en [5] aplicada sobre la ecuación usada para calcular dichas frecuencias relativas:

$$p_i^{(k)}(j) = \frac{|A_i^{(k)} \cap C_j|}{|A_i^{(k)}|}, \tag{4}$$

donde $(k = 1,2,...,n), (i = 1,2,...,m_k)$ y $(j = 1,2,...,m)$, siendo n el número de atributos, m_k el número de etiquetas lingüísticas del atributo k y m el número de clases. Mediante $|.|$ denotamos el cardinal de un conjunto difuso, por ejemplo para un conjunto difuso A: $|A| = \sum_{x \in X} \mu_A(x)$

Además, debemos extender el método de razonamiento difuso utilizado para clasificar nuevos ejemplos.

En un ADD, una regla es la conexión entre el nodo raíz y cada uno de los nodos hoja. Suponemos que el ADD tiene l hojas, por tanto l reglas $(Regla^i, i = 1,2,...,l)$. Para cada hoja, la regla i tiene N_i nodos ($Regla_1^i$, $Regla_2^i$, ...,$Regla_{N_i}^i$). Recordamos que cada nodo está modelado por medio de un CDIV, por tanto, $Regla_k^i(e_k) = [\underline{Regla_k^i}(e_k), \overline{Regla_k^i}(e_k)] \in L([0,1])$ con $k = 1,...,N_i$. Para clasificar un nuevo ejemplo e en una de las m clases, el *método de razonamiento difuso intervalo-valorado* sigue los 4 siguientes pasos:

- Calcular el grado de emparejamiento entre cada regla y el nuevo ejemplo.

$$E_i = \mathbf{T}_{T_a,T_b}([\underline{Regla_1^i}(e_1), \overline{Regla_1^i}(e_1)], ...,$$
$$[\underline{Regla_{N_i}^i}(e_{N_i}), \overline{Regla_{N_i}^i}(e_{N_i})]), i = 1,2,...,l$$

- Calcular la certeza de cada clase en cada hoja. Ponderamos el grado de emparejamiento del ejemplo, para cada una de las reglas, con las frecuencias relativas de sus correspondientes hojas.

$$Cert_j^i = \mathbf{T}_{T_c,T_d}([\underline{E_i}, \overline{E_i}], [\underline{p_i}(j), \overline{p_i}(j)]), j = 1,...,m, i = 1,...,l$$

- Calcular la certeza total de cada clase. Utilizamos la suma de intervalos para agregar la certeza de cada clase en cada una de las hojas.

$$Certeza_Total_j = \sum_{i=1}^{l} [\underline{Cert_j^i}, \overline{Cert_j^i}], j = 1,...,m$$

- Clasificar el ejemplo en la clase que maximice la certeza total utilizando la relación de orden total de intervalos presentada en la Sección 2.1.

3.2 Ajuste evolutivo del factor de ponderación de la ignorancia

Tal y como hemos introducido en la Sección 2.3, para la construcción inicial de todos los CDIVs consideramos el

grado de ignorancia medio ($\delta = \gamma = 0.5$). Sin embargo, este grado de ignorancia puede variar en función de la etiqueta lingüística, puesto que la cantidad de información disponible para definirlas puede que no sea la misma.

Por tanto, proponemos aplicar una etapa de ajuste evolutivo en la cual modificar los valores de los parámetros δ y γ (ver Teorema 1), manteniendo la restricción $\delta \geq \gamma \geq \delta \cdot x$, para cada CDIV. Como consecuencia, el grado de ignorancia que representa cada CDIV variará en función de la adecuación de las funciones de pertenencia al problema específico que estemos tratando.

Para llevar a cabo esta propuesta de ajuste, utilizamos el algoritmo evolutivo CHC [6]. Este algoritmo presenta un buen equilibrio entre diversidad y convergencia, lo que lo hace apropiado para afrontar este tipo de problema [11]. Utilizamos una codificación real en la que cada par de genes representan la modificación de los valores de los parámetros δ y γ de cada CDIV. Como resultado, el cromosoma tiene la siguiente forma:

$$C_{ET} = (\delta_{11}, \gamma_{11}, ..., \delta_{1m_1}, \gamma_{1m_1}, ..., \delta_{n1}, \gamma_{n1}, ..., \delta_{nm_n}, \gamma_{nm_n})$$

Una descripción detallada del modelo evolutivo se puede encontrar en [11, 12].

La Figura 2 muestra el comportamiento del ajuste evolutivo. El CDIV inicial está mostrado en gris claro mientras que su forma final, tras aplicar el ajuste evolutivo, está mostrada en gris oscuro.

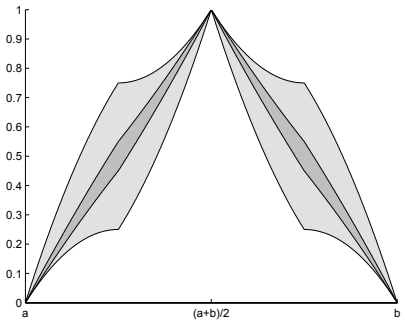


Figura 2: Ajuste evolutivo del factor de ponderación del grado de ignorancia. Los valores finales de los parámetros son: $\delta = 0.1$, $\gamma = 0.1$.

4 Estudio experimental

El objetivo de este estudio es demostrar que la cooperación entre los CDIVs y el ajuste evolutivo permite mejorar el rendimiento de ADDs. Para ello, comparamos los resultados de nuestra propuesta con los obtenidos por el algoritmo ID3 difuso y con los de las propuestas de aprendizaje evolutivo de ADDs definidas por Kim [9] y Chang [4].

En el resto de esta sección, primero describimos el marco experimental junto con los valores de los parámetros de los ADDs usados en el estudio, después introducimos los test estadísticos utilizados y para finalizar mostramos los resultados empíricos obtenidos.

4.1 Marco experimental

Hemos analizado el rendimiento de las diferentes propuestas en 20 conjuntos de datos seleccionados del repositorio KEEL[1] [1]. En la Tabla 1 podemos ver las siguientes características de cada conjunto de datos: número de ejemplos (#Ej.), número de atributos (#Atr.) y número de de clases (#Clas.). Debemos señalar que los conjuntos de datos *magic, page-blocks, ring* y *shuttle* han sido estratificados al 10 % para reducir su tamaño en la etapa de entrenamiento. Además, eliminamos del conjunto de datos las instancias que contengan valores perdidos (*cleveland* y *wisconsin*).

Tabla 1: Resumen de las características de los conjuntos de datos.

Id.	Conjunto de datos	#Ej.	#Atr.	#Clas.
app	Appendicitis	106	7	2
bal	Balance	625	4	3
bup	Bupa	345	6	2
cle	Cleveland	297	13	5
eco	Ecoli	336	7	8
gla	Glass	214	9	6
hab	Haberman	306	3	2
hea	Heart	270	13	2
iri	Iris	150	4	3
mag	Magic	1,902	10	2
new	New-Thyroid	215	5	3
pag	Page-blocks	548	10	5
pim	Pima	768	8	2
rin	Ring	740	20	2
shu	Shuttle	2,175	9	7
tae	Tae	151	5	3
tit	Titanic	2,201	3	2
win	Wine	178	13	3
wis	Wisconsin	683	9	2
yea	Yeast	1484	8	10

Para realizar la experimentación hemos empleado el modelo de validación cruzada de 5 particiones, es decir, 5 particiones con el 20 % de los datos cada una, combinando 4 de ellas para el entrenamiento y usando la restante para test. Como resultado de cada conjunto de datos empleamos la media de las 5 particiones.

Todos los ADDs considerados en este trabajo utilizan 3 etiquetas lingüísticas por variable y como operador de conjunción la t-norma producto, excepto el ADD propuesto por Kim que utiliza la t-norma mínimo. En el algoritmo ID3 difuso hacemos un α-corte con $\alpha = 0.4$ y el umbral de verdad es de 0.95. La configuración utilizada para las propuestas definidas por Kim y Chang es la recomendada en sus respectivos artículos y utilizamos $w = 0.98$ en el método de Kim. Con respecto a la propuesta de ajuste evolutivo presentada en la Sección 3.2, utilizamos poblaciones compuestas por 50 individuos, el número de evaluaciones

[1] http://www.keel.es/dataset.php

Tabla 2: Resultados en Entrenamiento (Ent.) y Test (Tst) conseguidos por las diferentes propuestas.

Conjunto	F-ID3_IVFS_ET		F-ID3		GAFDT_Kim		GAFDT_Chang	
de Datos	*Tr.*	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*
Bal	96.60	**90.72**	91.72	90.08	84.80	82.72	86.76	86.40
Bup	79.93	**67.25**	61.52	58.26	67.97	62.03	71.23	62.90
Cle	93.18	55.88	87.29	53.87	62.21	**57.59**	84.43	53.53
Eco	82.14	**77.69**	80.21	76.80	89.14	74.13	90.40	73.81
Gla	77.57	60.80	60.17	52.80	67.41	61.22	73.25	**62.65**
Hab	79.90	72.87	75.08	71.90	78.02	72.22	78.02	**72.89**
Iri	98.83	96.00	63.07	95.33	98.67	**96.67**	97.67	92.67
Mag	82.75	78.91	95.50	78.28	78.89	77.65	81.69	**79.65**
New	97.79	94.88	79.80	91.16	95.35	**95.35**	97.33	**95.35**
Pag	95.03	**94.16**	92.67	92.15	93.52	93.06	94.39	93.61
Pim	83.56	**76.43**	92.84	75.77	77.08	73.83	79.33	75.13
Rin	97.09	**90.81**	79.95	49.59	83.34	81.62	88.61	85.81
Win	100.00	**98.87**	49.59	97.75	96.49	90.40	97.89	95.49
Wis	98.28	96.04	90.63	94.58	97.11	95.90	97.99	**96.49**
Shu	97.74	**97.98**	98.74	90.57	94.56	94.34	85.95	85.47
Yea	63.01	**57.41**	95.90	55.46	49.88	48.38	43.48	44.95
App	93.16	**86.80**	90.57	84.94	91.03	85.84	92.92	84.94
Hea	94.81	78.52	93.52	79.26	85.09	76.30	92.96	**81.48**
Tae	70.86	**58.95**	63.40	54.99	65.73	49.74	67.72	56.34
Tit	78.33	78.33	78.33	78.33	79.05	**79.06**	78.90	78.33
Media	88.03	**80.46**	81.02	76.09	81.77	77.40	84.05	77.89

es de 5.000 por el número de variables y el valor del umbral tanto inicialmente como en el rearranque se calcula de la siguiente forma: $L = (\#Genes \cdot BITSGEN)/4.0$, donde *#Genes* es la longitud del cromosoma y $BITSGEN$ es el número de bits utilizados para realizar las codificación *gray* que en nuestro caso es 30.

4.2 Tests estadísticos utilizados para la comparativa del rendimiento

En esta contribución, usamos las técnicas de comprobación de hipótesis para proporcionar un soporte estadístico al análisis de los resultados. Concretamente, usaremos tests no paramétricos, dado que las condiciones iniciales que garantizan la fiabilidad de los tests paramétricos pueden no satisfacerse, haciendo que el análisis estadístico pierda credibilidad con estos tipos de tests [7]. Concretamente, usaremos el test de rangos alineados de Friedman para detectar diferencias estadísticas entre un grupo de resultados y el test post-hoc de Holm para encontrar qué algoritmos rechazan la hipótesis de igualdad con respecto a un método de control seleccionado. Además, calculamos el p-valor ajustado (APV) para tener en cuenta que se usan tests múltiples. De esta forma, podemos rechazar la hipótesis nula comparando el APV con el nivel de significancia α.

4.3 Resultados empíricos

La Tabla 2 muestra los resultados obtenidos tanto en entrenamiento como en test por las diferentes propuestas en cada conjunto de datos. En primer lugar aparecen los resultados obtenidos por nuestra nueva metodología, después los del algoritmo ID3 difuso y por último los de las propuestas definidas por Kim [9] y Chang [4]. Observamos que nuestro método obtiene una clara mejora con respecto al resto de métodos tanto en precisión como en número de victorias.

Para llevar a cabo nuestro estudio, primero mostramos en la Figura 3 el ranking de los distintos métodos utilizados asignando las posiciones de estos algoritmos en función del rendimiento para cada conjunto de datos y obteniendo a continuación el valor medio. Podemos observar que nuestra propuesta obtiene claramente la mejor posición.

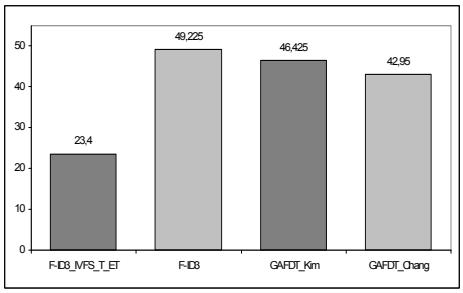


Figura 3: Rankings de los diferentes métodos.

A continuación, realizamos un test de rangos alineados de Friedman para detectar diferencias significativas entre los resultados de estas propuestas. El p-valor obtenido es cercano a cero (0.001), lo que implica que hay diferencias significativas entre los resultados y así, debemos aplicar un test post-hoc, en este caso el test de Holm (Tabla 3). A partir de los resultados de dicho test, demostramos que nuestra nueva propuesta es estadísticamente mejor que el resto.

5 Conclusiones

En este trabajo hemos propuesto una metodología para mejorar el rendimiento de los ADDs mediante CDIVs y un proceso de ajuste evolutivo. Para ello, hemos modelado las

Tabla 3: Test de Holm para comparar nuestra propuesta (método de control) con el resto de métodos.

i	Algoritmo	APV	Hipótesis ($\alpha = 0.05$)
3	F-ID3	0.001	Rechazado para F-ID3_IVFS_ET
2	GAFDT_Kim	0.003	Rechazado para F-ID3_IVFS_ET
1	GAFDT_Chang	0.008	Rechazado para F-ID3_IVFS_ET

etiquetas lingüísticas con CDIVs de forma que tenemos en cuenta el grado de ignorancia asociado a la definición de las funciones de pertenencia. El grado de ignorancia está ponderado por dos parámetros y, de esta forma, queda definida la forma de los CDIVs. Esta parametrización nos permite calcular el mejor grado de ignorancia que representa cada CDIV por medio de una etapa de ajuste evolutivo.

El estudio experimental ha demostrado la calidad de nuestra propuesta ya que, aplicado al ADD generado por el algoritmo ID3 difuso, permite mejorar los resultados de éste algoritmo y el de dos propuestas de aprendizaje evolutivo de ADD. Estos resultados nos permiten determinar, con el correspondiente aval estadístico, que el uso de los CDIVs es muy útil para afrontar problemas de clasificación ya que mejoramos el rendimiento de métodos difusos optimizados mediante algoritmos evolutivos.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el Ministerio de Educación y Ciencia con los proyectos TIN2011-28488 y TIN2010-15055.

Referencias

[1] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera: KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing* 17(2–3), pp. 255–287, 2011.

[2] E. Alpaydin: *Introduction to Machine Learning*, 2nd ed. The MIT Press, Cambridge, 2010.

[3] H. Bustince, M. Pagola, E. Barrenechea, J. Fernandez, P. Melo-Pinto, P. Couto, H. R. Tizhoosh, J. Montero: Ignorance functions. An application to the calculation of the threshold in prostate ultrasound images. *Fuzzy Sets and Systems* 161(1), pp. 20–36, 2010.

[4] J.-Y. Chang, C.-W. Cho, S.-H. Hsieh, S.-T. Chen: Genetic algorithm based fuzzy ID3 algorithm, in *Neural Information Processing*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, vol. 3316, pp. 989–995, 2004.

[5] G. Deschrijver: Generalized arithmetic operators and their relationship to t-norms in interval-valued fuzzy set theory. *Fuzzy Sets and Systems* 160(21), pp. 3080–3102, 2009.

[6] L. J. Eshelman: *Foundations of Genetic Algorithms*. Ch. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, Morgan Kaufman, pp. 265–283, 1991.

[7] S. García, F. Herrera: An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research* 9, pp. 2677–2694, 2008.

[8] S. García, A. Fernández, J. Luengo, F. Herrera: Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences* 180(10), pp. 2044–2064, 2010.

[9] M. Kim, J. Ryu: Optimized fuzzy decision tree using genetic algorithm, in *Neural Information Processing*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, vol. 4234, pp. 797–806, 2006.

[10] R. Sambuc: "Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne," Ph.D. dissertation, University of Marseille, 1975.

[11] J. Sanz, A. Fernández, H. Bustince, F. Herrera: Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning. *Information Sciences* 180(19), pp. 3674–3685, 2010.

[12] J. Sanz, A. Fernández, H. Bustince, F. Herrera: A genetic tuning to improve the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets: Degree of ignorance and lateral position. *International Journal of Approximate Reasoning* 52(6), pp. 751–766, 2011.

[13] X.-Z. Wang, D. S. Yeung, E. C. C. Tsang: A comparative study on heuristic algorithms for generating fuzzy decision trees. *IEEE Transactions on Systems, Man and Cybernetics - Part B* 31(2), pp. 215–226, 2001.

[14] Z. Xu, R. Yager: Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems* 35(4), pp. 417–433, 2006.

[15] Y. Yuan, M. J. Shaw: Induction of fuzzy decision trees. *Fuzzy Sets and Systems* 69(2), pp. 125–139, 1995.

APRENDIZAJE EVOLUTIVO DE SISTEMAS APROXIMATIVOS DE TIPO TSK PARA PROBLEMAS DE ALTA DIMENSIONALIDAD

M.J. Gacto[1], **R. Alcalá**[2], **F. Herrera**[2]

[1]*Dept. Informática, Universidad de Jaén, Jaén, España*, mgacto@ujaen.es
[2]*Dept. Ciencias de la Computación e I. A., Universidad de Granada, Granada, España*, {alcala,herrera}@decsai.ugr.es

Resumen

En esta contribución se propone un método en dos fases para el modelado difuso preciso en problemas de regresión con alta dimensionalidad mediante sistemas basados en reglas difusas aproximativos de tipo Takagi-Sugeno-Kang. En la primera fase se realiza un aprendizaje evolutivo de la base de datos (selección de variables, granularidad y parámetros de desplazamiento de las particiones difusas) y de la base de reglas en el mismo proceso. En la segunda fase se aplica un método de ajuste evolutivo combinado con una selección de reglas que incorpora un filtro de Kalman eficiente para mejorar la precisión del modelo. Ambas fases incluyen mecanismos que permiten trabajar con problemas con un elevado número de variables y de ejemplos.

El método propuesto ha sido comparado con el método clásico ANFIS y con dos algoritmos de aprendizaje evolutivo en siete problemas con diferente número de ejemplos y variables, mostrando mejores resultados en precisión en un tiempo razonable.

Palabras Clave: Sistemas basados en reglas difusas aproximativos, reglas Takagi-Sugeno-Kang, algoritmos evolutivos multi-objetivo, precisión, regresión y escalabilidad.

1 Introducción

En los Sistemas Basados en Reglas Difusas (SBRDs) de tipo Takagi-Sugeno-Kang (TSK) [13, 14], la estructura de regla está constituida por variables difusas en el antecedente y una función polinomial de las variables de entrada en el consecuente. Debido a este tipo de consecuente, se pierde interpretabilidad en el modelo, pero a cambio, se permite un mayor grado de precisión. El aprendizaje de las premisas y los consecuentes generalmente se realiza por separado. Lo ideal es aprender las dos partes (antecedentes y consecuentes) de forma conjunta, ya que son dependientes entre si, pero ésto representa un espacio de búsqueda muy complejo dado que cada parte tiene diferente naturaleza. Los Algoritmos Evolutivos (AEs) permiten el aprendizaje de forma conjunta, pero en problemas con alta dimensionalidad y/o con un número elevado de ejemplos, presentan serios problemas [8].

En esta contribución, se presenta un método escalable en dos fases para el modelado difuso preciso con reglas TSK. En la primera fase se propone un Algoritmo Evolutivo Multi-Objetivo (AEMO) [5, 4] basado en el aprendizaje evolutivo de la Base de Datos (BD), (que permite determinar qué variables intervienen, qué granularidades deben utilizarse y realiza un ligero desplazamiento de la partición difusa). La Base de Reglas (BR) se obtiene dentro del mismo proceso mediante un algoritmo Ad-Hoc eficiente que estima los coeficientes de la función del consecuente TSK. Este algoritmo incluye algunos mecanismos para garantizar un aprendizaje rápido del SBRDs de tipo TSK que permiten obtener la estructura del modelo y evitan la convergencia prematura en problemas con un elevado número de variables y de ejemplos. En la segunda fase se aplica un método de ajuste evolutivo combinado con una selección de reglas que incorpora un filtro de Kalman eficiente para mejorar la precisión del modelo.

Para mostrar el buen comportamiento del método propuesto se compara con el método clásico ANFIS [10] y con dos algoritmos de aprendizaje evolutivo dedicados al modelado preciso, una red neuronal (Regr-NNEP [12]) y un algoritmo de aprendizaje de SBRDs tipo TSK (LEL-TSK [3]), ambos propuestos con el objetivo de obtener los mejores resultados en precisión. Los experimentos han sido realizados sobre siete problemas con diferente número de variables y de ejemplos. El método propuesto obtiene mejores resultados en precisión en un tiempo razonable.

Este trabajo se organiza como se describe a continuación.

En la siguiente sección se introducen los SBRDs de tipo TSK. La sección 3 muestra el método propuesto describiendo sus principales características. Finalmente, las secciones 4 y 5 muestran un estudio experimental y las conclusiones obtenidas.

2 Sistemas basados en reglas difusas de tipo Takagi-Sugeno-Kang

En [13, 14], Takagi y Sugeno propusieron un modelo de SBRD que empleaba reglas en las que el antecedente estaba constituido por variables lingüísticas o difusas y el consecuente es una combinación lineal de las variables de entrada. La forma más general de este tipo de reglas es:

$$\begin{array}{c} Si\ X_1\ es\ A_1\ y\ \ldots\ y\ X_n\ es\ A_n\ entonces \\ Y = p_1 \cdot X_1 + \ldots + p_n \cdot X_n + p_0, \end{array} ,$$

donde X_i son las variables de entrada del sistema, Y es la variable de salida, p_i son parámetros reales y A_i son las etiquetas lingüísticas o conjuntos difusos. Este tipo de reglas son denominadas *reglas difusas de tipo TSK*, en alusión a sus creadores [14].

La salida de un SBRD TSK que emplee una Base de Conocimiento (BC) formada por m reglas se obtiene como la media ponderada de las salidas individuales aportadas por cada regla, Y_i, $i = 1 \ldots m$, del siguiente modo:

$$\frac{\sum_{i=1}^{m} h_i \cdot Y_i}{\sum_{i=1}^{m} Y_i},$$

donde $h_i = T(A_1(x_1), \ldots, A_n(x_n))$ es el grado de emparejamiento entre la parte antecedente de la regla i y las entradas actuales al sistema, $x = (x_1, \ldots, x_n)$. T es un operador de conjunción que se modela mediante una t-norma.

Los SBRDs TSK han sido aplicados con éxito a una gran cantidad de problemas prácticos. La ventaja principal de estos sistemas es el hecho de presentar una ecuación del sistema compacta que permite estimar los parámetros p_i empleando métodos clásicos, lo que permite obtener sistemas muy precisos, que pueden ser muy útiles para el modelado difuso preciso, especialmente cuando se usan variables difusas libres de semántica. Nosotros nos centraremos en este último tipo, SBRDs aproximativos de tipo TSK.

3 Método propuesto

En esta sección se presenta el método en dos fases propuesto para problemas de regresión con alta dimensionalidad y/o alto número de ejemplos. En la primera fase se aplica un AEMO para el aprendizaje de una BC inicial con reglas candidatas de tipo TSK de orden cero basadas en una rejilla difusa (BD) y en la segunda fase se aplica un AEMO de post-procesamiento para el ajuste evolutivo preciso. Este algoritmo incorpora algunas de las ideas del método FSMOGFS^e [2] para SBRDs de tipo Mamdani.

A continuación, se incluye una sección preliminar en la que se describirá un mecanismo para la estimación de errores y una versión modificada del método de Wang y Mendel [15] (WM) para obtener reglas TSK de orden cero. Posteriormente, se presentarán las dos fases del método propuesto.

3.1 Preliminares: Mecanismos integrados en el AEMO

En esta sección se presentan dos mecanismos utilizados en el algoritmo propuesto. El primero de ellos es un mecanismo de estimación del error que se usa en las dos fases del algoritmo. El segundo se utiliza sólo en la primera fase para la generación de la BR.

3.1.1 Mecanismo de estimación del error

Para manejar el problema de la escalabilidad con respecto al número de ejemplos disponibles en el conjunto de datos, se utiliza un mecanismo propuesto en [2] que estima el error a partir de un subconjunto reducido de los ejemplos. Este procedimiento se basa en tomar un pequeño porcentaje de los ejemplos de entrenamiento para estimar el error de las soluciones generadas. Una vez se han estimado dichos errores, se evalúan al completo sólo aquellas soluciones pertenecientes al conjunto élite (no dominadas).

El subconjunto de ejemplos E^e para la estimación de error se obtiene por selección aleatoria de $\lfloor r^e * m \rfloor$ ejemplos en cada generación. Donde r^e es el porcentaje de ejemplos usados para la estimación del error. Si $\lfloor r^e * m \rfloor \geq 1000$ *entonces* $r^e = 1000/m$, es decir, no más de 1000 ejemplos seran considerados. E^e se mantiene fija durante toda una generación. Después de cada generación los ejemplos se sustituyen por selección aleatoria de los ejemplos que no se utilizaron en la generación anterior. De esta manera, se promueve la rotación de los ejemplos seleccionados.

3.1.2 Generación eficiente de reglas TSK candidatas

Con el fin de obtener una BC completa a partir de un determinado cromosoma (que codifica una BD), se aplicará una variante del método de generación de reglas WM [15]. A diferencia de WM los consecuentes se obtienen con todos los coeficiente con valor 0 y el termino independiente se calcula como la media ponderada por el matching de los ejemplos cubiertos por la regla. Adicionalmente, se genera una regla por defecto que se utiliza en el caso de que un ejemplo se encuentre no cubierto por ninguna otra regla. Esta regla se calcula aplicando el filtro de Kalman [11] a los ejemplos cuya cobertura esta por debajo del $0,2$.

Este método en problemas con gran cantidad de variables y ejemplos consume mucho tiempo en derivar miles de re-

glas, por este motivo se ha añadido un criterio de cropping o corte en el número de reglas. De manera que WM se detiene si la BR llega a un máximo de 100 reglas y marca la BR como incompleta. Este valor (100 reglas) se fijó experimentalmente ya que los modelos obtenidos con más reglas no presentan diferencias significativas en los errores obtenidos. Con el fin de penalizar este tipo de soluciones (que deben desaparecer durante la evolución del algoritmo), se estima el número de reglas como el producto de la cantidad de etiquetas de las variables de entrada. En el caso del ECM, es penalizado con un valor muy grande.

3.2 Fase 1: AEMO para el aprendizaje de la BC

El AEMO propuesto está basado en el aprendizaje envolvente de la BD [8] (variables que intervienen, granularidades y desplazamientos de las particiones difusas) que permite una rápida obtención de la estructura del SBRD aproximativo de tipo TSK, reduciendo su dimensionalidad y haciendo uso de algunos mecanismos eficaces para asegurar una convergencia rápida.

En las siguientes subsecciones se describen sus principales características: esquema de codificación de la BD, objetivos, población inicial, operadores de cruce y mutación, mecanismo de prevención de incesto y criterio de parada.

3.2.1 Esquema de codificación de la BD

Se utiliza un esquema de codificación doble ($C = C_G + C_L$) para la *granularidad* y los *ajustes laterales*, donde:

- Número de Etiquetas (C_G): Es un vector de números enteros de tamaño N (donde N es el número de variables de entrada) en el que las granularidades de las diferentes variables se codifican, $C_G = (L^1, \ldots, L^N)$. Cada gen L^i representa el número de etiquetas usadas por los i-ésima variable y toma valores en el conjunto $\{2, \ldots, 7\}$. Puede tomar el valor igual a 1 para indicar que la variable correspondiente no se utiliza.
- Desplazamientos Laterales (C_L): Es un vector de números reales de tamaño N en el que se codifican los desplazamientos laterales de las distintas etiquetas para cada variable [1]. Así, la parte C_L tendrá la siguiente estructura (donde cada gen es el valor ajustado de la etiqueta correspondiente, y toma valores de $[-0,1;0,1]$), $C_L = (\alpha^1, \ldots, \alpha^N)$.

3.2.2 Objetivos

Para evaluar una BD codificada en un cromosoma, se decodifica la BD asociada, se aplica WM modificado para obtener la correspondiente BR, y finalmente (una vez que tenemos una BC) se calculan los siguientes objetivos:

1. Minimizar el Error Cuadratico Medio (*ECM*):

$$\text{ECM} = \frac{1}{2 \cdot |E|} \sum_{l=1}^{|E|} (F(x^l) - y^l)^2,$$

   donde $|E|$ es el tamaño del conjunto de datos, $F(x^l)$ es la salida obtenida por el SBRD para un ejemplo l y y^l es la salida esperada.
2. Miminizar el Número de Reglas (*NR*), para controlar la dimensionalidad.
3. Maximizar el grado de cubrimiento medio de los ejemplos, para controlar el sobreaprendizaje.

3.2.3 Población inicial

La población inicial está compuesta por dos subconjuntos diferentes de los individuos:

- En el primer subconjunto, cada cromosoma tiene el mismo número de etiquetas para todas las variables de entrada. Con el fin de ofrecer diversidad en la parte C_G, se han generado soluciones considerando todas las posibles combinaciones en el antecedente, es decir, de 2 a 7 etiquetas en todas las variables de entrada. Además, para cada una de las combinaciones anteriores se incluyen dos copias con diferentes valores en la parte C_L. La primera de ellas con valores aleatorios en $[-0,1;0,0]$ y el segundo con valores aleatorios en $[0,0;0,1]$. Si no hay espacio para estas soluciones, se incluyen desde las más pequeñas granularidades (las combinaciones más interesantes, en principio) a las más altas posibles.
- En el segundo subconjunto, se generan soluciones aleatoriamente para completar la población (valores en $\{2, \ldots, 7\}$ para C_G y valores en $[-0,1;0,1]$ para C_L).

Excepto en problemas con menos de tres variables de entrada, una variable de entrada v se extrae al azar, $L^v = 1$. Esta acción se repite hasta que no quedan más de 5 variables en todos los individuos. Este proceso se aplica con el fin de evitar la generación de soluciones con un número exorbitante de reglas.

3.2.4 Operadores de Cruce y Mutación

El operador de cruce depende de la parte del cromosoma en la que se aplica. Se utiliza el cruce clásico en un punto en la parte C_G y el operador Parent Centric BLX (PCBLX) [9] en la parte C_L.

De esta manera, cuatro nuevos individuos se obtienen mediante la combinación de los dos hijos generados a partir de C_G con dos hijos generados a partir de C_L. Para cada uno de ellos, el operador de mutación se aplica con una probabilidad de P_m. El operador de mutación disminuye en 1 la granularidad de un gen g seleccionadas al azar

($L^g = L^g - 1$) o aleatoriamente determina una granularidad mayor en $\{L^g + 1, \ldots, 7\}$ con la misma probabilidad. No se realiza la disminución de la granularidad cuando esto provoca BDs con sólo una variable de entrada. El mismo gen también se cambia de forma aleatoria en C_L. Finalmente, después de aplicar la mutación, sólo los dos individuos más precisos son considerados como descendientes.

3.2.5 Mecanismo de Prevención de Incesto y Criterio de Parada

Se incluye un mecanismo para la prevención de incesto en la parte C_L basado en CHC [6], para mantener la diversidad de la población y evitar la convergencia prematura. Sólo los padres cuya distancia de hamming dividida por 4 sea mayor que un umbral son cruzados. Debido a que se utiliza un esquema de codificación real en C_L, cada gen es transformado a código gray con un número fijo de bits por gen (*BGenes*). De esta forma el valor del umbral es inicializado así: $L = (\#GenesC_L * BGene)/4$, donde $\#GenesC_L$ es el número de genes de la parte C_L del cromosoma.

Este algoritmo termina cuando el número máximo de evaluaciones se alcancen o cuando L es inferior a cero.

3.3 Fase 2: AEMO de post-procesamiento para el ajuste evolutivo preciso

Una vez obtenida la BC en la fase 1, se aplica un AEMO de post-procesamiento que realiza un ajuste de las Funciones de Pertenencia (FP) y una selección de reglas, que ayuda a mejorar significativamente la precisión. Para ello se aplica una adaptación del AEMO $SPEA2_{E/E}$ [7] para trabajar con SBRDs aproximativos de tipo TSK. El nuevo AEMO propuesto utiliza el mecanismo del cálculo de errores rápido, explicado en 3.1.1.

En las siguientes subsecciones se describen los principales componentes del AEMO de post-procesamiento.

3.3.1 Esquema de codificación y Objetivos

Se utiliza un esquema de codificación triple ($C = C_T + C_S + C_C$) para los *ajustes de las FP*, la *selección de reglas,* y los *valores de los coeficientes de la función del consecuente*.

- Ajuste de las FP (C_T): es un vector de números reales que codifica los parámetros de las FP para cada regla, $C_i = (\ldots, a_1^i, b_1^i, c_1^i, \ldots, a_N^i, b_N^i, c_N^i, \ldots),\ \ i = 1, \ldots, m$, donde N es el número de variables de entrada y m es el número de reglas.
- Selección de reglas (C_S): es un vector binario que representa la existencia o no de la regla correspondiente.
- Coeficientes de la función del consecuente (C_C): es un vector de números reales de tamaño $(N+1) * m$ en el que se codifican los coeficientes de la función polinomial del consecuente TSK de cada una de las m reglas.

Este algoritmo considera los mismos tres objetivos explicados en la sección 3.2.2.

3.3.2 Población inicial

En la parte C_T la BD inicial se incluye como primer individuo y el resto de individuos se generan aleatoriamente dentro de los intervalos de variación. En la parte C_S la población inicial se obtiene con todos los individuos presentando en todos sus genes el valor '1'.

Por último, en la parte C_C se incluye como primer individuo los consecuentes obtenidos en la fase anterior. Como estos consecuentes son estimados, para obtener coeficientes optimos se aplica un filtro de Kalman sobre el individuo inicial y el resto de individuos se inicializan con esos coeficientes. No se obtienen los coeficientes de todos los individuos usando el filtro de Kalman debido a que incrementaría significativamente el tiempo de ejecución del algoritmo.

3.3.3 Operadores de cruce y mutación

Se aplica el cruce BLX-0.5 [9] para obtener la parte C_T. La parte binaria (C_S) se obtiene aprovechando la información de la parte real en padres e hijo. El padre más cercano determina si una regla es seleccionada o no para el hijo, copiando su valor en C_S [7]. El operador de mutación solo se aplica en la parte C_S y no añade reglas, directamente fija a cero un gen seleccionado aleatoriamente en C_S.

3.3.4 Aplicación eficiente del filtro de Kalman

El filtro de Kalman [11] es una técnica clásica que permite estimar los valores de los coeficientes de la función polinomial del consecuente de forma muy efectiva. Este filtro se ha hibridado con el AEMO propuesto. Si el error estimado de una nueva solución la sitúa como la solución más precisa se calculan los coeficientes de los consecuentes de las reglas aplicando dicho filtro.

El filtro de Kalman obtiene buenos resultados en entrenamiento aunque suele presentar un alto sobreaprendizaje. Una forma de evitarlo es utilizar únicamente el porcentaje reducido de ejemplos que se utiliza para la estimación del error (sección 3.1.1). Así el resto de ejemplos se utilizan cómo validación al calcular el error completo de la solución no dominada. Además, para evitar un excesivo consumo de tiempo el filtro sólo se aplica con una iteración sin reiniciar los parámetros calculados de una vez para otra, evitando así la convergencia prematura y permitiendo alcanzar los coeficientes óptimos por evolución.

3.3.5 Mecanismo de reinicio

El método se reinicia cuando L sea inferior a cero. El operador de reinicio se aplica manteniendo los tres mejores individuos en cada uno de los objetivos. Los restantes individuos toman el valor del individuo más preciso en C_S y valores

generados aleatoriamente en C_T. En la parte C_C se aplica el filtro iterativo de Kalman sobre el mejor individuo y se inicializan los restantes individuos con esos coeficientes.

4 Experimentación

Para evaluar la bondad y la escalabilidad de la técnica propuesta se han realizado experimentos considerando 7 problemas con distintas complejidades, desde 2 hasta 26 variables y desde 1056 hasta 22784 ejemplos. La tabla 1 muestra un resumen con las características de los problemas.

Tabla 1: Problemas considerados en la experimentación

Problema	**Abbr.**	**Variables**	**Ejemplos**
Presión plástico	PLA	2	1650
Coste línea eléctrica	ELE	4	1056
Edad Abalon	ABA	8	4177
Temperatura Izmir	WIZ	9	1461
Temperatura Ankara	WAN	9	1609
Precio casa	HOU	16	22784
Telecomunicaciones	POL	26	14998

Disponibles en `http://www.keel.es/`

Los métodos considerados en los experimentos son:

- **ANFIS** [10] es un método clásico de aprendizaje híbrido que permite obtener SBRDs de tipo TSK muy precisos gracias a las técnicas del gradiente descendiente y de mínimos cuadrados.
- **LEL-TSK** [3] es un algoritmo de aprendizaje evolutivo en tres etapas de SBRDs aproximativos de tipo TSK con el principal objetivo de la precisión. Incluye una etapa de generación iterativa de reglas para gestionar el espacio de búsqueda de manera eficiente.
- **Regr-NNEP** [12] es un algoritmo evolutivo para optimizar redes neuronales (aproximadores universales) de manera efectiva.
- **AEMO-TSK** propuesto en este trabajo para el modelado difuso preciso escalable.

Tanto LEL-TSK como Regr-NNEP están disponibles en *http://www.keel.es/* y consideran como único objetivo la precisión. Los valores de los parámetros considerados para los algoritmos LEL-TSK [3] y Regr-NNEP [12] son los propuestos por los autores de los métodos. Para el caso del método ANFIS [10] los parámetros considerados son: 5 etiquetas por variable, 100 iteraciones, $0,1$ tamaño del paso, $0,9$ y $1,1$ como grados de decremento e incremento del tamaño del paso, respectivamente. Los parámetros considerados para AEMO-TSK son: población de tamaño 200, población externa de tamaño 61, 100000 evaluaciones (sumando las dos fases), $0,2$ como probalidad de mutación y $r^e = 0,2$ para la estimación del error. El número de evaluaciones es 100000 en la red neuronal y más de 100000 en LEL-TSK por la naturaleza de sus etapas.

Se utiliza un modelo de validación cruzada de 5 particiones generadas aleatoriamente a partir del conjunto de datos al 20 % y se combinan 4 de ellas (80 %) para formar el conjunto de entrenamiento y la restante se usa en el test. Para cada partición los métodos se han ejecutado 6 veces, mostrando los resultados medios de 30 ejecuciones. En AEMO-TSK los valores medios se calculan considerando la solución más precisa de cada frente de Pareto.

La tabla 2 muestra los resultados obtenidos con los diferentes métodos, donde *NR* es el número de reglas, *NV* es el número de variables, $\text{ECM}_{ent/tst}$ son los errores medios en entrenamiento/test. Para el método Regr-NNEP se muestra el número de neuronas ocultas (*NOcultas*) y el número de vínculos efectivos (*NLinks*). Para ANFIS y LEL-TSK, no se muestran resultados en varios problemas debido a su alto número de variables y ejemplos (se producen problemas de memoría y/o serían necesarios muchos meses para obtener resultados).

Analizando los resultados mostrados en la tabla 2 podemos extraer las siguientes conclusiones:

- Los método ANFIS y LEL-TSK obtienen resultados muy precisos en entrenamiento, aunque presentan un serio problema de sobreaprendizaje.
- La red neuronal evolutiva (Reg-NNEP) obtiene resultados bastante competitivos tanto en entrenamiento como en test aunque no consigue alcanzar los resultados obtenidos por el método propuesto.
- El método propuesto obtiene los mejores resultados en test en todos los problemas.

Para analizar la escalabilidad de los algoritmos estudiados la tabla 3 muestra sus tiempos medios de ejecución, obtenidos en un ordenador con procesador Intel Core 2 Quad Q9550 2,83 GHz y 8GB de RAM, mediante el uso de sólo uno de los 4 núcleos. Los algoritmos ANFIS y LEL-TSK sufren un excesivo incremento en el tiempo cuando crece el número de variables e instancias que impide su ejecución en dichos problemas.

Finalmente, aunque REG-NNEP es el algoritmo más rápido, se puede observar como el algoritmo propuesto obtiene los mejores resultados en precisión en un tiempo razonable, sin verse afectado en los problemas más complejos.

Tabla 3: Tiempos medios de ejecución de los métodos

Problema	ANFIS	REG-NNEP	LEL-TSK	AEMO-TSK
PLA	00:01:09	0:18:09	0:03:14	0:05:16
ELE	39:42:50	0:09:38	0:01:57	0:10:04
ABA		2:41:28	0:24:22	0:56:49
WAN		1:40:56	0:13:48	0:17:01
WIZ		2:04:39	0:03:54	0:09:42
HOU			2:30:28	5:48:18
POL			1:20:26	5:10:00

Tabla 2: Resultados obtenidos por los métodos analizados.

	ANFIS			LEL-TSK			Regr-NNEP				AEMO-TSK			
Problema	**NR**	**ECM$_{ent}$**	**ECM$_{tst}$**	**NR**	**ECM$_{ent}$**	**ECM$_{tst}$**	**NOcultas**	**NLinks**	**ECM$_{ent}$**	**ECM$_{tst}$**	**NV**	**NR**	**ECM$_{ent}$**	**ECM$_{tst}$**
PLA	10	1,053	1,151	66	**1,032**	1,188	3,5	10,6	1,172	1,183	2,0	16	1,070	**1,137**
ELE	20	**1653**	2103471	45	2928	3752	1,5	7,2	15452	15710	4,0	36	1863	**2866**
ABA	-	-	-	107	**2,040**	2,412	3,0	20,6	2,595	2,616	4,4	33	2,168	**2,396**
WAN	-	-	-	123	**0,709**	1,632	3,9	26,7	1,304	1,435	4,1	47	0,748	**0,961**
WIZ	-	-	-	116	**0,700**	2,227	3,9	25,6	0,917	0,991	4,1	37	0,712	**0,989**
HOU	-	-	-	-	-	-	4,0	47,7	8,58E+08	9,27E+08	4,8	50	**8,48E+08**	**8,79E+08**
POL	-	-	-	-	-	-	2,9	44,8	248,92	249,95	7,0	51	**44,18**	**46,08**

5 Conclusiones

En este trabajo se presenta un método escalable en dos fases para el modelado difuso preciso mediante SBRDs aproximativos de tipo TSK para problemas de regresión con alta dimensionalidad.

El método propuesto se compara con el método clásico ANFIS y con dos algoritmos evolutivos dedicados al modelado preciso, mostrando mejores resultados en ECM en test comparando con el resto de los algoritmos analizados.

Este método es capaz de manejar problemas con un gran número de variables y de ejemplos permitiendo obtener buenos resultados en precisión en un tiempo razonable.

Agradecimientos

Soportado por el proyecto andaluz de excelencia no. P10-TIC-6858 y por el proyecto nacional no. TIN2011-28488.

Referencias

[1] R. Alcalá, J. Alcalá-Fdez, and F. Herrera. A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. *IEEE Trans. Fuzzy Syst.*, 15(4):616–635, 2007.

[2] R. Alcalá, M.J. Gacto, and F. Herrera. A fast and scalable multi-objective genetic fuzzy system for linguistic fuzzy modeling in high-dimensional regression problems. *IEEE Transactions on Fuzzy Systems*, 19(4):666–681, 2011.

[3] R. Alcalá, J. Casillas J. Alcalá-Fdez, O. Cordón, and F. Herrera. Local identification of prototypes for genetic learning of accurate tsk fuzzy rule-based systems. *International Journal of Intelligent Systems*, 22(9):909–941, 2007.

[4] C. A. Coello, D. A. Van Veldhuizen, and G. B. Lamont, editors. *Evolutionary algorithms for solving multi-objective problems*. Kluwer Academic Publishers, 2002.

[5] K. Deb. *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, NY, USA, 2001.

[6] L. J. Eshelman. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In G.J.E. Rawlin, editor, *Foundations of genetic Algorithms*, volume 1, pages 265–283. Morgan Kaufman, 1991.

[7] M.J. Gacto, R. Alcalá, and F. Herrera. A multi-objective evolutionary algorithm for an effective tuning of fuzzy logic controllers in heating, ventilating and air conditioning systems. *Applied Intelligence*, in press. doi: 10.1007/s10489-010-0264-x.

[8] F. Herrera. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence*, 1:27–46, 2008.

[9] F. Herrera, M. Lozano, and A. M. Sánchez. A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study. *International Journal of Intelligent Systems*, 18:309–338, 2003.

[10] J. S. R. Jang. ANFIS: Adaptive network based fuzzy inference system. *IEEE Trans. Syst., Man, Cybern.*, 23(3):665–684, 1993.

[11] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[12] A.C. Martínez-Estudillo, F.J. Martínez-Estudillo, C. Hervás-Martínez, and N. García. Evolutionary product unit based neural networks for regression. *Neural Networks*, 19(1):477–486, 2006.

[13] M. Sugeno and G.T. Kang. Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 28:15–33, 1988.

[14] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst., Man, Cybern.*, 15(1):116–132, 1985.

[15] L. X. Wang and J. M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. Syst., Man, Cybern.*, 22(6):1414–1427, 1992.

ALGORITMO GENÉTICO PARA DECISIONES INTELIGENTES EN CRUCES DE VEHÍCULOS CON FUNCIÓN DE COSTE DINÁMICA

E. Onieva[1], **J. Villagrá**[1], **V. Milanés**[1], **J. Godoy**[1]

[1]*Programa AUTOPIA, Centro de Automática y Robótica (UPM-CSIC) La Poveda-Arganda del Rey, 28500, Madrid.*,
{enrique.onieva,jorge.villagra,vicente.milanes,jorge.godoy}@csic.es

Resumen

Este trabajo analiza el caso donde un vehículo autónomo debe cooperar con otro manual para llevar a cabo una maniobra de cruce. La dificultad del escenario es la unidireccionalidad de la cooperación a llevar a cabo por el vehículo autónomo, dado que no puede asumir que el manual respetará las normas básicas de circulación en cruces.

La velocidad a seguir por el vehículo autónomo es calculada por un sistema basado en reglas difusas (SBRD); que será optimizado por medio de un algoritmo genético. La función de coste será calculada mediante la simulación incremental de escenarios de cruce, y evaluando si las decisiones tomadas son adecuadas para el escenario. Una vez calculada la velocidad a seguir, ésta se manda a una capa de bajo nivel, encargada de accionar los pedales del vehículo.

Palabras Clave: Algoritmos Genéticos, Lógica Difusa, Control de Vehículos, Cruces Inteligentes.

1 INTRODUCCIÓN

Los accidentes de tráfico en intersecciones suponen un alto coste en términos de muertes, lesiones y daño a la propiedad. La seguridad en intersecciones es vista como un área de investigación con alta prioridad, dado que los accidentes en intersecciones suponen entre un 25 % y 45 % del total de accidentes [7] y, dado que las intersecciones representan una proporción pequeña del trazado viario, ésto supone una cantidad altamente desproporcionada.

Los Sistemas Inteligentes de Transporte aplican técnicas de robótica e inteligencia artificial para lograr una conducción segura y eficiente [2]. Los estudios en conducción autónoma pueden dividirse en: (i) control de velocidad [9], sobre el acelerador y freno (y en algunos casos el cambio de marcha); (ii) control de dirección, sobre el volante [6].

En carretera abierta, los vehículos pueden ser completamente autónomos mediante una combinación apropiada de los sistemas mencionados. No obstante, hay una ligera necesidad de incluir comportamientos más avanzados cuando tratamos con entornos urbanos [3]. En estos entornos, puede darse el caso en que, aún actuando de manera *óptima* el vehículo se vea involucrado en una situación en la que un accidente es inevitable.

El presente trabajo está motivado por la futura necesidad de convivencia entre vehículos manuales y autónomos, y de cómo estos últimos deberán actuar bajo cooperación incompleta. Para ilustrar ésto, proponemos un escenario en el que dos vehículos se aproximan a una intersección sin señalizar; uno de ellos conducido por una persona y otro completamente autónomo. Las acciones a llevar a cabo por el primero son desconocidas para el autónomo, que únicamente tendrá la visión del estado actual del otro y deberá actuar unilateralmente para evitar riesgos. En el caso de dos vehículos autónomos, ambos llevarían a cabo una cooperación, cediendo uno de ellos el paso, o adaptando sus velocidades [5]. Pero, dado que uno de ellos es conducido manualmente, ésta cooperación se hace imposible.

Por ello proponemos un sistema de dos capas capaz de: (i) calcular la velocidad a seguir por el vehículo autónomo para realizar la maniobra de forma segura y eficaz y (ii) actuar sobre los pedales de éste para seguir la velocidad dada por la capa anterior.

La primera estará implementada por medio de un sistema basado en reglas difusas (SBRD) [10] de cuatro entradas referentes al estado del escenario.

Dichos controladores serán optimizados por medio de un algoritmo genético [4], que los evaluará en un conjunto de escenarios de cruce y calculará un coste parcial de acuerdo con resultados de simulación. Por ejemplo: si el con-

trolador provoca una colisión, éste será penalizado; en caso contrario, será recompensado en función de su efectividad.En total se optimizarán 81 estructuras de controladores, con diferentes granularidades para las entradas; por otra parte, la velocidad a seguir será dada en términos relativos (incremento de velocidad) o absolutos (velocidad a seguir); obteniéndose así 162 tipos de controladores.

La principal novedad en el diseño del sistema presentado radica en la elección de la función de coste. Ésta utiliza un conjunto de escenarios creciente y cambiante generación tras generación, con lo que se consigue reducir el tiempo de ejecución, así como maximizar los casos a considerar por el controlador. La función de coste evaluará al seguridad y eficacia de los controladores en cada uno de los escenarios.

2 ARQUITECTURA DE CONTROL PROPUESTA

Supondremos que el sistema de control dispone de las posiciones, direcciones y velocidades de ambos vehículos. En caso de no disponer de un mapa que nos indique dónde se encuentra la intersección, las distancias de ambos vehículos al punto de cruce deben ser calculadas.Las siguientes subsecciones detallan las dos capas de la arquitectura de control.

2.1 Determinación de la Velocidad Adecuada

Ambas distancias al punto de cruce ($D_{\{A|M\}}$) y velocidades ($S_{\{A|M\}}$) son utilizadas como entradas de un SBRD, éstas están representadas gráficamente en la figura 1.

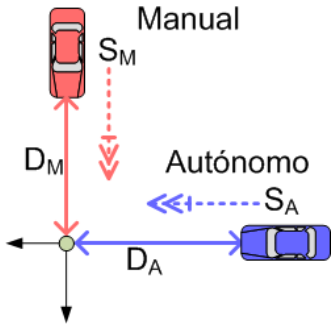


Figura 1: Cálculo del punto de cruce de ambos vehículos.

Todas las variables de entrada se codifican por medio de etiquetas triangulares uniformemente distribuidas en los intervalos $[-10,100]m$ para las distancias y $[0,50]km/h$ para las velocidades. Cada variable tiene su propio número de etiquetas a usar en la fuzzificación ($\#D_M$, $\#D_A$, $\#S_M$, $\#S_A$).

La base de reglas ($\#D_M \cdot \#D_A \cdot \#S_M \cdot \#S_A$) reglas formadas mediante la composición (operador Y) de todas las variables. Así pues, la estructura del conjunto de antecedentes quedaría de la forma:

SI $(D_M = D_M^i)$ Y $(D_A = D_A^j)$ Y $(S_M = S_M^k)$ Y $(S_A = S_A^l)$ donde $i = \{1 \ldots \#D_M\}$, $j = \{1 \ldots \#D_A\}$, $k = \{1 \ldots \#S_M\}$ y $l = \{1 \ldots \#S_A\}$.

La salida del sistema, llamada *out*, podrá representar: (i) La modificación requerida sobre la velocidad actual del vehículo. $S_{ref} = S_A + out$. (ii) La velodidad absoluta necesaria. $S_{ref} = out$. La salida será codificada por medio de $\#out$ singletons.

Finalmente, la asignación de singletons de salida a antecedentes de la base de reglas es codificada por medio de un vector, para así poder ser manejado por el algoritmo genético: $Ind_i \in \{1,...\#out\}, \forall i \in \{1,...\#D_M \cdot \#D_A \cdot \#S_M \cdot \#S_A\}$. Donde el valor asignado a Ind_i significa usar dicho singleton como consecuente de la regla número i.

2.2 Comportamiento Longitudinal del Vehículo

El comportamiento longitudinal de un vehículo autónomo puede ser aproximado (para dinámicas lentamente cambiantes y superficie plana) por la función de transferencia de segundo orden:

$$G(s) = \frac{K\omega_n^2}{s^2 + 2\eta\omega_n + \omega_n^2}$$

donde $K = 25.14$, $\eta = 160$, y $\omega_n = 55.87$ (véase [8]).

Para alcanzar las velocidades de referencia dadas por el SBRD, se implementa un controlador PI que, aparte de ser una técnica sencilla y bien conocida [1], permite describir el comportamiento en el dominio de Laplace como $C(s) = K_P E(s) + K_I \frac{E(s)}{s}$, siendo $E(s)$ el error de velocidad y usando $K_P = 0.3$ y $K_I = 0.1$ como ganancias.

Finalmente, la dinámica del sistema en lazo cerrado es discretizada y reescrita como la función lineal:

$$\begin{aligned} S_A(t_k) =& a_2 S_{ref}(t_{k-1}) + a_1 S_{ref}(t_{k-2}) + a_0 S_{ref}(t_{k-3}) - \\ & - b_2 S_A(t_{k-1}) - b_1 S_A(t_{k-2}) - b_0 S_A(t_{k-3}) \end{aligned}$$

con $a_0 = -5.467 \cdot 10^{-5}$, $a_1 = -0.2041$, $a_2 = 0.2495$, and $b_0 = 0$, $b_1 = 0.7421$, $b_2 = -1.697$, y donde $S_{ref}(t_k)$ y $S_A(t_k)$ representan las velocidades de referencia y actuales del vehículo en el instante t_k.

3 PROCESO DE OPTIMIZACIÓN GENÉTICO

El algoritmo genético optimiza la asignación de consecuentes al conjunto completo de antecedentes del SBRD (sección 2.1).

Se utiliza un esquema clásico de algoritmo genético estacionario, el cual, en cada generación, selecciona dos padres, aplica sobre ellos los operadores de cruce y mutación para generar dos hijos y reemplaza (en caso) dos individuos de la generación actual con ellos.

El proceso comienza con una población de 100 individuos generados aleatoriamente y se ejecuta durante 1000 generaciones. Los operadores implementados son:

- *Selección estocástica*: proporcional al valor de la función objetivo de un individuo.
- *Cruce uniforme*: para cada posición del vector hijo, se hace una selección aleatoria sobre de cuál de los padres el valor debe ser heredado. El segundo hijo es generado usando las elecciones inversas.
- *Mutación simple*: cada gen se cambia por un valor permitido con $p_m = 2/(\#D_M \cdot \#D_A \cdot \#S_M \cdot \#S_A)$.
- *Reemplazamiento*: cada uno de los hijos generados reemplazará a uno de sus padres en caso de que el valor de la función objetivo mejore.

3.1 Función Objetivo

Para evaluar el comportamiento de un controlador, éste será ejecutado en un número creciente de escenarios:

$$N_{sc} = 1 + 19 \cdot \lfloor g/G \rfloor$$

, representando g y G, respectivamente, la generación actual y el número total de generaciones. Dado que el algoritmo se ejecuta durante 1000 generaciones, se generarán un total de 10510 escenarios diferentes, a lo largo de todo el proceso.

En los escenarios, ambos vehículos se aproximan perpendicularmente al punto de cruce comenzando ambos a distancias y con velocidades aleatorias:

$$\{D_M, D_A\} = U(50,80) \quad \{S_M, S_A\} = U(10,40)$$

Cada escenario evolucionará dos veces:

1. Sin considerar las referencias dadas por el SBRD (ejecución libre o E_l), donde en vehículo mantiene constante su velocidad.
2. Siguiendo las referencias (ejecución controlada o E_c).

El propósito de E_l es el de determinar si se produciría una colisión si el vehículo autónomo no variara su velocidad; mientras que el de E_c es determinar si dicha colisión es evitada gracias a las indicaciones del SBRD, y su eficacia.

Cada una de las evoluciones del escenario inicial puede llevar a tres resultados diferentes según el momento en que el vehículo autónomo entra/sale de una *zona segura* definida 5 metros antes y después de punto de cruce:

1. Los vehículos no coinciden en la zona segura (no hay colisión, C_0). En este caso se espera que el SBRD haga que el vehículo no altere la velocidad.

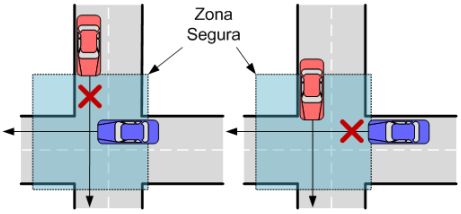


Figura 2: Situación con colisión lateral (izquierda) y frontal (derecha).

2. Se produce una colisión lateral, es decir ambos coinciden en la zona segura pero el autónomo ha entrado antes, C_L. Aquí se espera que el SBRD aumente ligeramente la velocidad, para salir de la zona segura antes de que el manual entre.
3. Se produce una colisión frontal ambos coinciden en la zona segura, habiendo entrado el manual antes, C_F. En este caso el SBRD deberá reducir la velocidad, para entrar en la zona segura una vez el manual haya salido.

Los casos que implican colisión se muestran gráficamente en la figura 2.

Dado que el esquema de generación de situaciones iniciales hace que sea más probable (sobre el 75 % de las veces) que se genere un escenario sin colisión en ejecución libre ($E_l = C_0$), dicha generación es restringida a que se generen situaciones que evolucionen a cada uno de los tres posibles casos con la misma probabilidad:

$$P(E_l \to C_0) = P(E_l \to C_L) = P(E_l \to C_F) = 1/3$$

La función objetivo será calculada por medio de la suma de objetivos parciales, representando cada uno de ellos la penalización obtenida en cada escenario inicial:

$$Fitness = \sum_{i=1,\ldots N_{sc}} F(E_l^i, E_c^i)$$

donde el resultado parcial se calcula en función de las nueve posibles combinaciones de resultados ($E_{l|c}^i = C_{O|L|F}$) de la siguiente manera:

- No se produce colisión alguna ($E_l^i = E_c^i = C_0$); entonces la penalización será menor cuanto menos se haya alterado la velocidad inicial, usando $F = |\int S_A^l - \int S_A^c|$. En la figura 3 se ilustra este criterio.
- El SBRD evitan una colisión lateral ($E_l^i = C_L \; Y \; E_c^i = C_0$); la penalización será de $F = |\int S_A^l - \int S_A^c|$ si se ha aumentado la velocidad inicial ($\int S_A^c > \int S_A^l$) y de $F = 2500$ en caso contrario.
- La situación especular se produce cuando ($E_l^i = C_F \; Y \; E_c^i = C_0$). En este caso, la penalización será de $F = |\int S_A^l - \int S_A^c|$ si se ha reducido la velocidad ($\int S_A^c < \int S_A^l$) y de $F = 2500$ en caso contrario.

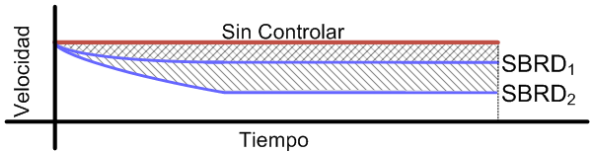


Figura 3: Ejecución de dos controladores: $SBRD_1$ frena menos que $SBRD_2$, por tanto recibe menor $|\int S_A^f - \int S_A^c|$.

- El SBRD no evita una colisión ($E_l^i = C_{L|F}\ Y\ E_c^i = C_{L|F}$); en este caso se penaliza con $F = 5000$.
- El SBRD produce una colisión ($E_l^i = C_0\ Y\ E_c^i = C_{L|F}$); en este caso se aplica la máxima penalización de $F = 10000$, dado que la colisión se ha producido por seguir las consignas del SBRD.

Las penalizaciones constantes $F = \{2500, 5000, 10000\}$ se establecen a partir del hecho que, en una ejecución promedio en la que el SBRD detiene inmediatamente el vehículo, se obtendría $F = |\int S_A^l - \int S_A^c| \approx 2000$ por eso, 2500 representa ligeramente más penalización que dicha ejecución, 5000 representa algo más del doble de penalización y 10000 representa unas cinco veces más penalización que un SBRD promedio con política de *detenerse siempre*. Estos valores se establecieron empíricamente.

El objetivo del proceso de ajuste será el de minimizar la suma de las funciones objetivo parciales. Finalmente resaltar que, dado que la función objetivo no es determinística (el número y situación inicial de escenarios cambia en cada generación), cada hijo es evaluado en los mismos escenarios que sus padres. Por otra parte el proceso de evaluación se repite en cada generación sobre toda la población, dado que un individuo con buen comportamiento en los escenarios de una generación no tiene por qué ser bueno en los de la siguiente.

4 EXPERIMENTACIÓN Y RESULTADOS

El número de etiquetas triangulares para codificar cada una de las variables de entrada utilizado fue 2, 3, y 4, o, de manera formal:

$$\{\#D_M, \#D_A, \#S_M, \#S_A\} \in \{2,3,4\}^4$$

, generándose así, 81 configuraciones de granularidad.

El número de consecuentes usados, así como sus rangos dependen del tipo de controlador a implementar:

- Para controladores que infieren la velocidad de manera relativa tenemos que $\#out = 3$, llamados: *frenar* en -10, *mantener* en 0 y *acelerar* en 10.
- Para controladores que devuelven la velocidad absoluta a seguir tenemos que $\#out = 4$, llamados: *para* en 0, *despacio* en 10, *medio* en 20 y *rápido* en 30.

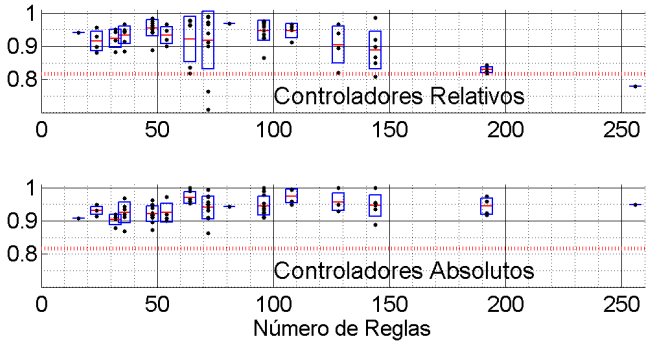


Figura 4: Seguridad de los Controladores. Las líneas rojas representan la media para un determinado número de reglas, las azules la desviación estándar. La línea roja punteada muestra el valor en ejecución sin controlar.

Cada ejecución del algoritmo genético fue repetida 10 veces. De ahora en adelante, usaremos $(REL|ABS)_{(\#D_M,\#D_A\#S_M,\#S_A)}$ como resultado promedio de controladores de relativos y absolutos, respectivamente.

En total, el proceso se ejecuta 162 veces usando las 2 clases de SBRD (REL y ABS) y las 81 granularidades.

Finalmente, los controladores resultantes fueron probados en escenarios generados mediante combinación de distancias al punto de cruce y velocidades iniciales, de la forma:

$$\begin{aligned} [D_A, D_M]_0 &\in \{50,55,60,65,70,75,80\}^2 \\ [S_M, S_A]_0 &\in \{10,15,20,25\}^2 \end{aligned}$$

haciendo así, un total de $7 \cdot 7 \cdot 4 \cdot 4 = 784$ escenarios se prueba . Los resultados serán comparados en términos de seguridad y eficacia en la realización de la maniobra.

4.1 Seguridad de los Controladores

En la figura 4 se muestra el porcentaje de escenarios finalizados sin colisión (sobre los 784), con los SBRD agrupados por número de reglas.

A primera vista puede decirse que con los controladores absolutos se obtienen mejores resultados; por otra parte, algunos controladores relativos obtienen peores resultados que no actuar sobre la velocidad inicial del vehículo. Cuatro de ellos obtienen una seguridad del 100%, que son: ABS_{4423}, ABS_{3433}, ABS_{3344} y ABS_{2442}.

Esta aparente dependencia entre seguridad y *tipo de controlador* puede llevar a confusión. Por ello, la figura 5 muestra la seguridad de los controladores en un plano, donde los ejes X-Y representan la seguridad de una determinada granularidad para controladores del tipo ABS y REL, respectivamente. No existen diferencias significantivas entre granularidades que funcionan mejor con un esquema que con otro (46% frente a 54%); por ello se deduce que algunas configuraciones de granularidad se comportan mejor

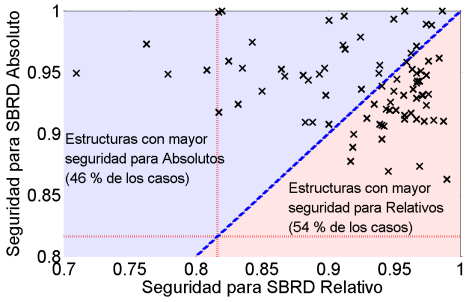


Figura 5: Correlación entre seguridad obtenida para una granularidad. La línea azul limita la región de no diferencia y las rojas, la seguridad en ejecución libre.

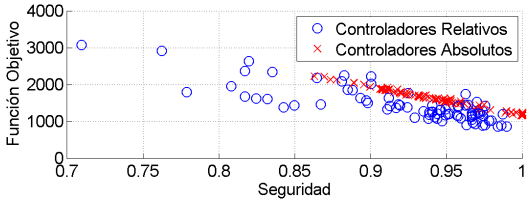


Figura 6: Correlación entre seguridad y función objetivo.

que otras, independientemente del tipo de salida implementada.

4.2 Eficacia de los Controladores

Esta sección pretende analizar los resultados para garantizar que los controladores, aparte de seguros, son eficaces. Por ejemplo: *detener siempre* el vehículo tendría máxima seguridad (nunca colisionaría), pero sería ineficaz.

En la figura 6 se muestra la correlación entre seguridad y eficacia, medida en términos de función objetivo (sección 3.1), ya que ésta recompensa individuos que llevan a cabo la maniobra sin provocar colisión, en el menor tiempo posible y respetando siempre el orden en que deben cruzar los vehículos. Se puede observar que el valor de función objetivo decrece casi linealmente conforme aumenta la seguridad; lo que implica que los individuos con bajo nivel de riesgo están actuando de manera *inteligente* para evitar colisiones.

4.3 Evaluación Cualitativa

En esta sección se pretende evaluar la manera en que ciertos SBRD evitan situaciones de riesgo. Se analizarán los sistemas $\{ABS|REL\}_{\{3344|4242\}}$ dado que la seguridad de $ABS_{\{3344|4242\}}$ es del 100 %, mientras que la de $REL_{\{3344|4242\}}$ es del 98 % y 99 %; por lo que parece que usar las granularidades $\{3344|4242\}$ es lo más acertado para el diseño de SBRD para llevar a cabo la maniobra.

Estos cuatro SBRD se prueban en cuatro situaciones, donde ambos vehículos comienzan a $50m$ del punto de cruce ($D_{M|A} = 50$), con diferentes velocidades. Las evoluciones se muestran en las figuras 7 a 10, que pasamos a analizar en adelante. En dichas figuras, la línea azul representa la evolución de la distancia en ejecución libre; las líneas rojas representan valores del vehículo manual; la franja gris representa la zona segura ($\pm 5m$) y, finalmente, el área rosada representa el intervalo durante el cual la zona segura es ocupada por el vehículo manual.

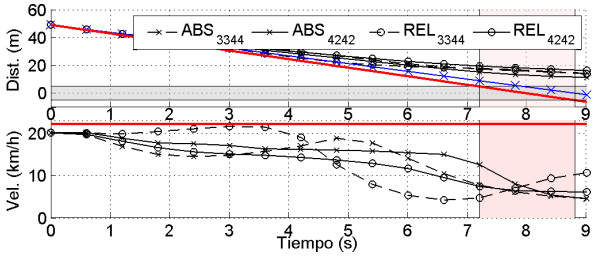


Figura 7: Evolución con el vehículo autónomo comenzando ligeramente más despacio ($E_l = C_F$).

La figura 7 muestra un experimento que llevaría a una colisión frontal, ya que el vehículo manual entra primero al área de cruce. No se encuentran diferencias mayores entre los comportamientos mostrados por los controladores. Todos ellos reducen la velocidad para evitar la colisión (sin detenerse), siendo REL_{3344} el único que acelera una vez pasado el riesgo.

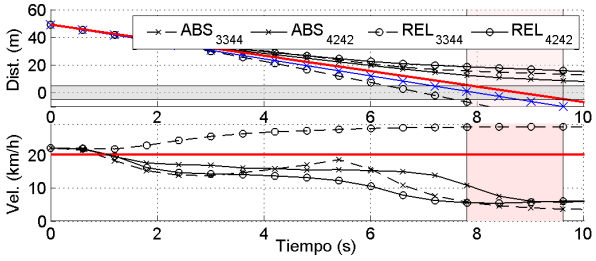


Figura 8: Evolución con el vehículo autónomo comenzando ligeramente más rápido ($E_l = C_L$).

En el segundo caso (figura 8) se produciría una colisión lateral, por lo que se espera que el vehículo autónomo acelere para abandonar la zona de riesgo antes de que entre el manual; únicamente REL_{3344} muestra este comportamiento mientras el resto, aunque consiguen evitar la colisión, lo hacen por medio de reducir la velocidad del vehículo.

Finalmente, las figuras 9 y 10 representan situaciones donde no existe riesgo de colisión, dado que el vehículo autónomo circula a mayor o menor velocidad que el manual de manera considerable y ambos no coinciden en la zona de riego. En estos casos se espera que los controladores varíen en la menor medida de lo posible la velocidad; lo cual es obtenido en mayor medida por REL_{3344}. Por tanto es éste controlador el que parece implementar el comportamiento que consideramos más inteligente, de una manera global.

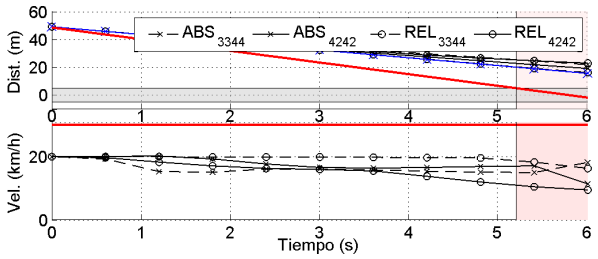


Figura 9: Evolución con el vehículo autónomo comenzando más despacio; sin riesgo de colisión ($E_l = C_0$).

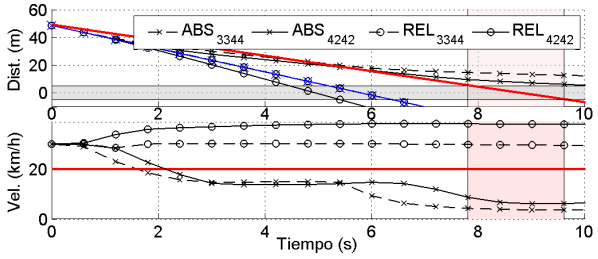


Figura 10: Evolución con el vehículo autónomo comenzando más rápido; sin riesgo de colisión ($E_l = C_0$).

5 Conclusiones y Trabajos Futuros

Este trabajo presenta un proceso de optimización genético de controladores difusos. Su objetivo es guiar un vehículo al aproximarse a un cruce donde otro vehículo no coopera para hacer la maniobra de manera óptima.

Está estructurado en dos capas. La primera de ellas fue un SBRD encargado de calcular la velocidad apropiada para circular sin riesgo, el cual fue objeto de optimización. La segunda representa la capa de bajo nivel del vehículo; la evolución de la dinámica de éste de acuerdo con las consignas recibidas.

Se han optimizado un total de 162 SBRD, con 81 granularidades de las variables de entrada y 2 significados de la salida (velocidad absoluta o relativa).

Para evaluar individuos, se implementa una función que simula la evolución del vehículo sin controlarlo y controlado en un número creciente de escenarios, para así considerar si: (i) el controlador evita una colisión que se hubiera producido sin él; (ii) el controlador no puede evitar dicha colisión; (iii) el controlador provoca una colisión que no se hubiera producido sin él. Con esto, la función objetivo valora la seguridad y eficacia del controlador durante la maniobra, reconpensándolo o penalizándolo en consecuencia.

Los controladores obtenidos son sometidos a un amplio estudio, evaluándolos en 784 escenearios y comparándolos en términos de seguridad y eficacia, para así asegurarnos que dichos controladores no implementan una política de *parar siempre*, la cual sería segura, pero poco inteligente.

Trabajos futuros irán orientados en la línea de la implementación de controladores más complejos para, por ejemplo, controlar también el volante del vehículo en caso de colisión inevitable, y realizar un giro para evitarlo (o minimizar consecuencias). Finalmente, se espera coordinar el presente sistema con otros propios de vehículos autónomos, como control de distancia con un vehículo precedente.

Agradecimientos

Este trabajo ha sido realizado gracias al CYCIT y al Plan Nacional a través de sus proyectos GUIADE (P9/08) y TRANSITO (TRA2008-06602-C03-01).

Referencias

[1] K. Aström and T. Hägglund: Advanced PID Controllers. *International Society of Automation*, 2006.

[2] R. Bishop: Intelligent vehicle technology and trends. *Artech House*, 2005.

[3] K. Dresner and P. Stone: A multiagent approach to autonomous intersection management. *Journal of Artificial Intelligence Research*, 31(1), pp. 591–656, 2008.

[4] D. Goldberg: Genetic algorithms in search, optimization, and machine learning. *Addison-wesley*, 1989.

[5] V. Milanés, J. Pérez, E. Onieva, and C. González: Controller for urban intersections based on wireless communications and fuzzy logic. *IEEE Transactions on Intelligent Transportation Systems*, 11(1), pp. 234–248, 2010.

[6] J. Pérez, V. Milanés, and E. Onieva: Cascade architecture for lateral control in autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 12(1), pp. 73-82, 2011.

[7] M. Staubach: Factors correlated with traffic accidents as a basis for evaluating advanced driver assistance systems. *Accident Analysis and Prevention*, 41(5), pp. 1025–1033, 2009.

[8] I. Tejado, V. Milanes, J. Villagra, H. Godoy, J. HosseinNia, and B. M. Vinagre: Low speed control of an autonomous vehicle by using a fractional pi. *2011 IFAC World Congress*. 2009.

[9] B. Yang, L. Keqiang, and F. Nenglian: Feedback Linearization Tracking Control of Vehicle Longitudinal Acceleration Under Low-Speed Conditions. *Journal of Dynamic Systems, Measurement, and Control*, 130(5), pp. 51–62, 2008.

[10] L. Zadeh: Fuzzy sets. *Information and Control*, 8(3), pp. 338-353, 1965.

UN OPERADOR DE CONJUNCIÓN ADAPTATIVO PARA MODELADO DIFUSO LINGÜÍSTICO DE PROBLEMAS DE ALTA DIMENSIONALIDAD

César Serrano[1], Antonio A. Márquez[1], Francisco A. Márquez [1], Antonio Peregrín [1]

[1] *Dept. de Tecnologías de la Información, DTI, Universidad de Huelva*, cesar.serrano@alu.uhu.es, {amarquez, alfredo.marquez, peregrin}@dti.uhu.es

Resumen

Los operadores de conjunción adaptativos mejoran la precisión de los Sistemas basados en Reglas Difusas. Sin embargo, su funcionalidad se ve dificultada en problemas de alta dimensionalidad debido a que el aprendizaje de sus parámetros, por su aumento, se hace mucho más difícil. En el presente trabajo presentamos un nuevo operador de conjunción adaptativo que permite obtener mejores resultados para este tipo de problemas con respecto a los operadores de conjunción adaptativos más utilizados en la literatura. La mejora se produce en dos aspectos importantes en este tipo de sistemas: la precisión y la escalabilidad. Los resultados obtenidos sobre 5 problemas de regresión de distintas características muestran la efectividad de la propuesta que se plantea en este trabajo.

Palabras Clave: Sistema de Inferencia Adaptativo, Operador de Conjunción Adaptativo, Problemas de Alta Dimensionalidad, Algoritmos Genéticos.

1 INTRODUCCIÓN

Uno de los retos más importantes hoy en día, en el modelado difuso lingüístico es el aprendizaje de sistemas para problemas de alta dimensionalidad. Esto es debido a que los algoritmos de obtención de reglas usuales utilizados en los Sistemas Basados en Reglas Difusas (SBRDs) generan un número desmesurado de éstas cuando el número de variables y ejemplos son elevados [3], [12]. Por tanto, se dificulta el aprendizaje de los SBRDs ya que un número elevado de reglas provoca un crecimiento crítico en el espacio de búsqueda, dando lugar a dos problemas [9]:

- Complejidad del espacio de búsqueda, debido al gran numero de variables consideradas.
- Escalabilidad, originado sobre todo por el gran número de instancias implicadas en estos conjuntos de datos.

Por otro lado, el uso de operadores adaptativos en los SBRDs para modelado y control, tanto en el Sistema de Inferencia [1] como en la Interfaz de Defuzzificación [4], o en ambos de modo simultáneo [13], ha demostrado ofrecer mejoras importantes en la precisión de los mismos.

Centrándonos en el Sistema de Inferencia Adaptativo, lo más frecuente es utilizar operadores de conjunción adaptativos [1]. Éstos normalmente emplean uno o más parámetros en sus expresiones con el fin de modificar su comportamiento, siendo habitual utilizar un parámetro por cada regla y por tanto adaptándose localmente.

En este trabajo proponemos un operador de conjunción adaptativo evolutivo que permite reducir el espacio de búsqueda y por tanto una más rápida convergencia, adecuado para problemas de alta dimensionalidad.

Los resultados obtenidos son comparados con dos de los operadores de conjunción adaptativos más utilizados en la literatura, las t-normas parametrizadas de Dubois y de Dombi [7, 14], sobre 5 problemas de distintas características, mostrando la efectividad de la propuesta.

La organización del trabajo es como sigue: en la siguiente sección se introducen brevemente los operadores de conjunción adaptativos utilizados en la literatura y su relación con los operadores de conjunción clásicos. En la Sección 3 se presenta nuestra propuesta, describiendo en detalle el modelo de aprendizaje evolutivo que utiliza. La Sección 4 muestra el comportamiento del modelo propuesto en cinco problemas. Por último, en la Sección 5, se muestran las conclusiones del trabajo.

2 PRELIMINARES: OPERADOR DE CONJUNCIÓN ADAPTATIVO

Los SBRDs para modelado utilizan reglas SI - ENTONCES de la forma siguiente:

$$R_i : Si\ X_{i1}\ es\ A_{i1}\ y\ ...\ y\ X_{im}\ es\ A_{im}\ entonces\ Y\ es\ B_i$$

con i = 1 hasta N reglas, siendo X_{i1} hasta X_{im} las entradas e Y la salida, y con A_{i1} hasta A_{im} y B_i los antecedentes y consecuentes respectivamente.

La expresión de la Regla Composicional de Inferencia en el modelado difuso con fuzzificación puntual es mostrada en la expresión (1), donde $\mu_{B'}$ es la función de pertenencia del consecuente inferido, I es el operador de implicación, C el operador de conjunción, μ_{Ai} al tratarse de fuzzificación puntual son los puntos de corte de las entradas discretas ($x_{1,...,}x_n$) del sistema con las funciones de pertenencia de los antecedentes de las reglas, y μ_B el consecuente de la regla.

$$\mu_{B'}(y) = I\,(C\,(\mu_{A1}(x_1)\,,\ ...\ ,\ \mu_{An}(x_n)),\ \mu_B(y)) \qquad (1)$$

Por tanto, el sistema de inferencia emplea dos operadores: el de conjunción, C(.) para calcular el grado de emparejamiento, y el de implicación, I(.) para calcular a continuación el consecuente inferido.

En [4, 8] se estudiaron las diferentes formas de parametrización del Sistema de Inferencia observándose menor influencia en el uso de operadores de implicación adaptativos frente al de los de conjunción. Por este motivo, en este trabajo se utiliza la parametrización del Sistema de Inferencia exclusivamente mediante el operador de conjunción.

Las T-normas son una familia de operadores generalmente usados en el diseño de SBRDs, modelando tanto el operador de conjunción como el de implicación [8] y ofreciendo buen comportamiento [5]. Las expresiones de las T-normas clásicas más utilizadas son:

Producto Lógico (minimo): $T_{Min}(x,y) = Min\,(x, y)$ (2)

Producto de Hamacher: $T_{Ham}(x,y) = \dfrac{x \cdot y}{x + y - x \cdot y}$ (3)

Producto Algebraico: $T_{Alg}(x,y) = x \cdot y$ (4)

Producto de Einstein: $T_{Eins}(x,y) = \dfrac{x \cdot y}{1 + (1-x)\cdot(1-y)}$ (5)

Producto Acotado: $T_{Aco}(x,y) = Max(0, x + y - 1)$ (6)

Producto Drástico: $T_{Dras}(x,y) = \begin{cases} x, & si\ y = 1 \\ y, & si\ x = 1 \\ 0, & en\ otro\ caso \end{cases}$ (7)

En este trabajo utilizaremos dos de las t-normas parametrizadas o adaptativas más utilizadas en la literatura [10,14]. Sus expresiones son:

Dombi ($\alpha>0$)
$$T_{Dombi}(x,y) = \frac{1}{1 + \sqrt[\alpha]{\left(\frac{1-x}{x}\right)^{\alpha} + \left(\frac{1-y}{y}\right)^{\alpha}}} \qquad (8)$$

Dubois ($0 \le \alpha \le 1$)
$$T_{Dubois}(x,y) = \frac{x \cdot y}{Max(x, y, \alpha)} \qquad (9)$$

La tabla 1 muestra la relación entre las seis t-normas clásicas y las dos t-normas adaptativas citadas anteriormente, dependiendo del valor del parámetro α.

Tabla 1: Relación entre las t-normas clásicas y las t-normas adaptativas en función de su parámetro α

	$T_{Mín}$	T_{Ham}	T_{Alg}	T_{Ein}	T_{Aco}	T_{Dras}
T_{Dombi}	∞	1				$\to 0$
T_{Dubois}	0		1			

El operador de conjunción adaptativo en definitiva, modula la influencia del grado de emparejamiento atenuándolo o acentuándolo. Esta modificación es no lineal debido a la expresión de las t-normas adaptativas.

Dos modelos de conjunción adaptativa podrían considerarse dependiendo de si se utiliza un sólo parámetro α, o distintos parámetros α_i para cada regla. Un sólo parámetro permite ajustar globalmente el comportamiento del conectivo mientras que los parámetros individuales ajustan dicho comportamiento para cada regla de la Base de Reglas (BR), es decir, permiten adaptar localmente la conjunción para cada regla. El modelo de un sólo parámetro, al ser menos flexible ofrece peores resultados en cuanto a precisión [1], por lo que en este trabajo se utiliza el modelo de adaptación local.

3 CAAD: UN OPERADOR DE CONJUNCION ADAPTATIVO PARA PROBLEMAS DE ALTA DIMENSIONALIDAD

En esta sección, proponemos introducir un operador de Conjunción Adaptativo para Problemas de Alta Dimensionalidad (CAAD) y mostraremos el modelo de aprendizaje evolutivo propuesto.

Tal y como se ha indicado en las secciones anteriores, los inconvenientes derivados de la utilización de operadores de conjunción adaptativos sobre problemas de alta dimensionalidad son consecuencia de la gran cantidad de reglas

derivada por el elevado número de variables y datos que tienen estos problemas.

Por otro lado, los métodos de aprendizaje empleados para aprender los valores de los parámetros de los operadores adaptativos deben realizar la búsqueda en un espacio continuo, que puede ser grande (véase expresión 8). Todo ello trae como consecuencia que los algoritmos empleados para aprenderlos tengan mayor dificultad y por ello necesiten más tiempo y consigan peores soluciones.

Con el fin de reducir el espacio de búsqueda, en este trabajo se propone utilizar una t-norma adaptativa que permita seleccionar entre distintas t-normas. Este conjunto de t-normas estará formado por las del Mínimo, Hamacher y Producto Algebraico, ya que son las más ampliamente empleadas como operadores de conjunción debido a su simplicidad computacional y a su buen comportamiento práctico [5]. Además, a estas tres se han añadido cuatro t-normas creadas a partir de la t-norma de Dubois con α=0.2, 0.4, 0.6 y 0.8 (véase expresión 9), de modo que se encontrarían entre el Mínimo y el Producto Algebraico. De esta forma, el algoritmo de aprendizaje debe sólo aprender una de estas 7 posibles t-normas para cada regla, lo que supone menos flexibilidad que una t-norma continua cuyos valores se encuentren en el mismo intervalo como es el caso de la de Dubois, pero un menor esfuerzo computacional cuando el número de reglas sea muy alto.

En la siguiente subsección se describirá el modelo evolutivo propuesto que lleva a cabo la selección de estas 7 t-normas posibles para cada regla.

3.1. MODELO EVOLUTIVO PROPUESTO PARA CAAD

En el presente trabajo utilizaremos un modelo evolutivo basado en un Algoritmo Genético (AG) de tipo CHC [6] para seleccionar la mejor t-norma como operador de conjunción para cada regla de entre las propuestas anteriormente. A continuación se describirán los componentes de este modelo: codificación y población inicial, evaluación, operador de cruce y rearranque.

Codificación y Población inicial

Cada cromosoma se codifica como un vector de enteros con *m* parámetros α_i, uno para cada regla de la BR. Cada uno toma valores en el intervalo [1, 7] que determinan el operador de conjunción, de entre los siete propuestos, que será utilizado por cada regla.

$$C = (\alpha_1, \ldots, \alpha_m) \mid \alpha_i \in \{1, 7\}$$

La población inicial se obtiene generando un individuo con todos los genes a '1' y por tanto coincidiendo con aplicar como operador de conjunción la t-norma del mínimo a todas las reglas, y los demás individuos son generados aleatoriamente dentro del conjunto {1, 7}.

Evaluación

Para evaluar un cromosoma utilizamos el clásico Error Cuadrático Medio (ECM), (10).

$$ECM\ (S)_B = \frac{\frac{1}{2}\sum_{k=1}^{P} (y_k - S(x_k))^2}{P} \qquad (10)$$

donde S es el modelo difuso cuyo Sistema de Inferencia utiliza la t-norma descrita por el valor del parámetro, la t-norma del mínimo como operador de implicación, y el método de defuzzificación *centro de gravedad ponderado por el grado de emparejamiento*, utilizando una base de datos concreta decodificada. Esta medida utiliza un conjunto de datos para evaluación constituido por P pares de datos numéricos Z_k=(x_k, y_k), k=1,...,P, siendo x_k los valores de las variables de entrada y y_k los correspondientes a las variables de salida.

Operador de cruce

Nosotros utilizamos el operador de cruce HUX [7]. Este operador intercambia exactamente la mitad de los genes que son distintos en los padres, asegurando la máxima distancia entre los hijos y sus padres (exploración). Para medirla se utiliza la distancia de Hamming y el cruce se produce cuando su valor es mayor que un umbral *L* dividido por dos. Este umbral, al empezar el algoritmo, se inicializa al número de genes dividido por cuatro.

Rearranque.

Para salir de óptimos locales, CHC utiliza un mecanismo de rearranque. En este caso, el mejor cromosoma se mantiene en la población y el resto son generados aleatoriamente dentro del conjunto {1, 7}. Este mecanismo será aplicado cuando el umbral *L* alcance un valor menor que 0.

4 ESTUDIO EXPERIMENTAL Y ANÁLISIS DE RESULTADOS

Para analizar el funcionamiento de la propuesta, hemos seleccionado 5 problemas de regresión con distintos tamaños. La Tabla 2 resume las principales características de estos 5 problemas seleccionados de la página web del proyecto KEEL [2] http://sci2s.ugr.es/keel/datasets.php desde donde pueden ser descargados.

Para cada problema se muestra el número de variables y el número de ejemplos. Los problemas están ordenados de menor a mayor dimensionalidad, considerándose que el problema ELE es de baja dimensionalidad, WAN y MOR de dimensionalidad media, y los dos últimos, CA y POLE, de alta dimensionalidad para problemas de regresión.

Tabla 2: Problemas considerados en el análisis.

Problema	Nombre	Variables	Ejemplos
Electrical Maintenance	ELE	4	1056
Ankara Weather	WAN	9	1609
Mortgage	MOR	15	1409
Computer-Activity	CA	21	8192
Pole	POLE	26	14998

Esta sección está organizada como sigue: En primer lugar, en la Subsección 4.1 se describe la configuración del estudio experimental. En la Subsección se 4.2 analiza el comportamiento del operador CAAD, para lo cual se ha comparado con otros operadores de conjunción adaptativos de la literatura. Finalmente, la Subsección 4.3 estudia gráficamente las frecuencias de los operadores aprendidos para los distintos problemas.

4.1. CONFIGURACION DEL ESTUDIO EXPERIMENTAL

En este estudio, comparamos la t-norma adaptativa propuesta con las t-normas adaptativas de Dombi y Dubois [10, 14], que hemos referenciado como DOM y DUB y cuyos modelos evolutivos de aprendizaje fueron evaluados y comparados junto con otras t-normas adaptativas y clásicas en [1]. Estos modelos utilizan también un modelo CHC pero en este caso con una codificación real.

En los distintos experimentos, se trabaja con un modelo de validación cruzada con 5 particiones de datos, esto es, 5 particiones aleatorias al 20%, y la combinación de cuatro particiones (80%) como entrenamiento y una partición como prueba. Así se tienen 5 particiones al 80% y 20% en entrenamiento y prueba. Para cada partición cada método ha sido ejecutado 6 veces y en las tablas se muestran los resultados medios de las 30 ejecuciones de cada algoritmo. Además hemos aplicado un t-test con un nivel de confianza del 95% ($\alpha = 0.05$) para asegurar que existen diferencias significativas entre los métodos.

Las particiones lingüísticas iniciales consideradas tendrán cinco términos lingüísticos con forma triangular, excepto para POLE, en el que hemos considerado cuatro particiones. La razón de esta elección viene determinada por el elevado número de reglas (y por tanto un elevado coste computacional) que se obtendrían utilizando cinco particiones en este problema. En todos los casos, se ha empleado el algoritmo de Wang y Mendel(WM) [15] para obtener la BR inicial. La elección de dicho algoritmo de aprendizaje es debido a su sencillez, y tal y como se ha comentado en la Sección 3, porque la complejidad del modelo viene determinada por el número de reglas. Por lo tanto, la elección de otro algoritmo influirá sólo en este sentido. En la Tabla 3 se muestra los resultados iniciales obtenidos, donde #R representa el número medio de reglas, y ECM_{ENT} y ECM_{PRU} el valor medio de error en entrenamiento y en prueba respectivamente. En este caso, se ha optado por ordenar los problemas por la complejidad de nuestro modelo.

Tabla 3: Precisión inicial calculada con la BR de WM.

Problema	#R	ECM_{ENT}	ECM_{PRU}
ELE	65	56135	56359
MOR	199	0,128	0,134
WAN	457	4,878	6,305
CA	1539	8,449	12,44
POLE	2047	281,8	290,8

Los valores de los parámetros utilizados en todos los experimentos son:

- CAAD: Codificación = Entera, Población = 50, evaluaciones= 10000, cruce = HUX
- DOM y DUB: Codificación = Real, Población = 50, evaluaciones =10000, cruce =BLX-α con α=0,5.

4.2. ANÁLISIS DE RESULTADOS

En esta Sección se muestran y analizan los resultados obtenidos por CAAD en comparación con DOM y DUB. La Tabla 4 muestra para cada problema, la precisión (medida ECM) alcanzada para cada uno de los operadores tanto en el conjunto de entrenamiento como prueba. En negrita se destacan aquellos valores que mejor comportamiento presentan.

Tabla 4: Precisión obtenida por los operadores adaptativos CAAD, DOM y DUB.

Problema	CAAD		DOM		DUB	
	ECM_{ENT}	ECM_{PRU}	ECM_{ENT}	ECM_{PRU}	ECM_{ENT}	ECM_{PRU}
ELE	33780	35723	**32054**	**34866**	34366	36933
MOR	**0,050**	**0,058**	0,060	0,066	**0,050**	0,059
WAN	2,119	3,681	3,127	4,550	**2,069**	**3,659**
CA	**3,873**	**8,105**	4,105	9,234	3,942	8,185
POLE	**119,81**	**130,63**	128,45	139,65	124,32	136,87

La Tabla 5 muestra los resultados de aplicar un test t-student (con 95% de confianza) al mejor resultado medio de la correspondiente columna comparado 1 a 1 con el resto de resultados medios para cada problema. La interpretación de ellas es:

+ El primero tiene un comportamiento significativamente mejor que el segundo.

- El primero tiene un comportamiento significativamente peor que el segundo.

= Indica un comportamiento estadísticamente sin diferencias significativas.

Los resultados muestran que CAAD, cuando el número de reglas es más alto (problemas CA y POLE), consigue mayor precisión que las otras t-normas adaptativas. Por el contrario, en los otros problemas con menos reglas (WAN, MOR y ELE) que son problemas de menor dimensionalidad, la menor flexibilidad de CAAD le impide alcanzar los mismos niveles de precisión.

Tabla 5: T-test de comparación entre CAAD y los operadores adaptativos de DOM y DUB.

Problema	Comparación	t-test en ECM_{ENT}	t-test en ECM_{PRU}
ELE	CAAD-DOM	-	-
	CAAD-DUB	=	=
MOR	CAAD-DOM	+	+
	CAAD-DUB	=	=
WAN	CAAD-DOM	+	+
	CAAD-DUB	=	=
CA	CAAD-DOM	+	+
	CAAD-DUB	+	+
POLE	CAAD-DOM	+	+
	CAAD-DUB	+	+

En la tabla 6 se muestran los tiempos medios de ejecución para cada problema y operador de conjunción adaptativo, obtenidos en un ordenador con CPU Intel Core™ 2 Duo a 2.66 GHz, con 4 Gb de memoria y con sistema operativo linux de 64 bits. Los operadores de DOM y DUB emplean en todos los casos tiempos significativamente mayores que CAAD, siendo de especial importancia con los dos mayores conjuntos de datos donde los tiempos empleados por CAAD son de magnitud de horas frente a magnitudes de días en DOM y DUB, donde la sencillez de la expresión del segundo de estos dos frente al primero también se deja apreciar. El número de reglas de cada problema, como se esperaba, es el factor determinante en los tiempos invertidos por los distintos operadores.

Tabla 6: Tiempos de ejecución (hh:mm:ss).

Problema	CAAD	DOM	DUB
ELE	00:00:48	00:03:06	00:02:35
MOR	00:06:33	00:34:24	00:16:25
WAN	00:16:31	00:71:06	00:38:03
CA	15:46:44	90:16:02	25:04:04
POLE	16:12:04	120:33:00	59:43:00

En definitiva, para los dos conjuntos de datos mayores, el operador CAAD muestra mayor precisión y tiempos de ejecución sensiblemente menores.

4.3. ESTUDIO GRÁFICO DE LOS PARÁMETROS OBTENIDOS POR CAAD

En esta subsección se realiza un estudio gráfico de los parámetros utilizados por CAAD en cada problema.

En la Figura 1, se muestra una representación gráfica de frecuencias en porcentaje del tipo de t-norma aprendida para cada conjunto de datos, es decir, el porcentaje de reglas que utilizan cada operador de conjunción.

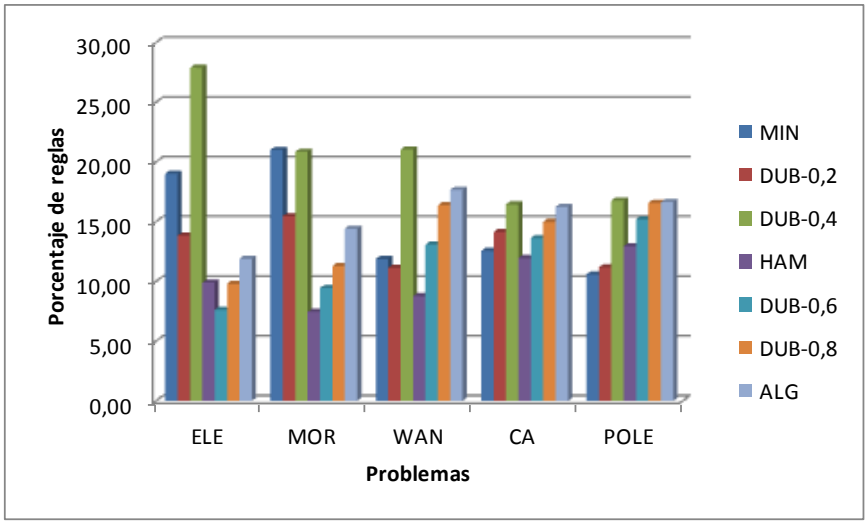


Figura 1: Porcentaje de reglas para los distintos operadores obtenidos por CAAD para cada problema.

Como se puede apreciar, en general, el operador de DUB-0.4 es el que más reglas utilizan en casi la totalidad de los problemas. En el extremo contrario, Hamacher es poco usado, seguido de DUB-0.6 y DUB-0.2, si bien no hay ningún conectivo no utilizado.

Por otro lado, las frecuencias de uso de las distintas t-normas dependen de la complejidad del problema, pues se aprecia que en los conjuntos de datos de mayor complejidad, los porcentajes de cada conectivo son más uniformes que en los problemas más sencillos donde se puede encontrar mayor diferencia entre las frecuencias de los conectivos más y menos utilizados por las reglas del sistema. Es posible que este efecto creemos que abre una puerta a estudiar más detenidamente sus orígenes.

5 CONCLUSIONES

En este trabajo se ha presentado un operador de conjunción adaptativo evolutivo para problemas de alta dimensionalidad, CAAD, que permite obtener SBRDs precisos con un coste computacional notablemente inferior en comparación con aquellos derivados de otros operadores de conjunción adaptativos clásicos de la literatura.

Este estudio preliminar debe ser completado con mayor número de problemas y bases de reglas generadas por diferentes algoritmos en un futuro, así como el estudio de las t-normas más o menos utilizadas para verficar si depende o no del problema y no sólo de su dimensionalidad.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio Español de Educación y Ciencia bajo el proyecto TIN2008-06681-C06-06.

Referencias

[1] J. Alcala-Fdez, F. Herrera, F. Márquez, A. Peregrín: Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *International Journal of Intelligent Systems* 22:9, pp. 1035-1064, 2007.

[2] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, F. Herrera: KEEL: A Software tool to asses evolutionary algorithms to data mining problems. *Soft Computing* 13:3, pp. 307-318, 2009.

[3] W. E. Combs, J. E. Andrews: Combinatorial rule explosion eliminated by a fuzzy rule configuration. *IEEE Transactions on Fuzzy Systems 6:1*, pp. 1-11, 1998.

[4] O. Cordón , F. Herrera , F. A. Márquez, A. Peregrín: A study on the evolutionary adaptive defuzzification methods in fuzzy modelling. *International Journal of Hybrid Intelligent Systems* 1:1, pp. 36-48, 2004.

[5] O. Cordón, F. Herrera. A. Peregrín: Searching for Basic Properties Obtaining Robust Implication Operators in Fuzzy Control. *Fuzzy Sets and* Systems 111, pp. 237-251, 2000.

[6] L.J. Eshelman: The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *Foundations of genetic Algorithms* 1, pp. 265-283, 1991.

[7] L.J. Eshelman, J.D. Schaffer: Real-coded genetic algorithms and interval-schemata. *Foundations of Genetic Algorithms* 2, pp.187-202, 1993.

[8] M. Gupta, J. Qi: Theory of T-norms and Fuzzy Inference Methods. *Fuzzy Sets and Systems* 40, pp. 431-450, 1991.

[9] F. Herrera: Genetic fuzzy systems: Taxonomy, current research trends and prospects. *Evolutionary Intelligence* 1, pp. 27-46, 2008.

[10] F. Herrera, M. Lozano y J.L. Verdegay: Dynamic and Heuristic Fuzzy Connectives-Based Crossover Operators for Controlling the Diversity and Convergence of Real-Coded Genetic Algoritms. *International Journal of Intelligent Systems* 11, pp. 1013-1040, 1996.

[11] F. Herrera, F. Márquez, A. Peregrín: Genetic Adaptation of Rule Connectives and Conjunction Operators in Fuzzy Rule Based System: An Experimental Comparative. En: *Proceedings of 3th International Conference of the European Society for Fuzzy Logic and Technology Study*, Zitau (Alemania), pp. 100-104, 2003.

[12] Y.Jin: Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement. *IEEE Transactions on Fuzzy Systems 8:2*, pp. 212-221, 2000.

[13] F.A. Márquez, A. Peregrín y F. Herrera: Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Linguistic Fuzzy Systems. *IEEE Transactions on Fuzzy Systems* 15:6, pp. 1162-1178, 2007.

[14] M. Mizumoto: Pictorial Representations of Fuzzy Connectives, Part I: Cases of T-norms, T-conorms and Averaging Operators. *Fuzzy Sets and Systems*, 31, Pág. 217-242, 1989.

[15] L.X. Wang, J.M. Mendel: Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man, and Cybernetics* 22:6, pp. 1414-1427, 1992.

UN SISTEMA DE CLASIFICACIÓN BASADO EN REGLAS DIFUSAS JERÁRQUICO CON PROGRAMACIÓN GENÉTICA PARA PROBLEMAS DE CLASIFICACIÓN ALTAMENTE NO BALANCEADOS

Victoria López[1], **Alberto Fernández**[2], **María José del Jesus**[2], **Francisco Herrera**[1]

[1]*Departamento de Ciencias de la Computación e Inteligencia Artificial, CITIC-UGR, Granada, España*, {vlopez,herrera}@decsai.ugr.es

[2]*Departamento de Informática, Universidad de Jaén, Jaén, España*, {alberto.fernandez,mjjesus}@ujaen.es

Resumen

En muchas aplicaciones del mundo real aparece el problema de clasificación con datos no balanceados. Este problema se da cuando la distribución de datos entre las clases es bastante diferente. Los algoritmos de clasificación no suelen funcionar adecuadamente al abordar esta situación ya que tienden a clasificar incorrectamente las instancias de la clase minoritaria que a menudo es el foco de interés de la aplicación.

Los sistemas de clasificación basados en reglas difusas han conseguido un buen rendimiento en el marco de la clasificación con datos no balanceados. En esta contribución, nos centramos en el uso de un sistema jerárquico con diferentes niveles de granularidad combinado con un procedimiento genético que une la selección de reglas y el ajuste lateral con el modelo lingüísitico de 2-tuplas. Los resultados empíricos en conjuntos con alto desbalanceo reflejan un mejor comportamiento de la propuesta frente a otros sistemas de clasificación basados en reglas difusas y C4.5.

Palabras Clave: Sistemas de clasificación basados en reglas difusas, sistemas difusos evolutivos, particiones difusas jerárquicas, selección, ajuste, clasificación no balanceada

1 Introducción

El problema de clasificación con datos no balanceados, también conocido como de desequilibrio entre clases, aparece cuando el número de ejemplos de una clase es superado notablemente por el número de ejemplos de la otra clase [13, 18]. Este problema se da en aplicaciones reales donde la clase minoritaria es el foco de interés del problema. Los algoritmos de clasificación tradicionales no suelen funcionar correctamente con el desbalanceo ya que se suelen ignorar los ejemplos de la clase minoritaria en favor de los de la clase mayoritaria.

Los Sistemas de Clasificación Basados en Reglas Difusas (SCBRDs) [14] son herramientas útiles dentro del ámbito del aprendizaje automático ya que presentan un buen equilibrio entre interpretabilidad y precisión. El buen comportamiento de los SCBRDs en conjuntos de datos no balanceados ha sido analizado en [10, 11]. La definición automática de los componentes de un sistema difuso utilizando algoritmos evolutivos ha dado lugar a la aparición de los sistemas difusos evolutivos. Dentro de ellos, la programación genética evoluciona un conjunto de soluciones en forma de árbol con longitud variable, permitiendo obtener una base de reglas adecuada para un SCBRD [5]. Sin embargo, una desventaja de los SCBRDs es la inflexibilidad del concepto de variable lingüística porque impone fuertes restricciones en la estructura de la regla difusa que pueden suponer una pérdida de precisión en problemas de alta complejidad.

Para mejorar el rendimiento de los SCBRDs se han propuesto diversos mecanismos que incrementan la cooperación entre las reglas de la Base de Conocimiento (BC). En este trabajo se presenta un SCBRD que promueve la sinergia entre varias de estas estrategias: el uso de una Base de Conocimiento Jerárquica (BCJ) [8], la selección evolutiva de reglas y el ajuste lateral. En concreto, se propone el uso de un SCBRD Jerárquico dividido en tres etapas para abordar el problema de clasificación no balanceada. En un primer paso, la técnica "Synthetic Minority Over-sampling Technique" (SMOTE) [6] equilibra la distribución del conjunto de entrenamiento. A continuación, se genera una BCJ con el algoritmo "Genetic Programming-based learning of COmpact and ACcurate fuzzy rule-based classification systems for High-dimensional problems" (GP-COACH) [5] modificado para permitir la expansión de una BC clásica en una BC jerárquica. Finalmente, se añade un postprocesamiento evolutivo que combina selección de reglas con un ajuste lateral de las etiquetas basado en el modelo lingüístico de 2-tuplas para incrementar el rendimiento final del modelo [1]. Este modelo se ha denominado "GP-COACH Hierarchical" (GP-COACH-H).

Para mostrar la bondad de la propuesta hemos realizado una comparación frente a otros SCBRDs y C4.5 sobre 44 conjuntos de datos de alto no balanceo, escogidos del repositorio KEEL [2]. Este estudio ha sido contrastado a través de test estadísticos no paramétricos [12, 17].

La organización del trabajo es como sigue. En la Sección 2 se introduce el problema de clasificación con datos no balanceados. La Sección 3 describe el método propuesto detallando sus etapas. A continuación, en la Sección 4, mostramos el estudio experimental llevado a cabo.

2 Conjuntos de datos no balanceados en clasificación

Comenzamos esta sección introduciendo el problema de clasificación con datos no balanceados. Seguidamente, describimos qué técnicas se utilizan para abordar este problema destacando el algoritmo SMOTE [6]. Terminamos explicando las métricas de evaluación para estos problemas.

2.1 El problema de clasificación no balanceada

El aprendizaje a partir de conjuntos de datos no balanceados es un problema que ha cobrado importancia en los últimos tiempos en minería de datos [13, 18]. Este problema es muy representativo ya que aparece en muchas aplicaciones del mundo real como en usos médicos, detección de anomalías, telecomunicaciones o finanzas. Nos referimos a conjuntos de datos no balanceados cuando la distribución entre clases no es uniforme. En esta situación el número de ejemplos que representan a una clase (normalmente, la que acapara el mayor interés) es mucho menor que el del resto.

Los algoritmos clásicos para clasificación tienden a clasificar correctamente los ejemplos de la clase negativa (clase predominante) ya que dichos algoritmos suelen buscar reglas generales que abarquen el mayor número de ejemplos, de acuerdo a las métricas utilizadas tradicionalmente. De esta manera, las instancias que pertenecen a la clase positiva (clase minoritaria) se clasifican incorrectamente con más frecuencia que las que pertenecen a la clase negativa.

En este trabajo utilizamos el "imbalance ratio" (IR) [15] para distinguir distintos grados de no balanceo en los datos. El IR se define como la razón entre el número de ejemplos de la clase negativa y el número de ejemplos de la clase positiva. Consideraremos que un conjunto de datos presenta un alto grado de no balanceo cuando su IR es mayor que 9 (menos de un 10% de instancias de la clase positiva).

2.2 Aproximaciones para la clasificación de conjuntos de datos no balanceados: el algoritmo SMOTE

Se han utilizado un gran número de enfoques para abordar el problema de clasificación con datos no balanceados. Estos enfoques se pueden catalogar en dos grupos: enfoques a nivel algorítmico que crean o modifican procedimientos teniendo en cuenta el no balanceo de los datos o enfoques a nivel de datos que preprocesan el conjunto de entrenamiento para disminuir el efecto del desbalanceo. En este trabajo se utilizan métodos de sobremuestreo que son un referente en el área: el algoritmo SMOTE [6] y una variante, "SMOTE + Edited Nearest Neighbor" (SMOTE+ENN) [4].

El algoritmo SMOTE aumenta el número de ejemplos de la clase positiva introduciendo instancias sintéticas en los segmentos que unen cualquiera de las instancias positivas con alguno de sus k vecinos más cercanos, consiguiendo que la región de instancias positivas sea más general. En ocasiones, los ejemplos generados por SMOTE ocupan zonas del espacio de la clase negativa afectando negativamente al resultado final de la clasificación. Para evitarlo, se utiliza el enfoque híbrido SMOTE+ENN en el que se aplica el algoritmo ENN tras el uso de SMOTE.

2.3 Métricas de evaluación

Muchas de las propuestas de aprendizaje automático de clasificadores utilizan alguna medida de precisión como el porcentaje de ejemplos bien clasificados. Sin embargo, el uso de este tipo de medidas puede conducir a conclusiones erróneas en conjuntos de datos no balanceados ya que no tienen en cuenta la proporción de ejemplos de cada clase.

La medida utilizada en este trabajo es la media geométrica (MG) [3], que se define como $MG = \sqrt{VP_{tasa} \cdot VN_{tasa}}$, donde VP_{tasa} es el porcentaje de ejemplos de la clase positiva bien clasificados y VN_{tasa} es el porcentaje de ejemplos de la clase negativa bien clasificados. Esta medida trata de maximizar el acierto de cada una de las dos clases equiparando estos dos objetivos en una única medida.

3 GP-COACH-H: Un clasificador basado en reglas difusas jerárquico con programación genética

En esta sección se describe nuestra propuesta de un clasificador basado en reglas difusas jerárquico para problemas de clasificación con alto no balanceo, GP-COACH-H. Este método está compuesto por tres etapas:

1. Preprocesamiento del conjunto de entrenamiento con el algoritmo SMOTE.
2. Generación de la Base de Reglas Jerárquica (BRJ) con el algoritmo GP-COACH modificado.
3. Post-procesamiento evolutivo: selección de reglas y ajuste lateral 2-tuplas.

El funcionamiento de este algoritmo se encuentra resumido en la Figura 1. A continuación se describen con mayor detalle las operaciones de este algoritmo.

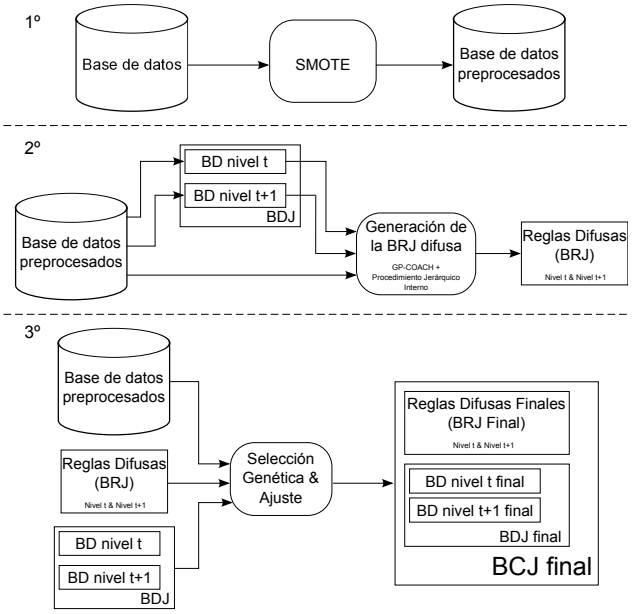


Figura 1: Funcionamiento de GP-COACH-H

3.1 Sistemas de clasificación basados en reglas difusas

Un SCBRD tiene dos componentes principales: el Sistema de Inferencia y la BC. En un SCBRD, la BC está compuesta de la Base de Reglas (BR), constituida por el conjunto de reglas difusas y la Base de Datos (BD), que contiene las funciones de pertenencia de las particiones difusas asociadas a las variables de entrada. Si no existe información experta sobre el problema, es necesario utilizar algún proceso de aprendizaje automático de la BC a partir de ejemplos.

En este trabajo se usan reglas difusas tipo "Disjunctive Normal Form" (DNF) (Ecuación 1). El peso de una regla se calcula con el factor de certeza (Ecuación 2). GP-COACH-H utiliza la suma normalizada como método de razonamiento difuso para clasificar nuevos ejemplos (Ecuación 3).

$$\begin{array}{c}\text{Regla } R_j: \text{ Si } x_1 \text{ es } \hat{A}_1^j \text{ y } \ldots \text{ y } x_n \text{ es } \hat{A}_n^j \\ \text{entonces Clase = } C_j \text{ con } RW_j\end{array} \quad (1)$$

$$RW_j = CF_j = \frac{\sum_{x_p \in Clase C_j} \mu_{\hat{A_j}}(x_p)}{\sum_{p=1}^{N} \mu_{\hat{A_j}}(x_p)} \quad (2)$$

$$\begin{array}{c} Sum_{Clase\ h}(x_p) = \frac{\sum_{R_j \in BR, C_j = h} \mu_{\hat{A_j}}(x_p) \cdot CF_j}{\max_{c=1,\cdots,M} \sum_{R_j \in BR, C_j = c} \mu_{\hat{A_j}}(x_p) \cdot CF_j} \\ Clase(x_p) = \arg\max(Sum_{Class\ h}(x_p)) \end{array} \quad (3)$$

En estas ecuaciones, R_j es la etiqueta de la j-ésima regla, $x = (x_1, \ldots, x_n)$ es un vector de muestras n-dimensional (variables de entrada), $\hat{A}_i^j$ es el conjunto de etiquetas lingüísticas $\{L_i^1 \text{ o } \ldots \text{ o } L_i^{l_i}\}$, C_j es la etiqueta de una clase, RW_j es el peso de la regla y $\mu_{A_j}(x_p)$ es el grado de emparejamiento del ejemplo x_p con el antecedente de la regla difusa R_j. Representamos las funciones de pertenencia por medio de funciones triangulares.

3.2 El algoritmo GP-COACH

El algoritmo GP-COACH [5] es un algoritmo de programación genética para el aprendizaje de bases de reglas difusas. Utilizamos este método como base de nuestra propuesta jerárquica modificando su comportamiento para incluir varios niveles de granularidad dentro de su comportamiento.

Este algoritmo trabaja con una población de reglas tipo DNF (Ecuación 1) que permite la ausencia de algunas variables de entrada. Cada individuo de la población codifica una regla y la BR está formada por todos la población al completo. Durante la evolución de la población se utilizan dos funciones de evaluación: una función local que mide la idoneidad de cada regla y una función global que evalúa la población al completo. Además, se incorpora un mecanismo para mantener la diversidad en la población y eliminar reglas irrelevantes, el procedimiento de 'token competition". GP-COACH incorpora un procedimiento de inferencia en dos fases utilizando primero reglas de tipo 1 y si no es posible su aplicación se aplican reglas de tipo 2. Para una descripción más detallada del algoritmo GP-COACH se recomienda consultar [5].

3.3 Construcción del Sistema Basado en Reglas Difusas Jerárquico

Los Sistemas Basados en Reglas Difusas Jerárquicos tratan de mejorar el rendimiento de los sistemas basados en reglas difusas en zonas del espacio de búsqueda especialmente difíciles. Para ello, en lugar de utilizar la definición clásica de la BC, utilizamos una expansión conocida como BCJ [8], que se compone de una serie de niveles. Cada uno de estos niveles contiene una BD y una BR asociadas. La BD de cada nivel contiene tantas particiones lingüísticas como indique la granularidad del nivel y la BR contiene las reglas que toman valores en las particiones lingüísticas anteriores. Este conjunto de niveles está organizado en una jerarquía de manera que la granularidad de los diferentes niveles aumenta conforme lo hace su orden.

GP-COACH-H utiliza una BCJ con dos niveles de granularidad. Las BD de ambos niveles se generan a partir del conjunto de entrenamiento con etiquetas triangulares simétricas y uniformemente distribuidas de forma similar a la utilizada en [8, 10]. El uso de diferentes niveles de jerarquía obliga a introducir modificaciones en GP-COACH.

GP-COACH-H mantiene una población mixta que contiene reglas de los diferentes niveles de granularidad, adapta los operadores genéticos a esta situación y modifica la función de evaluación de la población en consecuencia. El procedimiento para jerarquizar las reglas actuales comienza tras la creación de reglas de tipo 2. Para ello, se cataloga cada una de las reglas como *buena* o *mala*:

if $(Ac(X, R_j) \leq (1-\alpha) \cdot Ac_i(X_i, RB))$ **then**
R_j = *regla buena*

else
R_j = *regla mala*
end if

$$Ac(X,R_j) = \frac{|X^+(R_j)|}{|X(R_j)|} \quad (4)$$

$$Ac_i(X_i,RB) = \frac{|x_p \in X_i / MRD(x_p,RB) = Clase(x_p)|}{|X_i|} \quad (5)$$

donde $|\,|$ es el número de ejemplos, $X^+(R_j)$ y $X(R_j)$ son los conjuntos de ejemplos en los que la regla colabora en su clasificación (correcta clasificación para $X^+(R_j)$), X_i es conjunto de ejemplos de entrenamiento pertenecientes a la i-ésima clase, $MRD(x_p,RB)$ es la clase predicha para un ejemplo y $Clase(x_p)$ es la clase del ejemplo x_p.

Las reglas que se han considerado *buenas* se mantienen en la población mientras que las reglas *malas* se eliminan de la población actual. Se añaden a la población nuevas reglas de alta granularidad generadas con el algoritmo Chi [7] sobre los ejemplos que eran clasificados por las reglas *malas*.

Gracias al uso de estas operaciones, conseguimos que GP-COACH-H obtenga un buen conjunto de reglas de diferentes granularidades. Para una información más extensa de la estructura de la BCJ y la generación de reglas de alta granularidad a partir de un conjunto de reglas ya generado se sugiere la lectura de [8, 10].

3.4 Selección Evolutiva de Reglas Jerárquicas y Ajuste Lateral

En este último paso, se presenta el uso de algoritmos genéticos para seleccionar un subconjunto de reglas difusas compacto y cooperativo y ajustar las funciones de pertenencia de las etiquetas en la BD realizando desplazamientos laterales en las mismas según el modelo 2-tuplas [1, 10, 11]. La representación de las etiquetas según el modelo 2-tuplas permite realizar la búsqueda del desplazamiento lateral teniendo en cuenta un único parámetro de transición simbólica: un número del intervalo $[-0.5, 0.5)$ que expresa el dominio en el que se puede mover una etiqueta lingüística.

Para llevar a cabo este proceso de selección de reglas y ajuste lateral utilizamos el modelo genético del método CHC [9]. Los individuos de la población codifican los dos objetivos a alcanzar: una parte del cromosoma indica si una regla ha sido seleccionada para la solución final o no y la otra parte del cromosoma contiene valores reales que indican el desplazamiento para cada una de las etiquetas de las variables en las dos BD de diferentes niveles de jerarquía.

Para evaluar un cromosoma se calcula el acierto sobre el conjunto de entrenamiento usando la BC con las modificaciones representadas en su codificación. El cruce entre dos individuos se realiza de manera distinta para cada parte del cromosoma: se usa el operador HUX para la parte con codificación binaria y el operador PCBLX para la parte con codificación real del cromosoma. Se recomienda consultar trabajos previos que se usan un esquema similar [1, 10, 11].

Tabla 1: Conjuntos de datos altamente no balanceados

Nombre	#Ej.	#Var.	Clase (+,-)	%Clase(+; -)	IR
ecoli034vs5	200	7	(p,imL,imU; om)	(10.00, 90.00)	9.00
yeast2vs4	514	8	(cyt; me2)	(9.92, 90.08)	9.08
ecoli067vs35	222	7	(cp,omL,pp; imL,om)	(9.91, 90.09)	9.09
ecoli0234vs5	202	7	(cp,imS,imL,imU; om)	(9.90, 90.10)	9.10
glass015vs2	172	9	(build-win-non_float-proc,tableware, build-win-float-proc; ve-win-float-proc)	(9.88, 90.12)	9.12
yeast0359vs78	506	8	(mit,me1,me3,erl; vac,pox)	(9.88, 90.12)	9.12
yeast02579vs368	1004	8	(mit,cyt,me3,vac,erl; me1,exc,pox)	(9.86, 90.14)	9.14
yeast0256vs3789	1004	8	(mit,cyt,me3,exc; me1,vac,pox,erl)	(9.86, 90.14)	9.14
ecoli046vs5	203	6	(cp,imU,omL; om)	(9.85, 90.15)	9.15
ecoli01vs235	244	7	(cp,im; imS,imL,om)	(9.83, 90.17)	9.17
ecoli0267vs35	224	7	(cp,imS,omL,pp; imL,om)	(9.82, 90.18)	9.18
glass04vs5	92	9	(build-win-float-proc,containers; tableware)	(9.78, 90.22)	9.22
ecoli0346vs5	205	7	(cp,imL,imU,omL; om)	(9.76, 90.24)	9.25
ecoli0347vs56	257	7	(cp,imL,imU,pp; om,omL)	(9.73, 90.27)	9.28
yeast05679vs4	528	8	(me2; mit,me3,exc,vac,erl)	(9.66, 90.34)	9.35
ecoli067vs5	220	6	(cp,omL,pp; om)	(9.09, 90.91)	10.00
vowel0	988	13	(hid; remainder)	(9.01, 90.99)	10.10
glass016vs2	192	9	(ve-win-float-proc; build-win-float-proc, build-win-non_float-proc,headlamps)	(8.89, 91.11)	10.29
glass2	214	9	(ve-win-float-proc; remainder)	(8.78, 91.22)	10.39
ecoli0147vs2356	336	7	(cp,im,imU,pp; imS,imL,om,omL)	(8.63, 91.37)	10.59
led7digit02456789vs1	443	7	(0,2,4,5,6,7,8,9; 1)	(8.35, 91.65)	10.97
glass06vs5	108	9	(build-win-float-proc,headlamps; tableware)	(8.33, 91.67)	11.00
ecoli01vs5	240	6	(cp,im; om)	(8.33, 91.67)	11.00
glass0146vs2	205	9	(build-win-float-proc,containers,headlamps, build-win-non_float-proc;ve-win-float-proc)	(8.29, 91.71)	11.06
ecoli0147vs56	332	6	(cp,im,imU,pp; om,omL)	(7.53, 92.47)	12.28
cleveland0vs4	177	13	(0; 4)	(7.34, 92.66)	12.62
ecoli0146vs5	280	6	(cp,im,imU,omL; om)	(7.14, 92.86)	13.00
ecoli4	336	7	(om; remainder)	(6.74, 93.26)	13.84
yeast1vs7	459	8	(nuc; vac)	(6.72, 93.28)	13.87
shuttle0vs4	1829	9	(Rad Flow; Bypass)	(6.72, 93.28)	13.87
glass4	214	9	(containers; remainder)	(6.07, 93.93)	15.47
page-blocks13vs2	472	10	(graphic; horiz.line,picture)	(5.93, 94.07)	15.85
abalone9vs18	731	8	(18; 9)	(5.65, 94.25)	16.68
glass016vs5	184	9	(tableware; build-win-float-proc, build-win-non_float-proc,headlamps)	(4.89, 95.11)	19.44
shuttle2vs4	129	9	(Fpv Open; Bypass)	(4.65, 95.35)	20.5
yeast1458vs7	693	8	(vac; nuc,me2,me3,pox)	(4.33, 95.67)	22.10
glass5	214	9	(tableware; remainder)	(4.20, 95.80)	22.81
yeast2vs8	482	8	(pox; cyt)	(4.15, 95.85)	23.10
yeast4	1484	8	(me2; remainder)	(3.43, 96.57)	28.41
yeast1289vs7	947	8	(vac; nuc,cyt,pox,erl)	(3.17, 96.83)	30.56
yeast5	1484	8	(me1; remainder)	(2.96, 97.04)	32.78
ecoli0137vs26	281	7	(pp,imL; cp,im,imU,imS)	(2.49, 97.51)	39.15
yeast6	1484	8	(exc; remainder)	(2.49, 97.51)	39.15
abalone19	4174	8	(19; remainder)	(0.77, 99.23)	128.87

4 Experimentos y análisis de resultados

En esta sección, nuestro objetivo consiste en analizar el comportamiento del algoritmo GP-COACH-H. Para ello, escogemos 44 conjuntos de datos con un grado alto de no balanceo ($IR > 9$) del repositorio KEEL[1] [2] mostrados en la Tabla 1, donde se indica el número de ejemplos de cada conjunto (#Ex.), número de variables (#Var.), nombre de cada clase (positiva y negativa), distribución de ejemplos y tasa de no balanceo (IR). Los conjuntos de datos están ordenados en orden ascendente de IR.

El rendimiento del algoritmo GP-COACH-H se compara con otros SCBRDs, en concreto, comparamos la propuesta frente a la versión básica del algoritmo GP-COACH [5] con diferentes niveles de granularidad (5 y 9 etiquetas), el SCBRD jerárquico HFRBCS(Chi) [10] y el árbol de decisión C4.5 [16]. Los parámetros de configuración de los algoritmos considerados (abreviados como GP-C-H, GP-C-5 y 9, Chi-H, y C4.5 en las tablas) se muestran en la Tabla 2 y corresponden a los parámetros recomendados por los autores de los algoritmos. Para GP-COACH-H, escogemos los parámetros asociados a su algoritmo base GP-COACH y al

[1] http://www.keel.es/imbalanced.php

Tabla 2: Parámetros de configuración de los algoritmos

Algoritmo	Parámetros
Parámetros de los SCBRD	
GP-C y GP-C-H	T-norma mínimo, T-conorma máximo, Peso de reglas = Factor de certeza, Método de razonamiento difuso = Suma normalizada, Número de etiquetas difusas = 5 ó 9 (en GP-C), Número de etiquetas difusas = 5 para reglas de baja granularidad y 9 para reglas de alta granularidad (en GP-C-H)
Chi-H	T-norma producto, Peso de reglas = Factor de certeza penalizado, Método de razonamiento difuso = Regla ganadora, Número de etiquetas difusas = 3 para reglas de baja granularidad y 5 para reglas de alta granularidad
Parámetros de las versiones GP-COACH	
GP-C y GP-C-H	Evaluaciones = 20000, Tamaño de la población inicial = 200, α (raw fitness) = 0.7, Probabilidad de cruce = 0.5, Probabilidad de mutación = 0.2, Probabilidad de pérdida de una condición = 0.15, Probabilidad de inserción = 0.15, Tamaño del torneo = 2, $w_1 = 0.8$, $w_2 = w_3 = 0.05$, $w_4 = 0.1$
Parámetros asociados al proceso de expansión jerárquica	
Chi-H y GP-C-H	α (expansión regla) = 0.2, Evaluaciones CHC = 10000, Tamaño de la población CHC = 61, bits por gen CHC (GP-C-H) = 30
Parámetros de C4.5	
C4.5	Podado = true, Confianza = 0.25, Número mínimo de conjuntos por hoja = 2

proceso de expansión jérarquica de HFRBCS(Chi).

Para contrarrestar el efecto del no balanceo en los datos se utiliza el algoritmo SMOTE [6] para preprocesar los conjuntos de datos a utilizar por los SCBRDs y la variante SMOTE+ENN [4] para los conjuntos de datos con C4.5. Ambos algoritmos equilibran el número de ejemplos en el conjunto de entrenamiento al 50 % utilizando únicamente el vecino más cercano según la distancia euclídea.

Los distintos experimentos se llevan a cabo bajo un esquema de validación cruzada con 5 particiones, esto es, 5 particiones con el 20 % de los datos siendo la combinación de 4 de ellas (80 %) el conjunto de entrenamiento y la restante el conjunto de test. Para cada conjunto de datos se muestran los resultados medios de las cinco particiones.

Además, se utilizan técnicas estadísticas para encontrar diferencias significativas entre los métodos estudiados [12]. Consideramos el uso de test estadísticos no paramétricos debido a que las condiciones iniciales de aplicabilidad de test paramétricos podrían no ser satisfechas por lo que el análisis estadístico perdería credibilidad [17]. Concretamente, usamos el test de Iman-Davenport para detectar si existen diferencias significativas entre un grupo de algoritmos y el test post-hoc de Holm que muestra las diferencias que hay entre los algoritmos y un método de control.

La Tabla 3 muestra la medida MG por conjunto de datos en test para cada uno de los algoritmos utilizados en la comparación. Observando los resultados medios que aparecen en dicha tabla, podemos ver que GP-COACH-H es el algoritmo que obtiene el mayor valor de MG. Para contrastar estos resultados por medio de test estadísticos comenzamos con la aplicación del test de Iman-Davenport que determina la existencia de diferencias significativas entre algoritmos. El p-valor obtenido $(0,0779)$ es suficientemente bajo como para rechazar la hipótesis de igualdad con un nivel

Tabla 3: Tabla de resultados para los SCBRDs y C4.5 en los conjuntos de datos con alto desbalanceo. Preprocesamiento con SMOTE para los SCBRDs y SMOTE+ENN para C4.5

Conjunto de datos	GP-C-5	GP-C-9	Chi-H	GP-C-H	C4.5
ecoli034vs5	0,9018	**0,9250**	0,8421	0,8660	0,8761
yeast2vs4	0,8987	0,9036	0,8932	**0,9304**	0,9029
ecoli067vs35	0,8185	**0,8509**	0,8267	0,7286	0,7206
ecoli0234vs5	0,8286	0,8472	0,8425	0,8473	**0,8861**
glass015vs2	0,3732	0,2115	0,5590	0,6301	**0,7788**
yeast0359vs78	0,5111	0,4467	**0,7330**	0,7189	0,6894
yeast02579vs368	0,9087	0,9044	0,8946	0,9107	**0,9125**
yeast0256vs3789	0,7954	0,7955	0,7927	**0,7982**	0,7707
ecoli046vs5	**0,9171**	0,8793	0,8800	0,8677	0,8776
ecoli01vs235	0,8682	**0,9138**	0,8709	0,8471	0,8277
ecoli0267vs35	0,8407	0,8365	0,8247	**0,9028**	0,8061
glass04vs5	0,9064	0,9632	0,7092	0,9429	**0,9748**
ecoli0346vs5	0,8772	0,8934	0,8729	0,8847	**0,8946**
ecoli0347vs56	0,8525	0,8571	**0,9007**	0,8767	0,8413
yeast05679vs4	**0,8136**	0,8080	0,7318	0,6988	0,7678
ecoli067vs5	0,8356	**0,8923**	0,8559	0,8671	0,8376
vowel0	0,9598	0,9400	**0,9882**	0,9465	0,9417
glass016vs2	0,5419	0,4211	0,5837	**0,6467**	0,6063
glass2	0,6223	0,3949	0,5484	0,5886	**0,7377**
ecoli0147vs2356	0,7976	0,7972	**0,8477**	0,8263	0,8119
led7digit02456789vs1	**0,9011**	0,8874	0,8276	0,9000	0,8370
glass06vs5	**0,9949**	0,9060	0,8907	0,9120	0,9628
ecoli01vs5	0,8739	**0,9190**	0,8689	0,8946	0,8081
glass0146vs2	0,6651	0,5675	0,5117	**0,7300**	0,6157
ecoli0147vs56	0,8472	0,8577	**0,8886**	0,8372	0,8250
cleveland0vs4	0,7232	0,7578	0,3961	**0,8646**	0,7307
ecoli0146vs5	0,8832	0,8862	0,8674	**0,9194**	0,8880
ecoli4	**0,9373**	0,9008	0,9302	0,9357	0,8947
yeast1vs7	0,6802	0,6093	0,7074	0,6900	**0,7222**
shuttle0vs4	0,9991	0,9954	0,9912	**1,0000**	0,9997
glass4	0,7231	**0,8811**	0,7039	0,7303	0,7639
page-blocks13vs4	0,9035	0,8706	0,9864	0,9482	**0,9909**
abalone9-18	0,6922	0,6699	0,6756	**0,7500**	0,6884
glass016vs5	**0,9644**	0,9090	0,7796	0,8550	0,7738
shuttle2vs4	0,9568	0,9960	0,9749	0,9918	**1,0000**
yeast1458vs7	0,3546	0,5353	0,6249	**0,6304**	0,3345
glass5	0,5801	0,6758	0,6873	**0,7877**	0,5851
yeast2vs8	0,7274	0,7601	0,7247	0,7381	**0,8033**
yeast4	0,7923	0,7807	**0,8264**	0,8175	0,6897
yeast1289vs7	0,5262	0,5860	0,6937	**0,6939**	0,5522
yeast5	0,9020	0,9229	0,9420	**0,9428**	0,9390
ecoli0137vs26	**0,7215**	0,7203	0,7148	0,7067	0,7062
yeast6	**0,8856**	0,8574	0,8492	0,8170	0,8029
abalone19	0,6425	0,5828	**0,7019**	0,5532	0,1550
Media	0,7897	0,7845	0,7901	**0,8175**	0,7848

alto de confianza. Por ello, podemos concluir que existen diferencias significativas especificadas en la Tabla 4 con los ranking medios de los algoritmos y su p-valor ajustado calculado según el test post-hoc de Holm. En esta tabla se puede ver que GP-COACH-H obtiene el menor ranking de los algoritmos, lo que lo convierte en el método de control utilizado en el cálculo del procedimiento post-hoc. Como todos los p-valores son bajos, se puede rechazar la hipótesis nula en todos los casos, reforzando así la suposición de que GP-COACH-H es el método que mejor se comporta para conjuntos de datos altamente no balanceados.

Tabla 4: Ranking medio y p-valores ajustados según el procedimiento post-hoc de Holm para los SCBRDs y C4.5

Algoritmo	Ranking medio	p-valor ajustado
GP-COACH-H	2,4091	
GP-COACH-9	3,0227	0,0862
GP-COACH-5	3,0909	0,0862
C4.5	3,2045	0,0549
HFRBCS(Chi)	3,2727	0,0416

5 Conclusiones

En este trabajo hemos presentado un SCBRD con diferentes niveles de granularidad que integra la selección de reglas y el ajuste lateral 2-tuplas para mejorar el rendimiento en conjuntos de datos altamente no balanceados, GP-COACH-H. La propuesta combina el muestreo de datos con modificaciones algorítmicas del enfoque básico y adapta su comportamiento a los diferentes niveles de granularidad considerados, añadiendo un paso de postprocesamiento que ayuda al sistema de clasificación basado en reglas difusas jerárquico a adecuarse al contexto del problema y por tanto, mejorar el comportamiento global.

La propuesta se ha comparado con la versión básica de GP-COACH, el algoritmo HFRBCS(Chi) y con C4.5 para demostrar su buen comportamiento. Los resultados obtenidos muestran el buen rendimiento de la propuesta en el ámbito de la clasificación sobre los conjuntos altamente no balanceados superando al resto de algoritmos de la comparación. Estos buenos resultados se han obtenido gracias a una integración adecuada de las diferentes estrategias que se han usado en la propuesta, ya que han sido adaptadas teniendo en cuenta las características del problema a resolver.

Agradecimientos

Este trabajo ha sido posible gracias a la subvención del Ministerio de Ciencia e Innovación bajo los proyectos TIN2011-28488 y TIN2008-06681-C06-02; y la Junta de Andalucía mediante los proyectos P10-TIC-6858 y TIC-3928. V. López es beneficiaria de una beca FPU del Ministerio de Educación.

Referencias

[1] R. Alcalá, J. Alcalá-Fdez, F. Herrera: A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, *IEEE Transactions on Fuzzy Systems* 15 (4), pp. 616–635, 2007.

[2] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera: KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework, *Journal of Multi-Valued Logic and Soft Computing* 17 (2–3), pp. 255–287, 2011.

[3] R. Barandela, J. S. Sánchez, V. García, E. Rangel: Strategies for learning in class imbalance problems, *Pattern Recognition* 36 (3), pp. 849–851, 2003.

[4] G. E. A. P. A. Batista, R. C. Prati, M. C. Monard: A study of the behaviour of several methods for balancing machine learning training data, *SIGKDD Explorations* 6 (1), pp. 20–29, 2004.

[5] F. J. Berlanga, A. J. Rivera, M. J. del Jesus, F. Herrera: GP-COACH: Genetic programming-based learning of compact and accurate fuzzy rule-based classification systems for high-dimensional problems, *Information Sciences* 180, pp. 1183–1200, 2010.

[6] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer: SMOTE: Synthetic minority over-sampling technique, *Journal of Artificial Intelligent Research* 16, pp. 321–357, 2002.

[7] Z. Chi, H. Yan, T. Pham: *Fuzzy algorithms with applications to image processing and pattern recognition*. World Scientific, 1996.

[8] O. Cordon, F. Herrera, I. Zwir: Linguistic modeling by hierarchical systems of linguistic rules, *IEEE Transactions on Fuzzy Systems* 10 (1), pp. 2–20, 2002.

[9] L. J. Eshelman: The CHC adaptive search algorithm: How to have safe serach when engaging in nontraditional genetic recombination. En: G. J. E. Rawlins (Ed.) *Foudations of Genetic Algorithms*, Morgan Kaufmann, pp. 265–283, 1990.

[10] A. Fernández, M. J. del Jesus, F. Herrera: Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, *International Journal of Approximate Reasoning* 50, pp. 561–577, 2009.

[11] A. Fernández, M. J. del Jesus, F. Herrera: On the 2-tuples based genetic tuning performance for fuzzy rule based classification systems in imbalanced data-sets, *Information Sciences* 180 (8), pp. 1268–1291, 2010.

[12] S. García, A. Fernández, J. Luengo, F. Herrera: A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability, *Soft Computing* 13 (10), pp. 959–977, 2009.

[13] H. He, E. A. Garcia: Learning from imbalanced data, *IEEE Transactions on Knowledge and Data Engineering* 21 (9), pp. 1263–1284, 2009.

[14] H. Ishibuchi, T. Nakashima, M. Nii: *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining*. Springer-Verlag, 2004.

[15] A. Orriols-Puig, E. Bernadó-Mansilla: Evolutionary rule-based systems for imbalanced datasets, *Soft Computing* 13 (3), pp. 213–225, 2009.

[16] J. Quinlan: *C4.5: Programs for Machine Learning*. Morgan Kauffman, 1993.

[17] D. Sheskin: *Handbook of parametric and nonparametric statistical procedures*. Chapman & Hall/CRC, 2006.

[18] Y. Sun, A. K. C. Wong, M. S. Kamel: Classification of imbalanced data: A review, *International Journal of Pattern Recognition and Artificial Intelligence* 23 (4), pp. 687–719, 2009.

Linear and nonlinear regression models based in a concept of distance for imprecise response

Concepción Roldán[1], **Antonio Roldán**[1], **Juan Martínez-Moreno**[2]

[1]*Department of Statistics and Operations, University of Jaén, Spain*, {iroldan,afroldan}@ujaen.es
[2]*Department of Mathematics, University of Jaén, Spain*, jmmoreno@ujaen.es

Abstract

This paper proposes a new approach for the problem of analyzing fuzzy regression models with imprecise response and p real explanatory variables that is simple, gives good results and may be applied in different contexts. In practice there are experiments whose results can be described by means of triangular fuzzy set numbers which can be determined by 3 random values: the center, the left and the right spread. In this context, separate models for the center and the spreads can be analyzed. The fuzzy regression problem is solved using the effective and well-known traditional least-squares method and formulating the fuzzy regression based on a minimum criterion which uses different concepts of distances. The results are illustrated by means of examples.

Keywords: Fuzzy regression, Least-squares method, Distance, Minimization problem.

1 Introduction

Statistical regression analysis plays a fundamental role in decision making in areas such as Economics, Chemistry, Environmental Sciences, etc. Model identification consists in determining, among several models, one that best explains the behavior of the system. This functional relationship between independent (or explanatory) and dependent (or response) variables is performed assuming that the difference between the observed and estimated variables is due to random errors.

In many practical applications in medical science, social, chemistry or economic problems, some useful variables present vagueness/imprecision (see [4]; [17]; [15]). Formally, any $[0,1]$-valued function determines a fuzzy set. However, many researchers prefer to use simple shapes, such as triangular fuzzy numbers, trapezoidal fuzzy numbers, and LR-fuzzy numbers which satisfy the need of modeling fuzzy problems and are easier to fix and handle.

A simpler but important problem that arises in many real-life situations considers that the independent variable can be observed precisely (for instance, the time) and the response is usually described by approximate values instead of exact values.

In relation with fuzzy regression problems, the conventional methods for parameters estimation are: a) linear programming methods; b) least-squares methods; and c) support vector machines (SVMs). Tanaka et al. [19] proposed the formulation of possibilistic linear regression analysis and determined the fuzzy parameters by applying linear programming models. However it is known that this method has several drawbacks (see [13], [3]; [20]; [10]; [11]). It is very sensitive to outliers and some coefficients tend to become crisp due to the characteristics of linear programming. Another problem observed is, when the coefficients are fuzzy numbers the spread of the estimated response becomes wider as the magnitudes of the explanatory variables increase, even though the spreads of the observed responses decrease, or as more observations are included in the model. This contradicts intuition. To prevent this problem, Diamond [3], Wu and Tseng [20] and Kao and Chyu [10] considered numeric coefficients to describe the fuzzy relationship between the fuzzy response variable and fuzzy (or numeric) exploratory variables. All of them used the concept of least squares to determine the regression coefficients.

SVM is a technique for solving pattern recognition problems, used for classification and regression analysis. When SVM is applied to regression estimation problems, it is referred to as support vector regression (SVR). The SVR [9] is a nonparametric, regularized, and nonlinear regression tool which has been applied to optimal control, time-series prediction, interval regression analysis, determination of initial structures for fuzzy neural networks, radial basis function networks, etc.

From the statistical point of view, the least squares method to estimate the boundaries of a fuzzy variable is usually more appropriate and its computation is also often more efficient than the traditional linear programming method [20]. Recently, when both input and output data are assumed to be pairs of trapezoidal fuzzy numbers, simple linear regression models have been solved in terms of the moments of the involved random elements [7]. Least squares problems involving fuzzy data, have also been studied by Näther [14] and Krätschmer [12].

Recent papers develop regression models using fuzzy numbers in the response variable and crisp input. In this respect, Coppi et al. [2] studied the dependence of a LR fuzzy response variable on a set of crisp explanatory variables proposing an iterative least squares estimation procedure which imposes non-negativity restrictions. To overcome the non-negativity restrictions, in 2010, Ferraro et al. [6] solved a linear regression model for imprecise response and p real explanatory variables analyzing the problem within the context of classical multiple and multivariate regression models. A fuzzy regression model has also been analyzed minimizing the sum of the average squared distance between the observed and estimated responses using a few α-cuts [1]. In general, when researchers are dealing with fuzzy random variables (shortly, FRVs), fuzzy regression analysis for constructing the relationship between explanatory and response variables become more complex and they find difficulties in finding optimal solutions, for non-linear problems specially.

This paper offers a new approach. The fuzzy regression will be formulated based on a minimum criterion which uses different known distances between fuzzy numbers. We generalize this case to a more general situation considering other fitted models obtained by least squares method which uses the spreads as response variable. This new technique can be used not only with triangular fuzzy output data, is very easy to compute in practice, may be applied in different contexts and a comparative study with the proposals of different authors shows that the total error obtained is similar or lower to other complex techniques.

The paper is organized as follows. In Section 2 some notations, the fuzzy regression model considered and preliminary results are presented. The procedure estimation is stated in Section 3 for crisp input and triangular fuzzy output data. The aim is to minimize an objective function on training data set rather than minimize an upper bound on the generalization error of a model which is the aim of SVM. Therefore the comparison studies of Section 3 have been done with techniques that have the same aim.

2 The fuzzy regression model for response variable

In this section, we will briefly describe some notations, the fuzzy model considered and the main approaches in the specialized bibliography developed for the linear case.

Consider a random experiment in which a fuzzy response variable $\mathscr{Y}$ and p real explanatory variables $X_1, X_2, ..., X_p$ are observed on n independent units, that is, $\{\underline{X}_i, \mathscr{Y}_i\}_{i=1}^n$ is a random sample obtained from $(\underline{X}, \mathscr{Y})$ with $\underline{X}_i = (X_{1i}, X_{2i}, \ldots, X_{pi})'$. We will focus on the class of *triangular FRV*: $\mathscr{Y} = (Y^I, Y^c, Y^S) \in \mathscr{T}$ where $\mathscr{T}$ denotes the family of triangular fuzzy numbers on $\mathbb{R}$, Y^c is the *center*, and Y^I and Y^S are the *left* and the *right spreads*, that is:

$$\mathscr{Y}(x) = \begin{cases} \frac{x-Y^I}{Y^\ell}, & \text{if } Y^I \le x \le Y^c, \\ \frac{Y^S-x}{Y^r}, & \text{if } Y^c \le x \le Y^S, \\ 0, & \text{in any other case.} \end{cases}$$

We consider asterisk function defined as $\mathscr{Y}^*(x) = Y(x + Y^c)$ for all $x \in \mathbb{R}$, that is, $\mathscr{Y}^* = (-Y^\ell, 0, Y^r)$. Let $\mathscr{T}_0$ denote the set of triangular fuzzy numbers with $Y^c = 0$.

The generic form of the regression model we consider can be formalized as:

$$\mathscr{Y} = f(\underline{X}, \underline{a}) + (\varepsilon^I, \varepsilon^c, \varepsilon^S) \tag{1}$$

where $\underline{a}$ is the vector of the parameters, and $\varepsilon^I, \varepsilon^c$ and ε^S are the residuals, that is, real-valued RVs whose variances are finite.

Since $\mathscr{Y} = (Y^I, Y^c, Y^S)$, separate models for the center and the spreads can be analyzed. The center Y^c can be related to the explanatory variables $X_1, X_2, ..., X_p$ through the classical regression model by a (linear, parabolic, hyperbolic, logarithmic, exponential, etc.) function. The fitted model is denoted by $\widehat{Y}^c_{\underline{x}} = f(\underline{x}, \widehat{\underline{a}}) + \widehat{b}$.

3 The estimation procedure

In this section we develop the theoretical study and illustrate it by means of numerical examples.

3.1 Distances and objective function when the expected values of the spreads do not depend on the explanatory variables

Suppose that the spreads satisfy: $\mathrm{E}(\varepsilon^c - \varepsilon^I | \underline{X}) = I$, $E(\varepsilon^S - \varepsilon^c | \underline{X}) = S$ with $I, S > 0$ being constants. We propose considering a function $\Phi : M \to \mathscr{T}$ defined by

$$\Phi_{\underline{x}} = \widehat{Y}^c_{\underline{x}} + (-I, 0, S), \quad \underline{x} \in M = \mathrm{dom}\,\underline{X}.$$

The total error between this function and the real values can be described by

$$\mathsf{E} = \mathsf{E}(I, S) = \sum_{i=1}^{n} d(\mathscr{Y}_i^*, \Phi^*_{\underline{x}_i}),$$

where $d : \mathscr{T}_0 \times \mathscr{T}_0 \to [0,+\infty]$ is a real distance (in practice, d is a *quasi-distance*). The problem consists in looking for $I,S>0$ solutions of the following minimization problem: $\min(\mathsf{E}) = \min\left\{\sum_{i=1}^{n} d(\mathscr{Y}_i^*, \Phi^*_{\underline{x}_i})\right\}$.

Theorem 1 *Let $\underline{X}, \mathscr{Y}$ be two variables where $\underline{X}$ is a RV and $\mathscr{Y}$ is a FRV fulfilling model (1) and let $\{\underline{X}_i, \mathscr{Y}_i\}_{i=1}^n$ be a random sample obtained from $(\underline{X}, \mathscr{Y})$ with $\mathscr{Y}_i = (Y_i - I_i, Y_i, Y_i + S_i)$. Let $\widehat{Y}^c_{\underline{x}}$ be a fitted model obtained from $\{\underline{X}_i, Y_i\}_{i=1}^n$ and consider the regression problem (in $I,S>0$) for some distance d:*

$$(P)\begin{cases} \Phi_{I,S} : M \to \mathscr{T}, \quad \Phi_{I,S}(\underline{x}) = (\widehat{Y}_{\underline{x}} - I, \widehat{Y}_{\underline{x}}, \widehat{Y}_{\underline{x}} + S), \\ \mathsf{E}(I,S) = \sum_{i=1}^{n} d(\mathscr{Y}_i^*, \Phi^*_{\underline{X}_i}) \qquad \textit{minimum.} \end{cases}$$

- *Using $d_1(A,B) = \|A-B\|_1 = \int_{\mathbb{R}} |A(t) - B(t)|\, dt$, (P) has a unique solution: $\widehat{I} = \text{Median}(I_1, ..., I_n)$ and $\widehat{S} = \text{Median}(S_1, ..., S_n)$.*

- *Using $d_1^2(A,B) = \left(\int_{-\infty}^{0} |A(t) - B(t)|\, dt\right)^2 + \left(\int_{0}^{+\infty} |A(t) - B(t)|\, dt\right)^2$, (P) has a unique solution: $\widehat{I} = \text{Mean}(I_1, ..., I_n)$ and $\widehat{S} = \text{Mean}(S_1, ..., S_n)$.*

- *Using $d_2(A,B) = \|A-B\|_2^2 = \int_{\mathbb{R}} (A(t) - B(t))^2\, dt$, (P) has a unique solution. The parameter estimators, $\widehat{I}$ and $\widehat{S}$, are the only positive solutions of $n-1$ cubic equations:*

$$\left\{ \begin{array}{c} 2\widehat{I}^3 \sum_{i=r+1}^{n} \frac{1}{I_i} + (3r-2n)\widehat{I}^2 - \sum_{i=1}^{r} I_i^2 = 0 \\ \textit{in the interval } [I_r, I_{r+1}] \end{array} \right\}_{r=1}^{r=n-1},$$

$$\left\{ \begin{array}{c} 2\widehat{S}^3 \sum_{i=r+1}^{n} \frac{1}{S_i} + (3r-2n)\widehat{S}^2 - \sum_{i=1}^{r} S_i^2 = 0 \\ \textit{in the interval } [S_r, S_{r+1}] \end{array} \right\}_{r=1}^{r=n-1},$$

respectively.

- *Using $d_\infty(A,B) = \sup(\{|A(t) - B(t)| : t \le 0\}) + \sup(\{|A(t) - B(t)| : t \ge 0\})$, (P) has solution. The parameter estimator, $\widehat{I}$, is the value $I_1, I_2, \ldots, I_n$ where the minimum of the following n real numbers is found:*

$$\left\{ n - 1 - \frac{1}{I_j} \sum_{i=1}^{j-1} I_i - I_j \sum_{i=j+1}^{n} \frac{1}{I_i} \right\}_{j=1}^{n}$$

and the parameter estimator, $\widehat{S}$, is the value $S_1, S_2, \ldots, S_n$ which reaches the minimum of a similar numerical set.

A simulation experiment was carried out in order to evaluate the efficiency of the parameter estimates obtained with the different distances considering linear and non-linear problems and different sample sizes. For robustness comparison, the residuals are normally distributed for the center and chi squared distributed for the spreads. The simulation experiment was repeated 10000 times for obtaining the mean squared error (MSE) of each estimate.

Firstly, we consider the following linear regression model:

$$\mathscr{Y} = 0.7X + \left(\varepsilon^I, \varepsilon^c, \varepsilon^S\right),$$

where X is a RV varying from 1 to 100 one at a time and $\varepsilon^c \sim N(1, 0.5)$, $\varepsilon^c - \varepsilon^I \sim \chi^2_{30}/40$, $\varepsilon^S - \varepsilon^c \sim \chi^2_{15}/10$, that is, the *regression function* at x is

$$\widehat{\mathscr{Y}}(x) = 0.7x + (0.25, 1, 2.5).$$

The average of the estimates and the estimated mean squared error for each of them calculated as $\text{MSE}(\theta) = (1/10000)\sum_{r=1}^{10000}(\theta - \hat{\theta})^2$, are presented in Table 1. The simple linear regression model for fuzzy random variables has also been studied by other authors [7]. The results obtained by this other method are presented in the last row of Table 1. Comparing the results obtained using the different distances and the method proposed by other authors we can conclude that d_1^2 and the method proposed by other authors produce the best estimates.

The next simulation experiment considers the following non-linear regression model:

$$\mathscr{Y} = 1.1X^2 + 0.5X + \left(\varepsilon^I, \varepsilon^c, \varepsilon^S\right),$$

where X is a RV varying from 1 to 20 (or 100) one at a time and $\varepsilon^c \sim N(-0.4, 0.5)$, $\varepsilon^c - \varepsilon^I \sim \chi^2_{30}/40$, $\varepsilon^S - \varepsilon^c \sim \chi^2_5/10$, that is, the *regression function* at x is given by the following quadratic functions:

$$\widehat{\mathscr{Y}}(x) = 1.1x^2 + 0.5x + (-1.15, -0.4, 0.1)$$

The average of the estimates and the MSE for each of them are presented in Table 2. The results obtained showed that d_1^2 gave the best estimates again.

3.2 Comparison study: Fuzzy regression when the expected values of the spreads depend on the explanatory variables

In this section an approach based on the results obtained in the previous section is presented. Suppose that

$$E(\varepsilon^c - \varepsilon^I | \underline{X}) = I(\underline{X}), \quad E(\varepsilon^S - \varepsilon^c | \underline{X}) = S(\underline{X}).$$

We propose to make the expected values of the spreads depend on the explanatory variables fitting the data to a model by least squares method considering as the response variable the values of the spreads on each side respectively. The fitted models for both functions obtained by least squares method will be denoted by $\widehat{I}(\underline{x})$ and $\widehat{S}(\underline{x})$

Table 1: Monte Carlo estimates and estimated MSEs of the parameters using sample sizes $n = 100$

	$\widehat{a}$	MSE$(a) \times 10^6$	$\widehat{b}$	MSE$(b) \times 10^6$	$\widehat{I}$	MSE$(I) \times 10^3$	$\widehat{S}$	MSE$(S) \times 10^3$
d_1	0.7000	3.063	1.0002	0.0104	0.734	0.827	1.435	8.556
d_1^2	Idem	Idem	Idem	Idem	0.750	0.372	1.500	2.992
d_2	Idem	Idem	Idem	Idem	0.739	0.495	1.456	4.805
d_∞	Idem	Idem	Idem	Idem	0.736	0.921	1.449	8.957
GBCL	0.7000	3.508	1.0010	0.0116	0.750	0.373	1.500	2.992

Table 2: Monte Carlo estimates and estimated MSEs of the parameters using sample sizes $n = 100$

	$\widehat{\underline{a}} = (\widehat{a}_1, \widehat{a}_2)$	MSE$(\underline{a}) \times 10^6$	$\widehat{b}$	MSE(b)	$\widehat{I}$	MSE$(I) \times 10^3$	$\widehat{S}$	MSE$(S) \times 10^3$
d_1	(0.500, 1.100)	(1.919,2.728)	-0.401	0.0231	0.734	0.843	0.435	5.372
d_1^2	Idem	Idem	Idem	Idem	0.750	0.377	0.499	0.992
d_2	Idem	Idem	Idem	Idem	0.739	0.502	0.458	2.667
d_∞	Idem	Idem	Idem	Idem	0.736	0.937	0.457	4.321

The applicability of the proposed method is shown with a comparison study. Tanaka et al. [19] designed an example to illustrate their fuzzy regression model for dealing with the problem of one numeric exploratory variable and one fuzzy dependent variable with symmetric triangular fuzzy numbers and considered the following five pairs of observations (listed in Table 3).

Table 3: Data and spreads

X	Y^I	Y^c	Y^S	I_i	S_i
1	6.2	8	9.8	1.8	1.8
2	4.2	6.48	6	2.2	2.2
3	6.9	9.5	12.1	2.6	2.6
4	10.9	13.5	16.1	2.6	2.6
5	10.6	13	15.4	2.4	2.4

Maybe these datasets do not have the level of detail and complexity than those used in other studies, but in the literature these data have been considered by many researchers for the experimental evaluation and comparison of their proposed methodology. Several authors pointed out that the first fuzzy regression model (proposed by Tanaka et al. in 1982) has several disadvantages and modified it or developed their own methodologies to prevent the problems. THW regression model [19] estimates the fuzzy regression coefficients by linear programming. This model has a numeric slope and a fuzzy intercept. KB, DM, WT and HBS models ([11], [3], [20] and [8], respectively) have a fuzzy slope and intercept. If the explanatory variables are numeric values and dependent variables are symmetric fuzzy numbers, WT model is the same as DM approach. NN and CH models ([13] and [1], respectively) have a numeric slope and KC model [10] has numeric coefficients and a fuzzy error term. Table 4 lists the regression models obtained by the methods proposed by these authors based on the five pairs of observations considered in Table 3.

With our methodology, fitting a fuzzy linear model, the parameter estimates are $\widehat{I} = \widehat{S} = 2.32$ and the linear model obtained is given by $\widehat{\mathscr{Y}}(x) = (2.63, 4.95, 7.27) + 1.71x$. However, considering that the spreads depend on the explanatory variables by a non-linear model, we have obtained the following linear expression for the center:

$$\widehat{Y}^c(x) = 4.95 + 1.71x, \quad \rho_{\mathscr{Y}^c X} = 0.872,$$

(where ρ denotes the correlation coefficient) and the following non-linear expression for the spreads:

$$\widehat{\mathrm{I}}(x) = \widehat{\mathrm{S}}(x) = \frac{1}{0.344 + 0.208/x}, \quad \rho_{IX} = 0.949.$$

These two models were chosen (using the statistical package SPSS) among several models considered as the best. Therefore the fuzzy regression model obtained using the methodology proposed in this paper is given by:

$$\widehat{\mathscr{Y}_{RRM}}(x) = \widehat{Y}^c(x) + (-\widehat{I}(x), 0, \widehat{S}(x)).$$

Let E_i denote the mean squared error between the estimated and observed fuzzy responses for the ith case. Table 5 shows the total errors in estimating the fuzzy responses for all the models based on the mean squared error criterion. Other criteria can also be considered. The last column gives the results obtained by our methodology. The total estimation error, $\sum_i E_i$, of the proposed approach is the lowest, 9.238. Note that the model obtained only gives a linear expression for the center. For the linear model ($\widehat{\mathscr{Y}}(x) = (2.63, 4.95, 7.27) + 1.71x$), the total estimation error is 9.485, a reasonable result in comparison with the errors obtained applying other approaches.

Analyzing these results, we can observe that our technique is very easy to compute in practice and gives a first and easy

Table 4: Regression models obtained by others authors for the five pairs of observations considered in Table 3

Author	Fuzzy regresion model
Tanaka et al. [19]	$\widehat{\mathscr{Y}_{THW}}(x) = (0, 3.85, 7.7) + 2.1x$
Kim and Bishu [11]	$\widehat{\mathscr{Y}_{KB}}(x) = (3.11, 4.95, 6.84) + (1.55, 1.71, 1.82)x$
Diamond [3]	$\widehat{\mathscr{Y}_{DM}}(x) = (3.11, 4.95, 6.79) + (1.55, 1.71, 1.87)x$
Nasrabadi and Nasrabadi [13]	$\widehat{\mathscr{Y}_{NN}}(x) = (2.36, 4.86, 7) + 1.73x$
Wu and Tseng [20]	$\widehat{\mathscr{Y}_{WT}}(x) = (3.11, 4.95, 6.79) + (1.55, 1.71, 1.87)x$
Hojati et al. [8]	$\widehat{\mathscr{Y}_{HBS}}(x) = (5.1, 6.75, 8.4) + (1.1, 1.25, 1.4)x$
Kao and Chyu [10]	$\widehat{\mathscr{Y}_{KC}}(x) = 4.81 + 1.72x + (-2.2, 0.12, 2.44)$
Chen and Hsueh [1]	$\widehat{\mathscr{Y}_{CH}}(x) = 1.71x + (2.63, 4.95, 7.27)$

Table 5: Comparison of the total estimation errors from several models

Explan.	Resp.	E_i								
var.	var.	THW	KB	DM	KC	NN	HBS	WT	CH	RRM
1	(6.2, 8, 9.8)	6.304	1.816	1.816	1.974	2.663	0.000	1.816	1.931	1.796
2	(4.2, 6.48, 6)	4.084	3.834	3.882	3.857	3.035	8.154	3.882	3.888	3.881
3	(6.9, 9.5, 12.1)	1.204	0.363	0.376	0.376	0.176	1.125	0.376	0.376	0.358
4	(10.9, 13.5, 16.1)	2.344	3.074	2.931	2.936	3.649	3.124	2.931	2.963	2.928
5	(10.6, 13, 15.4)	2.874	0.215	0.279	0.269	0.112	0.000	0.279	0.253	0.275
	$\sum_i E_i$	16.809	9.302	9.283	9.412	9.636	12.403	9.283	9.411	9.238

approach for the problem of analyzing regression problems with crisp input and triangular fuzzy output data.

Acknowledgments

This paper has been partially supported by Junta de Andalucía and projects FQM-268, FQM-245, FQM-178 of the Andalusian CICYE, Spain.

References

[1] L.H. Chen, C.C. Hsueh: Fuzzy regression models using the least-squares method based on the concept of distance. *IEEE Transactions on Fuzzy Syst* 17, no.6, pp. 1259–1272, 2009.

[2] R. Coppi, P. D'Urso, P. Giordani, A. Santoro: Least squares estimation of a linear regression model with LR fuzzy response. *Comp Stat Data Anal* 51, pp. 267–286, 2006.

[3] P. Diamond: Fuzzy least squares. *Inform Sci* 46, pp. 141–157, 1988.

[4] A. Colubi: Statistical inference about the means of fuzzy random variables: applications to the analysis of fuzzy- and real-valued data. *Fuzzy Sets Syst* 160, pp. 344–356, 2009 .

[5] I. Couso, L. Sánchez: Higer order models for fuzzy random variables. *Fuzzy Set and Syst* 159, pp. 237–258 2007

[6] M.B. Ferraro, R. Coppi, G. González-Rodríguez, A. Colubi: A linear regression model for imprecise response. *International Journal of Approximate Reasoning* 51, pp. 759–770, 2010.

[7] G. González-Rodríguez, A. Blanco, A. Colubi, M.A. Lubiano: Estimation of a simple linear regression model for fuzzy random variables. *Fuzzy Sets and Syst* 160, pp. 357–370, 2009.

[8] M. Hojati, C.R. Bector, K. Smimou: A simple method for computation of fuzzy linear regression, *Eur J Oper Res*, 166, pp. 172–184, 2005.

[9] D.H. Hong, C. Hwang: Support vector fuzzy regression machines. *Fuzzy Set and Syst* 138, pp. 271–281, 2002.

[10] C. Kao, C.L. Chyu: Least-squares estimates in fuzzy regression analysis. *Eur J Oper Res* 148, pp. 426–435, 2003.

[11] B. Kim, R.R. Bishu: Evaluation of fuzzy linear regression models by comparing membership functions, *Fuzzy Sets and Syst* 100, pp. 343–352, 1998.

[12] V. Krätschmer: Limit distribution of least squares estimators in linear regression models with vague concepts, *J Multivariate Anal* 97, pp. 1044–1069, 2006.

[13] M.M. Nasrabadi, E. Nasrabadi: A mathematical-programming approach to fuzzy linear regression analysis. *Appl Math Comput*, 155, pp. 873–881, 2004.

[14] W. Näther: Regression with fuzzy random data. *Comp Stat Data Anal* 51, pp. 235–252, 2006.

[15] A.M. Palacios, L. Sánchez, I. Couso: Diagnosis of dyslexia with low quality data with genetic fuzzy systems. *International Journal of Approximate Reasoning* 51, pp. 993–1009, 2010.

[16] C. Roldán, A. Roldán, J. Martínez-Moreno: A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: A case study in biomass production. *Soft Computing*, DOI: 10.1007/s00500-011-0769-1.

[17] L. Sánchez, I. Couso, J. Casillas: Genetic learning of fuzzy rules based on low quality data. *Fuzzy Sets and Syst* 160, pp. 2524–2552, 2009.

[18] H. Tanaka, I. Hayashi, J. Watada: Linear regression analysis with fuzzy model. *IEEE Trans Syst Man Cybern*, SMC-12, pp. 903–907, 1982.

[19] H. Tanaka, I. Hayashi, J. Watada: Possibilistic linear regression analysis for fuzzy data. *Eur J Oper Res* 40, pp. 389–396, 1989.

[20] B. Wu, N.F. Tseng: A new approach to fuzzy regression models with application to business cycle analysis. *Fuzzy Sets and Syst* 130, pp. 33–42, 2002.

[21] L.A. Zadeh: Fuzzy sets. *Inform Control* 8, pp. 338–353, 1965.

Extracción de reglas de asociación difusas a partir de datos de baja calidad

A.M. Palacios[1], **J. Alcalá-Fdez**[2]

[1] *Universidad de Oviedo, Departamento de Informática, 33204 Gijón, España*, palaciosana@uniovi.es
[2] *Universidad de Granada, Departamento de Ciencias de la Computación e Inteligencia Artificial, CITIC-UGR 18071 Granada, España*, jalcala@decsai.ugr.es

Resumen

Una de las técnicas más utilizas de la Minería de Datos consiste en inducir reglas de asociacion a partir de base de datos. Estas reglas pueden ayudarnos a analizar los datos y a tomar buenas decisiones dentro del ámbito del problema. Muchos de los trabajos han sido enfocados sobre bases de datos con valores precisos, sin embargo en muchas aplicaciones del mundo real los datos presentan un cierto grado de imprecisión. Además, en algunas ocasiones esta impresión puede llegar a ser significativa y no es natural modelarla con una distribucion de probabilidad. En este trabajo proponemos un nuevo algoritmo capaz de extraer conocimiento interesante a partir de bases de datos con datos imprecisos. Este algoritmo integra conceptos de imprecisión con el algoritmo Apriori difuso con el objetivo de obtener reglas de asociación difusas relevantes. Los resultados obtenidos en un problema real sobre el rendimiento de los deportistas en el atletismo muestran la efectividad del método propuesto.

Palabras Clave: Minería de Datos, Reglas de Asociación Difusas, Datos de Baja Calidad

1. Introducción

La Minería de Datos (MD) consiste en extraer conocimiento interesante a partir de conjuntos de datos del mundo real y es el paso central del proceso de Extracción de Conocimiento a partir de Base de Datos (BD). El descubrimiento de asociaciones es una de las técnicas de la MD que más ha sido utilizada para extraer conocimiento interesante a partir de BD.

Las reglas de asociación son utilizadas para representar e identificar dependencias entre ítems en una BD [18]. Estas reglas son expresiones del tipo $X \rightarrow Y$, donde X e Y son conjuntos de ítems que cumplen $X \cap Y = \emptyset$, es decir, si todos los ítems de X existen en una transacción entonces todos los ítems de Y, con una alta probabilidad, están en la transacción, y donde X e Y no tienen ningún ítem en común [1, 2]. Últimamente, la teoría de los conjuntos difusos ha sido utilizada más frecuentemente para describir relaciones entre los datos [9]. El uso de conjunto difusos nos permite extender los tipos de relaciones que se pueden representar, facilitando la interpretación de las reglas en términos lingüísticos y eludiendo las fronteras no naturales en el particionamiento de los dominios de los atributos [5, 6].

Muchos investigadores han propuesto métodos para la extracción de reglas difusas a partir de datos cuantitativos [8, 12, 17]. Estos métodos han sido enfocados sobre BD con valores precisos y exactos, sin embargo en muchas aplicaciones reales los datos presentan un cierto grado de imprecisión. En ocasiones, esta imprecisión es lo suficientemente pequeña y puede ser ignorada sin peligro. En otras ocasiones, la incertidumbre de los datos puede ser modelada con una distribución de probabilidad. Sin embargo, existen otro tipo de problemas donde la imprecisión es relevante y no puede ser modelada mediante una distribución de probabilidad [3]. Por esta razón, diseñar algoritmos que sean capaces de extraer conocimiento interesante a partir de Datos de Baja Calidad (DBC) representa un desafio para los investigadores [15, 16].

La estadística difusa considera el uso de los conjuntos difusos para representar la información imprecisa de los datos. Recientes trabajos en estadística difusa sugieren el uso de una representación difusa cuando el dato es desconocido a través de una familia de intervalos de confianza [4], considerando una representación posibilística para modelar los datos imprecisos.

En este trabajo, integramos conceptos de DBC con el algoritmo Apriori difuso propuesto por Hong y otros en [8] con el objeto de obtener reglas de asociación difusas de alta calidad a partir de BD con DBC. Extenderemos este algoritmo considerando una representación posibilística para modelar datos de entrada con valores imprecisos. Así, una entrada imprecisa será representada por un conjunto difuso, el cual será identificado por la familia de todas las distribuciones de probabilidad [4, 16]. La función de pertenencia de esta entrada imprecisa será otro conjunto difuso de acuerdo con el Principio de Extensión [11], el cual es compatible con la interpretación posibilística de los conjuntos difusos [16]. Esto implicará, que la confianza de una regla de asociación difusa será definida a partir de un conjunto de probabilidades.

Los resultados obtenidos sobre un problema real con DBC, basado en el rendimiento de los atletas en las pruebas de atletismo de 100 metros lisos [14], muestran la efectividad del método propuesto.

La estructura del trabajo es como sigue. En la siguiente sección se describe el algoritmo Apriori difuso propuesto por Hong y otros. La Sección 3 introduce la interpretación y representación posibilística de los DBC. El método propuesto es descrito en detalle en la Sección 4. La Sección 5, muestra los resultados obtenidos por el método propuesto sobre una BD real con DBC. Finalmente, la Sección 6 muestran algunas conclusiones.

2. Algoritmo Apriori Difuso

El objetivo principal del algoritmo Apriori difuso, introducido en [8] por Hong y otros, es encontrar ítems relevantes así como reglas de asociación difusas en las instancias con valores cuantitativos, descubriendo interesantes patrones.

Este método consiste en transformar cada valor cuantitativo en un conjunto difuso de etiquetas lingüísticas asumiento que las funciones de pertenencia son conocidas de antemano. El algoritmo posteriormente calcula la cardinalidad de cada item difuso, a lo que denomina "cuenta". Si el valor de cuenta del item difuso es superior o igual que el valor del mínimo soporte este item será considerado un item difuso frecuente. A continuación combina los items frecuentes y vuelve a repetir el proceso. Finalmente, este método obtiene las reglas de asociación difusas mediante el criterior del algoritmo Apriori [2].

3. Datos de Baja Calidad: Representación e Interpretación

En la actualidad existen problemas reales donde la información viene representada por DBC. En estos DBC no se puede observar con precisión las propiedades de un objeto, por lo que no se percibe con exactitud el valor del objeto ni se tiene un conocimiento completo sobre su distribución de probabilidad.

Para modelar y tratar estos datos imprecisos, en este trabajo consideramos una representación posibilísitica. De esta forma, una entrada imprecisa será representada por un conjunto difuso identificado por las familias de distribuciones de probabilidades, siendo cada α-corte de este conjunto difuso un conjunto aleatorio que contiene el valor exacto y desconocido de la variable con una probabilidad de al menos 1-α [4, 16] (ver Figura 1). Resaltar que, esta representación incluye los valores intervalares y precisos como casos particulares.

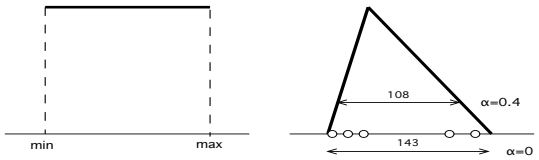


Figura 1: Representación difusa de datos imprecisos.

Esta representación proporciona un marco común para tratar con datos numéricos, palabras, intervalos, conjuntos difusos, valores perdidos (mediante un intervalo que agrupe todo el rango de la variable) o diferentes valores para un mismo atributo como se describe en [16], donde todos estos conceptos son agrupados bajo el término DBC.

En esta propuesta la representación y dominio de cada item de baja calidad puede ser definido mediante:

- Un intervalo $\overline{X} = [x_1, x_2]$ donde x_1 y x_2 están incluidos en el dominio del item. Por ejemplo, un item con dominio entre [0'0, 10'0] podrá ser definido como $\overline{X}$ = [1'5,3'6] o mediante un valor perdido $\overline{X}$ = [0'0, 10'0] (implicando que "x $\in \overline{X}$").
- Un conjunto difuso $\widetilde{X} = (x_1, x_2, x_3)$. Por ejemplo, un item definido mediante tres valores diferentes $\widetilde{X}$=(1'0,1'3,2'0).

4. Función de Pertenencia Difusa con Datos de Baja Calidad

En esta sección describiremos como se obtiene la función de pertenencia a partir de datos imprecisos. Si se tiene una percepción precisa "x" de las propiedades de un objeto y un conjunto difuso $\widetilde{A}$ con un conjunto finito de etiquetas lingüísticas, $L = \{l_1, \ldots, l_n\}$, donde "n" es el número de etiquetas, la función de pertenencia será:

$$pert(x)(l_i) = P_x(l_i) \mid \sum_{i=1}^{n} P_x(l_i) = 1. \quad (1)$$

Si el objeto observado es impreciso y toda la información que se tiene es que está en el conjunto $\overline{X}$ ($x \in \overline{X}$), la función de pertenencia de este conjunto de valores será un conjunto de funciones de pertenencias:

$$pert(\bar{X})(l_i) = \{pert(x)(l_i) \mid x \in \bar{X}\}. \quad (2)$$

Por ejemplo, imaginemos que disponemos de la entrada $\overline{X}$ = [2'1,3'0] y que se quiere calcular la función de pertenencia en la etiqueta "Baja". Toda la información de la que se dispone es que el valor real y desconocido (x_0) pertenece a $x_0 \in$ [2'1,3'0] y donde cada α-corte sobre $\overline{X}$ es un conjunto aleatorio que contiene el valor preciso del objeto con una probabilidad de al menos 1-α. En este caso, al tratarse de un valor intervalar, cada α-corte tiene una probabilidad del 100 % de contener el valor real x_0 ($P(x_0 \in$[2'1,3'0])=1). Esto implica que $pert$([2'1,3'0])(Baja) será un conjunto de probabilidades obtenidas a partir de cada valor x, donde $x \in$ [2'1,3'0].

En el caso de que la entrada sea un valor impreciso $\widetilde{X}$, la función de pertenencia será un conjunto difuso de acuerdo con el Principio de Extensión [11], el cual es compatible con la interpretación posibilística de los conjuntos difusos [16]:

$$\begin{aligned} \text{pert}(\widetilde{X})(l_i)(\text{t}) = \max\{\alpha \mid \text{t} = \text{fuzz}(\text{x})(l_i) \ \text{ y} \\ \text{x} \in [\widetilde{X}]_\alpha\}_{\alpha \in [0,1]}. \end{aligned} \quad (3)$$

Como podemos observar, el conjunto difuso $pert(\widetilde{X}(l_i))$ está asociado al conjunto de familias $\{pert(\widetilde{X}_\alpha(l_i))\}_{\alpha \in [0,1]}$.

5. Extracción de Reglas de Asociación difusas con Datos de Baja Calidad

Esta sección describe nuestra propuesta para obtener reglas de asociación difusas a partir de BD con DBC. Los parámetros de entrada para esta propuesta son:

- Un conjunto de DBC, $\widetilde{D}$, compuesto de t instancias, donde cada una de ellas contiene m atributos, y donde $\widetilde{X}_j^i$ representa el ítem j ($1 \le j \le m$) en la instancia i ($1 \le i \le t$).
- El conjunto S=$\{L_1, \ldots, L_m\}$, donde L_j=$\{l_1, \ldots, l_n\}$ representa el conjunto finito de etiquetas lingüísticas del atributo j siento n el número de etiquetas lingüísticas.
- Un valor predeterminado del mínimo soporte α.
- Un valor predeterminado de la confianza λ.
- El número de cortes disponibles para obtener los posibles valores reales de la variable.
- El número de veces γ que barremos las probabilidades de cada uno de los posibles valores obtenidos (ver paso 10 del método).

A continuación se describe el algoritmo propuesto:

P1: Transformar cada DBC $\widetilde{X}_j^i$, mediante la función de pertenencia, en un conjunto difuso interpretado como un conjunto de probabilidades para cada etiqueta de L_j ($\overline{P_j^i}$=$\{[p_{*1}, p_1^*]_{l_1}^i, \ldots, [p_{*n}, p_n^*]_{l_n}^i\}$).

P2: Calcular la frecuencia de ocurrencias del ítem j en cada término lingüístico k de L_j (L_{jk}), donde $L_j = \{l_1,\ldots,l_k,\ldots,l_n\}$ y $L_{j_k} = l_k$.

$$\overline{Cuenta}_{L_{j_k}}(\,\%) = \tfrac{1}{t} \bigoplus_{i=1}^{t} [p_{*k}, p_k^*]_{l_k}^i \quad (4)$$

donde t representa el número de instancias y $\bigoplus$ es la suma aritmética difusa [10].

Todos los L_{j_k} son recolectados para formar el conjunto de candidatos C_r donde r representa el número de ítems relacionados en el conjunto de candidatos, inicialmente r=1:

$$\begin{aligned} C_r = \{\cup\{L_{j_k},\ \forall \text{k} \mid 1 \le k \le n\}, \\ \forall \text{j} \mid 1 \le j \le m\}. \end{aligned} \quad (5)$$

P3: Comprobar si $\overline{Cuenta}_{L_{j_k}}(\,\%)$ para todo L_{j_k} de C_r es mayor o igual que el mínimo soporte α. Si $\overline{Cuenta}_{L_{j_k}}(\,\%)$ satisface esta condición entonces el conjunto L_r contendrá a L_{j_k}. $\overline{Cuenta}_{L_{j_k}}(\,\%)$ esta representado por un conjunto de probabilidades lo que implica que, esta condición será satisfacida si el límite superior (p_k^*) es mayor o igual que α. De este modo, todas las posibles ocurrencias del conjunto de ítems serán consideradas. Por tanto, L_r será el conjunto:

$$L_r = \{L_{j_k} \mid p_k^* \ge \alpha,\ L_{j_k} \in C_r\}. \quad (6)$$

P4: Si L_r no es nulo, entonces continuar con el algoritmo; en otro caso, salir del algoritmo.

P5: Generar el conjunto de candidatos C_{r+1} a partir de la unión de los ítems que componen L_r. C_{r+1} se obtiene de forma similar a la indicada por el algoritmo Apriori [2], con la excepción de que dos ítems con el mismo atributo j no pueden formar parte de C_{r+1} [8].

P6: Realizar los siguientes acciones para cada "(r+1)-itemset" obtenido en C_{r+1}:

a) Calcular la función de pertenencia de cada "(r+1)-itemset", $\overline{P}^i_s = \overline{P}^i_1 \wedge \ldots \wedge \overline{P}^i_{(r+1)}$. La t-normal del producto generaliza la agregación o combinación entre el conjunto de probabilidades de los ítems de "(r+1)-itemset":

$$\overline{P}^i_s = \bigotimes_{j=1}^{(r+1)} \overline{P}^i_j. \quad (7)$$

b) Calcular la frecuencia de ocurrencias de cada "(r+1)-itemset":

$$\overline{Cuenta}_s(\%) = \frac{1}{t} \bigoplus_{i=1}^{t} \overline{P}^i_s \quad (8)$$

Si $max\{\overline{Cuenta}_s(\%)\}$, es mayor o igual que el mínimo soporte α entonces, "(r+1)-itemset" forma parte del conjunto L_{r+1}.

P7: Si L_{r+1} es nulo continuar con el siguiente paso; en caso contrario actualizar el número de ítems relacionados en el conjunto de candidatos (r=r+1) y repetir los pasos 5 y 6.

P8: Recolectar en R los items relacionados de cada L_i, donde $2 \leq i \leq (r+1)$.

P9: Construir todas las reglas de asociación difusas (X $\rightarrow$ Y) a partir del conjunto R.

P10: Determinar, mediante los siguientes dos pasos, si las reglas de asociación difusas obtenidas son relevantes y proporcionan patrones e información interesante de los DBC:

- Calcular la confianza de la regla.
- Comparar la confianza obtenida en el paso anterior con la mínima confianza λ.

La presencia de entradas imprecisas implica que la confianza de una regla vendrá definida por un valor intervalar entre [0,1]:

$$\begin{aligned} &\overline{Confianza}(X \rightarrow Y)_{(\bar{P}_{s_1},\ldots,\bar{P}_{s_q})} = \\ &= \{\text{Confianza}(X \rightarrow Y)_{(x_{s_1},\ldots,x_{s_q})} \mid \\ &\quad Cuenta_{anteced.} > 0, \forall x_{s_j} \in \bar{P}_{s_j}\} \end{aligned} \quad (9)$$

donde Confianza($X \rightarrow Y$) es definida en [8] y $Cuenta_{anteced}$ representa la frecuencia de ocurrencias del antecedente de la regla.

El coste computacional de la ec. (9) es muy elevado, y contiene el valor real y desconocido de la confianza de la regla que dependiendo de los valores de x_{s_j}, la regla podría ser o no considerada relevante. Por esta rarón, nosotros proponemos la siguiente aproximación. Si se considera la regla $(X \rightarrow Y)$ con $\overline{Cuenta}_s = [x_1, x_2] = \overline{X}$ y $\overline{Cuenta}_{anteced.} = [y_1, y_2] = \overline{Y}$, algunos "cortes" (c) son aplicados para obtener los posibles valores reales de $\overline{X}$ e $\overline{Y}$, donde a cada corte c (x_j) se le asigna una probabilidad aleatoria (P_{x_j}) de ser el valor real, y donde la suma de todas las probabilidades, de todos los cortes, tiene que ser 1. Estos posibles valores reales de $\overline{X}$ e $\overline{Y}$ son V_X=$\{\overline{X}_w$, w =1,. . .,c$\}$ y V_Y=$\{\overline{Y}_w$, w =1,. . .,c$\}$.

A partir de un valor de V_Y ($y_j \in V_Y$) y de un valor del conjunto V_X ($x_j \in V_X$), para determinar si una regla es relevante no solamente hay que satisfacer la condición $x_j \geq y_j * \lambda$, $x_j \leq y_j$, $y_j > 0$, sino que, además, hay que considerar que la probabilidad de que x_j sea el valor real tiene que ser superior a 0,5 ($P_{x_j} \geq 0{,}5$). Resaltar que, para cada valor de V_Y ($y_j \in V_Y$), todos los posibles valores de V_X han sido considerados y un nuevo conjunto C = $\{C_{y_j} \mid y_j \in V_Y\}$ es obtenido a partir de todas las probabilidades que determinan si una regla es relevante o no a partir de cada valor de V_Y ($y_j \in Y$):

$$\begin{aligned} C_{y_j} = \textstyle\sum_{i=1}^{c} P_{x_i} \geq 0{,}5 \mid &x_i \geq y_j * \lambda, \\ &x_i \leq y_j,\ y_j > 0. \end{aligned} \quad (10)$$

Este proceso es repetido γ veces para barrer las posibles probabilidades de cada posible valor de V_X. Así, para cada valor de V_Y se obtiene un conjunto de posibles probabilidades que dependen de las probabilidades asignadas a cada uno de los posibles valores de V_X:

$$\overline{C}_{y_j} = \{C_{y_j}, \forall C_i,\ 1 \leq i \leq \gamma\}.$$

Finalmente, para determinar si la regla es relevante estos conjuntos de posibles probabilidades son ordenados, de acuerdo a la dominancia uniforme definida en [13]. El conjunto que representa la mediana será seleccionado $\overline{C}_{y_j}$ para determinar a la regla como relevante siempre y cuando su probabilidad mínima sea superior o igual a 0,5.

6. Experimentos

En esta sección se muestra la efectividad del método propuesto a partir de un problema real con DBC basado en el rendimiento de los atletas en la prueba de atletismo de 100 metros lisos ("100mlI") [14].

En las siguientes subsecciones se analizarán las reglas de asociación difusas obtenidas según el valor del mínimo soporte y de la mínima confianza. Además, se analizará el conocimiento obtenido por los expertos a partir de las reglas generadas por el método propuesto.

Las particiones linguísticas consideradas están compuestas por tres términos lingüísticos, con funciones de pertenencia triangulares uniformemente distribuidas. Los valores considerados como parámetros del método son: 7 α-cortes y 2 barridos de las probabilidades (γ).

Cuadro 1: Nivel de conocimiento e interpretabilidad de las reglas de asociación difusas para "100mlI".

Reglas	Sorp-Regla	$\{\overline{C}_{yi}\}$	Mediana$\{\overline{C}_{yi}\}$
Si Salto Es Medio y Reacción Es Media Entonces Carrera Es Media	[0'229,0'324]	{[0'19,0'4],[0'19,0'4] [0'50,0'76],[0'933,0'97], [0'96,0'97],[0'97,0'97],[1'0,1'0]}	[0'933,0'971]
Si Salto Es Medio y Atleta Es Relevante Entonces Carrera Es Media	[0'15,0'24]	{[0'046,0'167],[0'096,0'261] [0'09,0'26],[0'6,0'87], [0'73,0'90],[0'83,0'95],[0'94,0'99]}	[0'6,0'87]
Si Reacción Es Baja Entonces Carrera Es Media	[0'09,0'14]	{[0'92,0'99],[0'92,0'99],[0'94,0'99] [1'0,1'0],[1'0,1'0],[1'0,1'0],[1'0,1'0]}	[1'0,1'0]
Si Salto Es Bajo y Carrera Es Media Entonces Atleta Es No Relevante	[0'11,0'179]	{[0'0,0'38],[0'43,0'44] [0'43,0'44],[0'71,0'99], [0'82,1'0],[0'82,1'0],[0'92,1'0]}	[0'718,0'999]
Si Reacción Es Baja Entonces Atleta Es Relevante	[0'09,0'13]	{[0'08,0'08],[0'15,0'47],[0'15,0'47] [0'52,0'81],[0'52,0'81],[0'99,0'99],[1'0,1'0]}	[0'524,0'816]

6.1. Análisis de las reglas de asociación difusas según soporte y confianza

Varios experimentos se han realizado para analizar el número de reglas de asociación difusas obtenidas a partir del algoritmo propuesto capaz de soportar DBC. La relación entre el número de reglas de asociación difusas con respeto a varios valores del mínimo soporte considerando distintas confianzas es mostrada en la parte superior de la Figure 2. Se aprecia como el número de

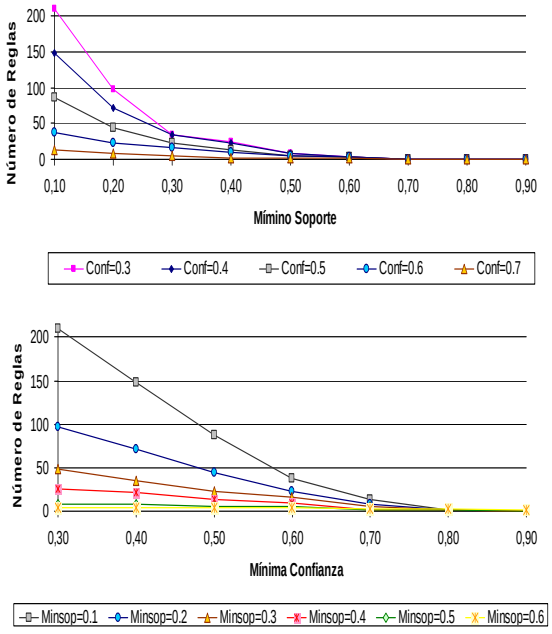


Figura 2: Relación entre reglas de asociación, α y λ.

reglas disminuye cuando el valor del mínimo soporte va aumentando. Además, se observa como las curvas obtenidas tienen forma similar y la separación entre ellas es más pequeña con valores superiores a 0,2 en el mínimo soporte. Sin embargo, se aprecia como la distancia entre las curvas es equidistante para distintas confianzas y un soporte fijo. Esto implica que el número de reglas, a partir de un soporte establecido, aumenta de una manera constante según va disminuyendo la confianza y por tanto, no se resalta la presencia de instancias especiales.

En la parte inferior de la Figure 2 se muestra la relación que existe entre el número de reglas de asociación difusas y varios valores de confianza con respecto a distintos valores del mínimo soporte. Se aprecia como el número de reglas aumenta según va disminuyendo el valor de la confianza. Resaltar como el valor de la confianza influye en el número de reglas cuando el mínimo soporte toma valores pequeños tales como 0,1 y 0,2.

6.2. Información extraida de "100mlI"

El análisis de las reglas de asociación difusas proporcionan altos niveles de conocimiento a partir del conjunto de datos "100mlI". El número de reglas obtenidas con un soporte de 0,2 y una confianza de 0,6 fue de 23, con respecto a las 38 obtenidas con un soporte de 0,1. Observamos como el valor predeterminado de la confianza implica la obtención de reglas de asociación relevantes debido al proceso de barrido de las posibles probabilidades de cada uno de los posibles valores que pueden tomar las variables (ver P10, Sección 5). Una muestra del nivel de interpretabilidad de la información obtenida es mostrada en la Tabla 1 obtenida a partir de varias reglas de asociación difusas relevantes. Estas reglas muestran como un atleta tiene características similares en las distintos parámetros que componen la prueba, es decir, un atleta tendrá una velocidad o carrera estimada como "Media" si su salto y su reacción también son "Medios". Además, gracias a la información obtenida se puede determinar que un

atleta no es relevante, para la prueba de 100ml si su salto es "Bajo", aunque tenga una velocidad considera "Media". Por el contrario, sí se considera relevante siempre que su reacción sea "Baja" y su carrera sea "Media". Esta información obtenida ha sido de gran utilidad para configurar el programa de entrenamiento de los deportistas de este equipo de atletismo.

7. Conclusiones

El objetivo principal de este trabajo es extraer reglas de asociación difusas a partir de DBC. Para ello, integramos conceptos de DBC con el algoritmo Apriori difuso propuesto por Hong y otros. Para alcanzar el objetivo, varios aspectos han sido considerados debido a la asusencia del valor real y preciso de los datos y a la interpretación posibilística aplicada. Esto ha afectado: a) al cálculo de la frecuencia de ocurrencias de los ítems, la cual es determinada a partir de las funciones de pertenencia que son interpretadas como conjuntos de probabilidades, b) al cálculo de la confianza de una regla la cual está contenida en un conjunto de probabilidades. El buen comportamiento y rendimiento de esta propuesta es mostrado a partir de un problema real con DBC, obteniendo como resultado información y patrones relevantes entre las dependencias y relaciones de los ítems.

Agradecimientos

Este trabajo está soportado por el Ministerio de Educación y Ciencia de España, TIN2008-06681-C06-{01 y 04} y TIN2011-28488 y por el proyecto andaluz TIC-2010-6858

Referencias

[1] R. Agrawal, T. Imielinski, A. Swami. Mining association rules between sets of items in large datasets. En: *SIGMOD*, Washington D.C. (USA), pp. 207-216, 1993.

[2] R. Agrawal, R. Srikant. Fast algorithms for mining association rules, En: *Int. Conf. on Very Large Data Bases*, pp. 487-499, 1994.

[3] C. Baudrit, D. Dubois, N. Perror. Representing parametric probabilistic models tainted with imprecision. *Fuzzy Sets and Systems*, vol. 15 (1), pp. 1913-1928.

[4] I. Couso, L. Sánchez. Higher order models for fuzzy random variables. *Fuzzy Sets and Systems*, vol. 159, pp. 237-258, 2008.

[5] M. Delgado, N. Marín, D. Sánchez, M.A. Vila. Fuzzy association rules: General model and applications. *IEEE Transactions on Fuzzy Systems*, vol. 11(2), pp. 214-225, 2003.

[6] D. Dubois, H. Prade, T. Sudamp. On the representation, measurement, and discovery of fuzzy associations. *IEEE Transactions on Fuzzy Systems*, vol. 13(2), pp. 250-262, 2005.

[7] J. Han, M. Kamber. *Data Mining: Concepts and Techniques.* Second Edition. M. Kaufmann, 2006.

[8] T.P. Hong, C.S. Kuo, S.C. Chi. Trade-off between time complexity and number of rules for fuzzy mining from quantitative data. *International Journal Uncertain Fuzziness Knowledge-Based Systems*, vol. 9(5), pp. 587-604, 2001.

[9] H. Ishibuchi, T. Nakashima, M. Nii. *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*, Springer-Verlag, 2004.

[10] A. Kaufmann, M. Gupta. *Introduction to Fuzzy Arithmetic: Theory and Applications.*, 1991

[11] G.J. Kulr, B. Yuan. Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[12] Y.C. Lee, T.P. Hong, W.Y. Lin. Mining fuzzy association rules with multiple minimum supports using maximum constraints. *LNCS*, vol. 3214, pp. 1283-1290, 2004.

[13] P. Limbourg. Multi-objetive optimization of problems with epistemic uncertainty, 2005.

[14] A. Palacios, L. Sánchez, I. Couso. Future performance modelling in athletism with low quality data-based gfss. *J. of Mult.-Valued Logic and Soft Computing*, vol. 17(2-3), pp. 207-228, 2011.

[15] A. Palacios, L. Sánchez, I. Couso. Linguistic cost-sensitive learning of genetic fuzzy classifiers for imprecise data. *IJAR*, vol. 5(6), pp. 841-862, 2011.

[16] L. Sánchez, I. Couso, J. Casillas. Genetic learning of fuzzy rules based on low quality data.*Fuzzy Sets and Systems*, vol. 160(17), pp. 2524-2552, 2009.

[17] S. Yue, E. Tsang, D. Yeung, D. Shi. Mining fuzzy association rules with weighted items. *IEEE International Conference on Systems, Man and Cybernetics*, pp. 1906-1911, 2000.

[18] C. Zhang, S. Zhang. *Association Rule Mining: Models and Algorithms.* Springer-Verlag, 2002.

Disimilitud esperada de una variable aleatoria difusa

Laura Garrido[1]

[1]*Dpto. de Estadística e I.O. y D.M, Universidad de Oviedo*, garridolaura@uniovi.es

Resumen

La validación de sistemas de aprendizaje de máquina para resolver problemas de clasificación supervisada, u otros problemas más generales de modelado involucra frecuentemente el uso de medidas de dispersión, como el error cuadrático medio o la varianza. Cuando los datos de aprendizaje son imprecisos, la función de *fitness* puede modelarse por medio de una variable aleatoria difusa. En estas circunstancias, es necesario disponer de medidas de dispersión adecuadas a este tipo de datos. En este trabajo se propone la construcción de medidas de dispersión de variables aleatorias difusas que se denominarán "medidas de disimilitud esperada". Como valor añadido respecto a otras medidas de dispersión escalares anteriormente introducidas en la literatura, permiten cuantificar la dispersión incluso en aquellos casos en que las imágenes de la variable aleatoria difusa no son subconjuntos difusos de la recta real. Esta innovación facilitará la validación de sistemas incluso en aquellos casos en que la función de fitness no admita una escala numérica, como por ejemplo, en algunos problemas de aprendizaje por refuerzo.

Palabras Clave: Medida de comparación, divergencia entre conjuntos difusos, disimilitud, variable aleatoria difusa.

1. Introducción

El concepto variable aleatoria difusa fue introducido por primera vez en 1976 por Féron [6], como una función que asigna un subconjunto difuso a cada posible resultado de un experimento aleatorio, extendiendo así los conceptos de variable aleatoria y de conjunto aleatorio. Con posterioridad se han propuesto diversas variantes del concepto. Estas variantes se diferencian, desde un punto de vista formal, en las condiciones de medibilidad impuestas a la aplicación y en las propiedades del espacio final, pero todas ellas tienen el objetivo de modelar situaciones que combinan aleatoriedad y borrosidad. Además de las diferencias entre las definiciones formales, también se le ha dado diversas interpretaciones al concepto variable aleatoria difusa. Puede ser interpretada ([3]) como un objeto aleatorio, como la representación de una variable aleatoria de la que tenemos información imprecisa o como una probabilidad superior condicionada ([4]). Cada una de esas interpretaciones da lugar a distintos modos de extender parámetros como la esperanza o la varianza ([3]). En este trabajo, trataremos las variables aleatorias difusas como objetos aleatorios. Introduciremos el concepto de disimilitud esperada de una variable aleatoria difusa, la cual promedia los grados de diferencia o divergencia entre las salidas de la variable. Con el objetivo de encontrar la medida más adecuada para cuantificar la diferencia entre dichas salidas, realizamos un breve estudio sobre algunas medidas de "desigualdad" existentes en la literatura, como las medidas de divergencia ([9]) o las medidas de distancia ([12, 13]), introduciendo el concepto de disimilitud. Probaremos cómo la disimilitud esperada generaliza diferentes parámetros según la medida concreta escogida, como la entropía, la varianza o una combinación de ambas. Además, nos permite cuantificar la diferencia esperada entre valores de variables aleatorias difusas no reales, dado que no es necesaria la existencia de la esperanza para calcularla. Observe el lector que, en las definiciones existentes en la literatura del concepto varianza escalar de una variable aleatoria difusa (ver, por ejemplo, [8]), se cuantifica la distancia cuadrática media con respecto a la esperanza de la variable aleatoria difusa.

El resto del trabajo está organizado del siguiente modo: en la Sección 2 realizamos un breve estudio sobre las diferentes medidas de comparación posibles, dando algunos resultados relacionando los conceptos de disimilaridad [2], divergencia [9], distancia [12, 13] y métrica, e introducimos el concepto de disimilitud. En la Sección 3, proponemos el concepto de disimilitud esperada de una variable aleatoria

difusa, estudiando algunas propiedades interesantes e ilustrándolas con ejemplos. Finalizamos el trabajo con algunas conclusiones.

2. Medidas de "diferencia" o "desigualdad" entre pares de conjuntos difusos

Como acabamos de mencionar en la Introducción, el primer paso en la construcción de una medida de disimilitud esperada será el estudio de las diferentes opciones disponibles para cuantificar la "diferencia" entre pares de valores de una variable aleatoria difusa.

A continuación, enumeramos una lista de axiomas que se utilizará para presentar diversas definiciones de "desigualdad" de una forma clara y compacta. Denotaremos por U el universo de referencia, y por $\mathscr{F}(U)$ la familia de los subconjuntos difusos de dicho universo.

En primer lugar, presentamos una serie de axiomas generales, que podrían ser exigidos tanto en medidas de "diferencia" como en medidas de "igualdad" (medidas que denotamos por m):

- G1.- $0 \leq m(A,B) \leq 1 \quad \forall A,B \in \mathscr{F}(U)$.
- G1*.- $0 \leq m(A,B) \leq 1 \quad \forall A,B \in \mathscr{F}(U)$ y existe algún $C,D \in \mathscr{F}(U)$ tal que $m(C,D) = 1$.
- G2.- $m(A,B) = m(B,A) \quad \forall A,B \in \mathscr{F}(U)$.
- G3.- Supongamos U finito. Sea $\rho : U \to U$ una permutación. Para cada $A \in \mathscr{F}(U)$ denotaremos por A^ρ el conjunto difuso cuya función de pertenencia se define a partir de μ_A como sigue: $\mu_{A^\rho}(x) = \mu_A(\rho(x))$. Entonces
$$m(A,B) = m(A^\rho, B^\rho).$$
- G3*.- U es finito y $m(A,B) = \sum_{x \in U} h[\mu_A(x), \mu_B(x)]$.
- G4.- Existe una aplicación $f : \mathscr{F}(U) \times \mathscr{F}(U) \times \mathscr{F}(U) \to \mathbb{R}$ tal que[1]
$$m(A,B) = f(A \cap B, A \setminus B, B \setminus A).$$
- G4*.- Existe una función $F_m : \mathbb{R}^3 \to \mathbb{R}$ y una medida difusa [2] $M : \mathscr{F}(U) \to \mathbb{R}$ tal que, para cada par $A,B \in \mathscr{F}(U)$, $m(A,B)$ puede ser expresada del siguiente modo:
$$m(A,B) = F_m(M(A \cap B), M(A \setminus B), M(B \setminus A)).$$

A continuación, enumeramos aquellos axiomas que pueden ser exigidos por una medida de "desigualdad", la cual denotamos por d:

- D1.- Si $A,B,C \in \mathscr{F}(U)$ satisfacen las restricciones $A \subseteq B \subseteq C$, entonces:
  (a) $d(A,C) \geq d(A,B)$ y
  (b) $d(A,C) \geq d(B,C)$.
- D2.- $d(D,D^c) = \text{máx}_{A,B \in \mathscr{F}(U)} d(A,B) \ \forall D \in \wp(U)$.
- D3.- $A = B \Longrightarrow d(A,B) = 0$.
- D3*.- $A = B \Longleftrightarrow d(A,B) = 0$.
- D4.- $\forall A,B,C \in \mathscr{F}(U)$, $d(A \cap C, B \cap C) \leq d(A,B)$.
- D5.- $\forall A,B,C \in \mathscr{F}(U)$, $d(A \cup C, B \cup C) \leq d(A,B)$.
- D6.- Si $A,B,C,D \in \mathscr{F}(U)$ satisfacen las siguientes restricciones: $A \cap B \subseteq C \cap D$, $A \setminus B \supseteq C \setminus D$, y $B \setminus A \supseteq D \setminus C$ entonces $d(A,B) \geq d(C,D)$.
- D7.- $d(A,B) = d(A \setminus B, B \setminus A)$.

Existen en la literatura diversas definiciones relacionadas con la noción de "diferencia". A continuación mencionaremos las definiciones axiómaticas más destacables. Dichas definiciones fueron dadas en contextos distintos, para distintos tipos de conjuntos (difusos, intervalo-valorados, intuicionistas, etc.) y con distintas notaciones. Para presentar todas las definiciones en un marco común, no utilizaremos su versión original, sino que nos restringiremos a los subconjuntos difusos y haremos referencia a los axiomas que acabamos de enumerar.

Antes de pasar a ver la primera de las definiciones relacionadas con la noción de "diferencia" entre conjuntos difusos, debemos conocer el concepto de *M-medida de comparación*, ya que la definición de nuestro interés es un caso particular de dicho tipo de medidas.

Definición 1. *([2] Bouchon-Meunier et al., 1996) Una* M-medida de comparación *sobre U es una función c sobre* $\mathscr{F}(U) \times \mathscr{F}(U)$ *satisfaciendo G1 y G4*.*

[1] Consideramos como operador diferencia, $\setminus$, una operación entre subconjuntos difusos verificando las propiedades: si $A \subseteq B$ entonces $A \setminus B = \emptyset$, $A \setminus B \subseteq A \setminus (A \cap B)$ y $A \setminus B$ es monótona con respecto a A.

[2] Pueden encontrarse diferentes definiciones de este concepto. Aquí, función de conjuntos monótona no decreciente satisfaciendo la restricción $M(\emptyset) = 0$.

El concepto de medida de comparación incluye tanto la idea de "igualdad" como de "desigualdad". En su artículo ([2]) los autores dan definiciones de ambos tipos. En este trabajo, es de nuestro interés únicamente el concepto de *disimilaridad* que vemos a continuación.

Definición 2. *([2], Bouchon-Meunier et al., 1996) Una M- medida de comparación d es una* M-medida de disimilaridad *si satisface el Axioma D3 y la función asociada $F_d : \mathbb{R}^3 \to \mathbb{R}$ no depende del primer argumento y es creciente en el segundo y tercer argumento.*

Montes et al. introdujeron en [9] una definición axiomática para la *divergencia* entre pares de conjuntos difusos basada en las siguientes propiedades: ha de ser no negativa y simétrica, ha de tomar el valor cero cuando los dos conjuntos comparados coinciden, y decrece a medida que los dos conjuntos son "más parecidos". Diferentes formalizaciones de esta tercera propiedad dan lugar a diferentes definiciones. A continuación tenemos las dadas por Montes et al.

Definición 3. *([9], S. Montes et al., 2002) Una función con valores reales d sobre $\mathscr{F}(U) \times \mathscr{F}(U)$ se denomina* medida de divergencia *si satisface las propiedades G2, D3, D4 y D5.*

Definición 4. *([9], S. Montes et al., 2002) Una función con valores reales d sobre $\mathscr{F}(U) \times \mathscr{F}(U)$ se denomina* medida de divergencia local *si satisface las propiedades de medida de divergencia (G2, D3, D4 y D5) y, además, la propiedad G3*.*

Las nociones de "divergencia" y "disimilaridad simétrica" son muy parecidas, como puede comprobarse en [7]. Toda disimilaridad simétrica, cuando se restringe a los conjuntos nítidos, satisface los axiomas de medida de divergencia. Inversamente, toda medida de divergencia que puede expresarse como una medida de comparación es una medida de disimilaridad. De hecho, las medidas de divergencia y disimilaridad se centran en las diferencias entre los conjuntos comparados, pero no en sus coincidencias. Podemos ilustrar este hecho mediante este sencillo ejemplo.

Ejemplo 1. *Consideremos como universo el siguiente conjunto de idiomas $U = \{$Chino (c), Español (e), Holandés (h), Francés (f), Italiano (i), Ruso (r)$\}$. Sean $E = \{h,e,f,i,c\}$, $G = \{h,e,f,r,c\}$, $A = \{i\}$ y $V = \{r\}$ los conjuntos nítidos que representan las habilidades comunicativas de cuatro personas llamadas Erik, Gerard, Angelo y Vladimir. Erik y Gerard comparten muchos más idiomas que Angelo y Vladimir, pero esas coincidencias no pueden detectarse a través de medidas de divergencia y disimilaridades. Si solamente queremos centrarnos en las diferencias, dichas medidas pueden ser útiles. Pero, si también queremos tener en cuenta sus habilidades comunes, debemos utilizar unas medidas de comparación diferentes.*

Podemos encontrar en la literatura otras medidas que tienen en cuenta las coincidencias entre los conjuntos comparados. A continuación vemos dos de ellas:

Definición 5. *([12] L. Xuecheng, 1992) Una función con valores reales d sobre $\mathscr{F}(U) \times \mathscr{F}(U)$ se denomina* medida de distancia *si satisface las propiedades G2, D1, D2 y D3.*

Definición 6. *([13], H. Zhang et al., 2009) Una función con valores reales d sobre $\mathscr{F}(U) \times \mathscr{F}(U)$ se denomina* medida de distancia *si satisface las propiedades G1*, G2, D1, D2 y D3*.*

Existen algunas relaciones entre las medidas de divergencia y las medidas de distancia, que pueden encontrarse en [7]. Si una medida d satisface la Definición 6, o bien la Definición 4, entonces satisface la Definición 5.

Hemos de remarcar que el término medida de distancia aparece en [12, 13], pero no se refiere al concepto matemático de métrica al que hacemos alusión a continuación. Podemos encontrar en la literatura reciente diversas métricas y pseudo-métricas definidas en la familia de los conjuntos difusos, como la distancia de Hamming, la pseudo-metrica de Puri-Ralescu [10] [3], así como otras familias de métricas propuestas en [1, 8, 11] definidas para familias específicas de conjuntos difusos convexos.

Aunque no entraremos en detalles en este trabajo, podemos encontrar algunas relaciones entre las mencionadas métricas y las nociones de divergencia, distancia y medida de comparación. Dichas relaciones se enumeran en el siguiente resultado.

Proposición 1. *Se verifica que:*

- *Todas las métricas y pseudo-métricas citadas pueden expresarse como medidas de comparación sobre sus respectivos dominios de definición. Además satisfacen los axiomas G2, D1, D3 y D4.*
- *Únicamente la distancia de Hamming satisface los axiomas D2 y D5. Además es la única que verifica que su función asociada al expresarla como medida de comparación no depende de su primer argumento y es creciente en los otros dos con respecto a la inclusión.*

De acuerdo con este resultado, exigir a una medida los axiomas D2 y D5, así como que al expresarla como medida de comparación no dependa de su primer argumento y sea creciente en los otros dos con respecto a la inclusión, excluye a la mayoría de las métricas y pseudo-métricas mencionadas. En el resto del trabajo, nos referiremos con el

[3] Puri y Ralescu introducen una métrica en la clase de los subconjuntos difusos de $\mathbb{R}^n$ cuyos α-cortes son compactos y no vacíos. Esta medida puede ser extendida facilmente a familias más generales de subconjuntos difusos como una pseudo-métrica.

término *medida de disimilitud* a aquellas medidas de comparación que satisfagan los axiomas G2, D1, D3 y D4.

3. Disimilitud esperada de una variable aleatoria difusa

Consideremos un espacio probabilístico $(\Omega, \mathscr{A}, P)$ y una variable aleatoria difusa definida sobre él, i.e., una aplicación $\mathscr{A}-\sigma$ medible $\tilde{X}:\Omega\to\mathscr{F}$, donde σ es una σ-álgebra definida en cierta clase de subconjuntos difusos $\mathscr{F}\subseteq\mathscr{F}(U)$. Toda variable aleatoria difusa induce una medida de probabilidad sobre σ dada por la siguiente fórmula:

$$P_{\tilde{X}}(\mathscr{C}) = P(\{\omega\in\Omega : \tilde{X}(\omega)\in\mathscr{C}\}),\ \forall\mathscr{C}\in\sigma.$$

Consideremos $P\otimes P:\mathscr{A}\otimes\mathscr{A}\to[0,1]$ la probabilidad producto, es decir, la única probabilidad satisfaciendo la siguiente restricción

$$(P\otimes P)(A\times B) = P(A)\cdot P(B)\ \forall A,B\in\mathscr{A}.$$

Sean $\tilde{X}_1$ y $\tilde{X}_2$ dos copias (idénticamente distribuidas) de $\tilde{X}$, y consideremos una medida de disimilitud definida sobre $\mathscr{F}$, $D:\mathscr{F}\times\mathscr{F}\to[0,1]$.

Definición 7. *Definimos la* disimilitud esperada *de $\tilde{X}$ como:*

$$E_D(\tilde{X}) = \int_{\Omega\times\Omega} D(\tilde{X}_1(\omega),\tilde{X}_2(\omega'))\,d(P\otimes P)(\omega,\omega')$$

Nótese que $E_D(\tilde{X})$ está bien definido dado que la aplicación $g(\omega,\omega') = D(\tilde{X}_1(\omega),\tilde{X}_2(\omega')),\ \forall(\omega,\omega')\in\Omega\times\Omega$ es $\mathscr{A}\otimes\mathscr{A}-\beta_{\mathbb{R}}$ medible.

Para justificar esta definición basta recordar las distintas formas de expresar la varianza de una variable aleatoria. La varianza de la variable con imágenes reales X puede obtenerse como la esperanza de la variable $(X-E(X))^2$, siendo necesario para este cálculo la existencia de $E(X)$. Pero otra forma de obtener la varianza es calculando la esperanza de la variable $(X_1-X_2)^2$, con X_1 y X_2 dos copias idénticamente distribuidas de X. Si en lugar de utilizar la diferencia consideramos una medida de disimilitud arbitraria, obtenemos la fórmula de disimilitud esperada dada en este trabajo. Notar que de este modo no es necesario el cálculo (ni siquiera la existencia) de la esperanza de la variable.

Podríamos utilizar, además de disimilitudes, otros tipos de medidas en la definición anterior. Según la medida escogida se generalizarían distintos conceptos existentes. Pero destaca el caso en el que la medida utilizada es una medida de disimilitud, y por ello ha sido el tipo de medida escogida en este trabajo. A continuación vemos los motivos.

Por un lado, la disimilitud esperada extiende el concepto de entropía de una variable aleatoria. Sea $\tilde{X}$ una variable aleatoria difusa representando a una variable aleatoria X definida en un universo finito $U=\{u_1,\ldots,u_n\}$ en el sentido de que $\tilde{X}(\omega)=\{X(\omega)\},\ \forall\omega\in\Omega$. Si consideramos como medida de disimilitud la distancia de Hamming, $D(A,B) = d_H(A,B) = \#A\triangle B$ entonces la disimilitud esperada de $\tilde{X}$ es 2 veces la entropía de X:

$$E_D(\tilde{X}) = \sum_{i=1}^{n}\sum_{j=1}^{n} d_H(\{u_i\},\{u_j\})p_i\cdot p_j = 2(1-\sum_{i=1}^{n} p_i^2),$$

donde p_i denota la probabilidad $P(X=u_i)$, $i=1,\ldots n$.

Por otro lado, la disimilitud esperada extiende algunos conceptos existentes de varianza escalar de una variable aleatoria difusa [3]: todas las (pseudo-)métricas consideradas en la Sección 2 satisfacen (como vimos) las propiedades de medida de disimilitud. Además, si consideramos una función no decreciente g satisfaciendo la condición $g(0)=0$, la composición de dicha función con una disimilitud es también una medida de disimilitud. Si construimos la siguiente medida $D=\frac{d^2}{2}$ donde d es cualquiera de las (pseudo-)métricas consideradas, y tenemos en cuenta la aritmética específica utilizada en cada contexto, para evitar el uso explicito de la esperanza, podemos extender los conceptos existentes de varianza escalar [3]. Aquí podemos ver la gran ventaja de la disimilitud esperada: ésta puede ser aplicada incluso cuando las imágenes de la variable aleatoria difusa no son convexas o no son subconjuntos de $\mathbb{R}^n$, ya que no es necesaria la existencia de la esperanza de la variable para obtener la disimilitud esperada.

Veamos a continuación algunos ejemplos del uso de la disimilitud esperada, aunque no entraremos en demasiados detalles debido a la brevedad exigida en este trabajo:

Ejemplo 2. *Consideremos una población finita Ω y el conjunto de idiomas $U=\{c, e, h, f, i, r\}$ del Ejemplo 1. Consideremos la aplicación multi-valuada $\mathscr{G}:\Omega\to\wp(U)$ que asigna a cada persona $\omega\in\Omega$ el subconjunto de idiomas de U que puede hablar* [4]*. Consideremos la distancia de Hamming $D(A,B)=d_H(A,B)=\#(A\triangle B)$. La disimilitud esperada $E_D(\mathscr{G})$ representa un grado de divergencia entre las destrezas comunicativas de las personas de la población.*

Ejemplo 3. *El ejemplo anterior puede modificarse si disponemos de información más refinada sobre las habilidades comunicativas de la población. Podríamos utilizar una variable aleatoria difusa $\tilde{X}:\Omega\to\mathscr{F}(U)$ para representar dichas habilidades. El valor de pertenencia $\tilde{X}(\omega)(u)$ representará un grado de preferencia ([5]) en una escala de 0 a 1 para el idioma $u\in U$. Entondes $\tilde{X}(\omega)(u)>\tilde{X}(\omega)(u')$ significará que la persona ω prefiere hablar u a u', porque está más familiarizado con ese idioma. Estos grados de*

[4] Supongamos que $\mathscr{G}(\omega)\neq\emptyset,\ \forall\omega\in\Omega$, es decir, suponemos que toda persona en la población es capaz de hablar al menos uno de los idiomas del conjunto U. Las aplicaciones multi-valuadas representan casos especiales de variables aleatorias difusas. Además, si consideramos la familia de todos los subconjuntos de la población como la σ-álgebra inicial, la aplicación es medible con respecto a cualquier σ-álgebra del espacio final.

preferencia pueden determinarse como una función de los niveles MCERL, por ejemplo. Para cierta medida de disimilitud específica, la disimilitud esperada de $\tilde{X}$ representa un grado esperado de diferencia respecto a las habilidades comunicativas entre un par de personas de la población.

En la siguiente proposición, se especifican algunas propiedades generales de la disimilitud esperada.

Proposición 2. *Sea $S : \mathscr{F} \times \mathscr{F} \to \mathbb{R}$ una medida de disimilitud. Entonces:*

- $E_S(A) = 0$, $\forall A \in \mathscr{F}$.
- $E_S(\tilde{X} \cup A) \leq E_S(\tilde{X})$, *para toda variable aleatoria difusa $\tilde{X}$ y todo subconjunto $A \in \mathscr{F}$.*
- *Si* $S(A,B) = \sum_{x \in U} g(A(x), B(x))$ *entonces* $E(\tilde{X} \cup \tilde{Y}) \leq E_S(\tilde{X}) + E_S(\tilde{Y})$, *para todo par de variables aleatorias difusas $\tilde{X}$ e $\tilde{Y}$.*

Ilustremos el significado de estas propiedades utilizando de nuevo el conjunto de idiomas de los ejemplos 1 y 2. La primer propiedad significa que la disimilitud esperada es nula cuando toda la población posee las mismas habilidades comunicativas. Para la segunda propiedad, supongamos que todas las personas de la población que no hablan español realizan un curso para aprender dicho idioma. Entonces, la disimilitud esperada debe decrecer. Finalmente, consideremos dos grupos separados de idiomas, y consideremos las habilidades comunicativas dentro de cada grupo. La variable $\tilde{X}$ denotará las habilidades dentro del primer grupo y la variable $\tilde{Y}$ las habilidades dentro del segundo grupo de idiomas. Entonces, la disimilitud esperada en el conjunto total de idiomas no puede ser estrictamente mayor que la suma de la disimilitud esperada dentro de cada grupo.

4. Conclusiones

Hemos comprobado que el concepto de disimilitud esperada generaliza varios parámetros conocidos asociados a variables aleatorias, conjuntos aleatorios y variables aleatorias difusas. En particular, generaliza el concepto de entropía de una variable aleatoria y algunas varianzas escalares para variables aleatorias difusas. Las definiciones de varianza escalar que podemos encontrar en la literatura [3, 8] se restringen a aquellas situaciones en las que los valores de las variable aleatoria difusa son subconjuntos convexos de $\mathbb{R}^n$. El nuevo concepto de disimilitud esperada, sin embargo, puede aplicarse en varias situaciones, incluso aunque los valores de la variable no sean reales. En trabajos futuros, estudiaremos nuevas propiedades de la disimilitud esperada para medidas de disimilitud concretas, tratando de establecer su conexión con las nociones de entropía y dispersión.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio Español de Educación y Ciencia bajo el proyecto MTM2010-17844.

Referencias

[1] C. Bertoluzza, N. Corral, A. Salas: On a new class of distances between fuzzy numbers. *Mathware & Soft Computing* 2, pp. 71-84, 1995.

[2] B. Bouchon-Meunier, M. Rifqi, S. Bothorel: Towards general measures of comparison of objects. *Fuzzy Sets and Systems* 84, pp. 143-153, 1996.

[3] I. Couso, D. Dubois: On the variability of the concept of variance of a fuzzy random variable. *IEEE Transactions on Fuzzy Systems* 17, pp. 1070 - 1080, 2009.

[4] A.P. Dempster: Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics* 38, pp. 325-339, 1967.

[5] D. Dubois, H Prade: The three semantics of fuzzy sets. *Fuzzy Sets and Systems* 90, pp. 141-150, 1997.

[6] R. Féron: Ensembles aléatoires flous. *C.R. Acad. Sci. Paris Ser. A* 282, pp.903-906, 1976.

[7] L. Garrido: Relación entre diferentes definiciones axiomáticas de medida de similaridad entre conjuntos difusos. Tesina de Licenciatura, 2011.

[8] R. Körner: On the variance of fuzzy random variables. *Fuzzy Sets and Systems* 92, pp. 83-93, 1997.

[9] S.Montes, I. Couso, P. Gil, C. Bertoluzza: Divergence measure between fuzzy sets. *International Journal of Approximate Reasoning* 30, pp. 91-105, 2002.

[10] M.L. Puri, D.A. Ralescu: Fuzzy random variables. *Journal of Mathematical Analysis and Applications* 114, pp. 409-422, 1986.

[11] W. Trutschnig, G. González-Rodríguez, A. Colubi, M.A. Gil: A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. *Information Sciences* 179, pp. 3964-3972, 2009.

[12] L. Xuecheng: Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems* 52, pp. 305-318, 1992.

[13] H. Zhang, W. Zhang, C. Mei: Entropy of interval-valued fuzzy sets based on distance and its relationship with similarity measure. *Knowledge-Based Systems 22*, pp. 449-454, 2009.

Análisis Espectral Singular para Datos de Baja Calidad

Luciano Sánchez[1], **Inés Couso**[2], **Luis Junco**[1], **Ana Palacios**[1]

[1]*Universidad de Oviedo. Departamento de Informática*, {luciano,lajunco,palaciosana}@uniovi.es
[2]*Universidad de Oviedo. Departamento de Estadística*, couso@uniovi.es

Resumen

En este estudio se describe cómo generalizar el Análisis Espectral Singular a series temporales compuestas por intervalos borrosos, bajo una interpretación posibilística de la función de pertenencia. La extensión aquí propuesta proporciona cotas, también borrosas, de la energía asociada a las oscilaciones irregulares o aperiódicas presentes en la serie. Estas cotas pueden compararse con las de un ruido coloreado de tipo AR(1), mediante un test estadístico para datos imprecisos. Si la señal original es reconstruida sólo mediante las componentes significativamente diferentes de las del mencionado ruido, se logra un filtrado eficiente de la señal aún en aquéllos casos en que la energía del ruido decrece con la frecuencia y la serie es percibida de forma imprecisa o incompleta.

Palabras Clave: Análisis Espectral Singular, Datos Borrosos, Datos de Baja Calidad.

1 Introducción

En este trabajo[1] se describirá cómo generalizar el Análisis Espectral Singular (o *Singular Spectral Analysis*, SSA) [7] a series temporales de datos de baja calidad, entendiendo como tales aquéllos que son conocidos de forma imprecisa y cuya incertidumbre no es modelable mediante una única distribución de probabilidad.

Entre los usos comunes del SSA está el tratamiento de series de datos meteorológicos y, en general, la extracción de conocimiento a partir de de otras señales con oscilaciones irregulares y/o aperiódicas. En general, se considera que esta técnica se adapta mejor a este tipo de señales que el análisis de Fourier, la transformada de Gabor o las transformadas wavelet [2]. La razón de este buen comportamiento está en que, a diferencia de las transformaciones mencionadas, SSA no se basa en una base predefinida de funciones, sino que genera esta base a partir de los datos de cada problema. SSA puede utilizarse para caracterizar señales casi caóticas, y también para separar la señal del ruido en los casos en que la energía del ruido decae con la frecuencia, circunstancia en la cual el análisis de Fourier es propenso a detectar erróneamente oscilaciones inexistentes a bajas frecuencias [1].

Adicionalmente, y como se detallará en las siguientes secciones, tiene interés extender la técnica SSA a problemas en los que la incertidumbre en los datos no sea puramente estocástica. En particular, se propondrá cómo acotar con intervalos borrosos la energía de cada una de las componentes obtenidas en la descomposición SSA, en los casos en que cada término de la serie esté descrito a su vez mediante un intervalo borroso. Este tratamiento permite adquirir conocimiento a partir de ciertos tipos de señales inciertas de gran interés práctico, como por ejemplo las producidas por algunos sistemas SCADA (Supervisory Control and Data Analysis), donde cada instancia es percibida mediante un triplete compuesto por el valor mínimo, el promedio y el máximo de la señal durante un período de tiempo.

Las generalizaciones de los diferentes tipos de análisis espectral a datos borrosos no han sido frecuentes, y se han ocupado sobre todo de las transformadas wavelet y de Fourier [3]. Hay trabajos previos relacionando la lógica borrosa y el Análisis Espectral Singular en problemas con datos nítidos [6] pero, a diferencia de éstos, la extensión propuesta en este estudio se aplica a datos imprecisos, fundamentándose en una interpretación epistémica de la función de pertenencia borrosa, identificando ésta con una distribución de posibilidad que, a su vez, coindice con la familia de probabilidades compatibles con el conjunto anidado de intervalos de confianza definido por sus α-cortes [4].

La estructura de este documento es la siguiente: en la sec-

[1]Este estudio se ha financiado con cargo al proyecto de investigación TIN2011-24302, del Ministerio de Ciencia e Innovación.

ción 2 se introduce SSA para datos precisos, y se compara con la transformada de Fourier sobre un problema de ejemplo. En la sección 3 se detalla la extensión de esta técnica a datos de baja calidad, y se ilustra el procedimiento con la ayuda de datos sintéticos. Este trabajo finaliza en la sección 4, con las conclusiones y el trabajo futuro que se pretende desarrollar en esta línea.

2 Análisis Espectral Singular para datos nítidos

Considere que la serie temporal

$$\{X(t) : t = 1,\ldots,N\} \tag{1}$$

es una secuencia de muestras de la variable de interés, cuya media es cero. Esta serie se envuelve en un espacio vectorial de dimensión M, usando copias desplazadas de los datos escalares,

$$\mathbf{X}(t) = (X(t), X(t+1),\ldots,X(t+M-1)) \tag{2}$$

de modo que

$$\{\mathbf{X}(t) : t = 1,\ldots,N-M+1\}. \tag{3}$$

SSA consiste en calcular las direcciones principales de la secuencia de vectores aumentados $\mathbf{X}(t)$ en el espacio de fase. En primer lugar, la matriz de covarianzas C_X se estima como [7]

$$C_X = \frac{1}{N-|i-j|}\sum_{t=1}^{N-|i-j|} X(t)X(t+|i-j|) \tag{4}$$

y los valores propios $\{(\lambda_k,\rho_k) : k = 1,\ldots,M\}$ de C_X se obtienen resolviendo la ecuación

$$C_X\rho_k = \lambda_k\rho_k. \tag{5}$$

Cada valor propio λ_k mide la varianza parcial en la dirección ρ_k. Como es bien conocido, esto es equivalente a formar una matriz E cuyas columnas son los vectores propios ρ_k y una matriz diagonal Λ, cuyos elementos son los λ_k en orden decreciente:

$$E'C_XE = \Lambda. \tag{6}$$

En la figura 1 se muestra una serie temporal que se usará, a modo de ejemplo, en este estudio. Dicha serie se construye por muestreo a 100KHz de una señal continua. De la misma forma que ocurre en muchos sistemas industriales de captura de datos, supondremos que el sistema de captura agrupa cada 100 muestras consecutivas en tres valores: promedio, mínimo y máximo, y que estos son los únicos valores transmitidos al observador. La tasa efectiva de muestreo de cada una de estas tres señales es, por tanto, de 1KHz.

La señal continua consta de una componente senoidal a 50Hz, a la que se añade entre los periodos 100 y 200 una segunda señal a 100Hz. Una tercera señal, con componentes a 100 y 150Hz, es inyectada también entre los períodos 400 y 500. A partir del periodo 500, hay impulsos de 10μs a 180Hz. Obsérvese que estos impulsos no se percibirán en los valores promediados, pero sí en los valores mínimo y máximo.

A la señal continua se le añade un fuerte ruido rojo de tipo AR(1), cuya energía decae con la frecuencia. Este ruido es mucho más común en la práctica que el ruido blanco, sin embargo es menos utilizado en estudios teóricos ya que al no tener un espectro de frecuencias plano es más difícil separarlo de la señal de interés.

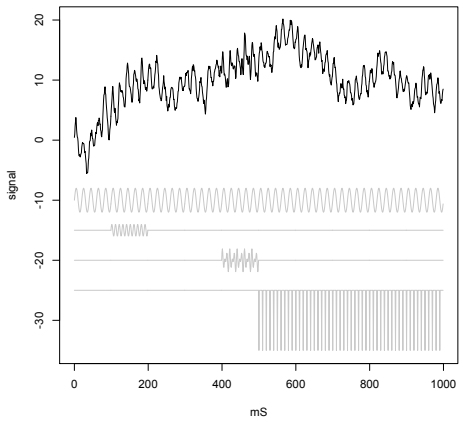


Figura 1: Serie temporal. Frecuencia de muestreo de 100KHz, promediada en bloques de 100 muestras. La señal tiene una componente de 50Hz, ruido rojo AR(1), segundo armónico entre las muestras 100 y 200 y tanto segundo como tercer armónicos entre los periodos 400 y 500. Se inyectan pulsos of 10μs a 180Hz entre las muestras 500 y 1000, que no son percibidos por el suavizado de la señal.

2.1 SSA para mejora adaptativa de la relación señal-ruido

Los vectores propios ρ_k de la matriz de covarianzas C_X se denominan también Funciones Ortogonales Empíricas (*Empirical Orthogonal Functions*, EOFs). Las parejas de EOFs en cuadratura son, aproximadamente, la contrapartida de un par seno-coseno en análisis de Fourier [7].

Las proyecciones de la serie sobre cada una de las EOFs se denominan Componentes Principales (*Principal Components*, PC):

$$A_k(t) = \sum_{j=1}^{M} X(t-j+1)\rho_k(j). \tag{7}$$

Para separar señal y ruido, se elige un grupo $\mathscr{E}$ de EOFs, y la parte de la serie que se corresponde con estas componentes se reconstruye combinando las PCs como sigue:

$$R_{\mathscr{E}}(t) = \frac{1}{M_t} \sum_{k \in \mathscr{E}} \sum_{j=L_t}^{U_t} A_k(t-j+1)\rho_k(j), \qquad (8)$$

donde los valores de M_t, L_t and U_t son

$$(M_t, L_t, U_t) = \begin{cases} (t,1,t) & 1 \le t \le M-1, \\ (M,1,M), & M \le t \le N-M+1, \\ (N-t+1, t-N+M, M), & N-M+1 \le t \le N. \end{cases} \qquad (9)$$

Retomando el ejemplo, en la figura 2 se observa que las primeras 5 EOFs de la función de ejemplo contienen la mayor parte de la energía de la señal, por lo que podría pensarse que $\mathscr{E} = \{1,\ldots,5\}$. Sin embargo, esta elección no sería sustancialmente diferente de la producida por un filtrado de las frecuencias altas mediante análisis de Fourier. En la siguiente sección se mostrará cómo incluir un modelo del ruido en el filtrado.

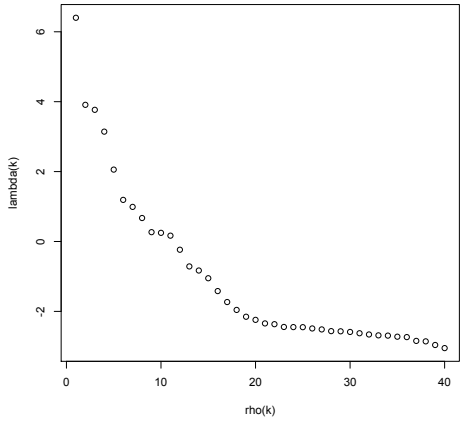


Figura 2: Espectro SSA: logaritmos de los valores propios, ordenados de mayor a menor

2.2 Comparación con el espectro del ruido para detectar oscilaciones ocultas

El espectro SSA es más legible si se asocia cada EOF con una frecuencia dominante (ver figura 3), obtenida por medio de una transformada de Fourier reducida, es decir el máximo con respecto de f de

$$e(\rho_k, f) = \left(\left(\sum_{j=1}^{M} \rho_{kj} \cos(2\pi f j) \right)^2 + \left(\sum_{j=1}^{M} \rho_{kj} \sin(2\pi f j) \right)^2 \right)^{\frac{1}{2}} \qquad (10)$$

Una pareja de valores propios aproximadamente iguales, asociados a EOFs con la misma frecuencia dominante, señalan la presencia de una oscilación periódica, como se muestra con los círculos rojos de la figura 3. Existen ciertas similaridades entre esta representación y la densidad espectral de potencia (o transformada de Fourier de la autocorrelación), mostrada en la figura 4.

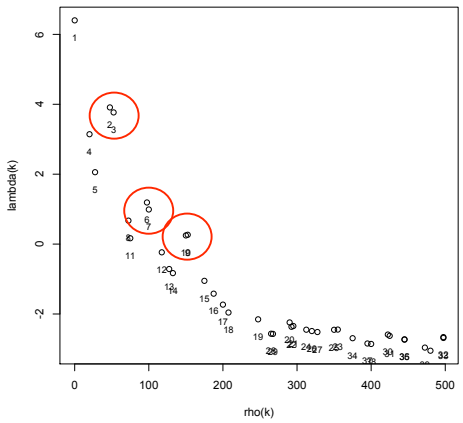


Figura 3: Espectro SSA: valores propios frente a frecuencias dominantes de las EOFs

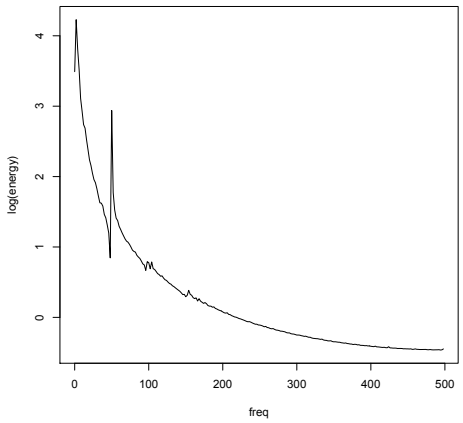


Figura 4: Densidad espectral de potencia

Por otra parte, SSA es capaz de detectar correctamente la presencia del primer armónico (ρ_2 and ρ_3), el segundo (ρ_6 and ρ_7) y el tercero (ρ_9 and ρ_{10}). Estos dos últimos grupos son más sencillos de identificar que en la figura 4. Además, la componente con la máxima energía (ρ_1) está aislada, por lo que se deduce que es una tendencia y no una oscilación; esta información no puede obtenerse de la figura 4.

Para determinar si una componente del grafo SSA es relevante en la reconstrucción de la señal, se representan en el mismo diagrama los intervalos de confianza de la energía del ruido, obtenidos con simulación Montecarlo. Por ejemplo, para determinar si la serie temporal es diferente

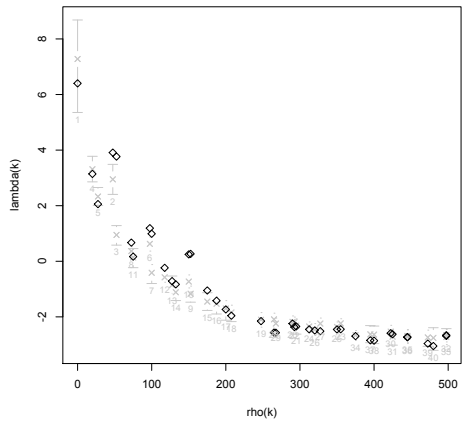


Figura 5: SSA e intervalos de confianza generados mediante simulación Monte-Carlo para un ruido rojo de tipo AR(1)

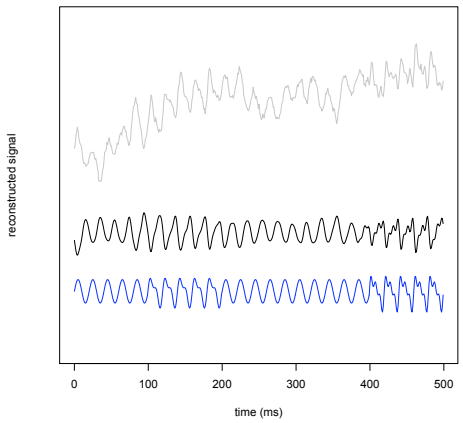


Figura 6: Parte Superior: serie temporal promedio. Centro: Filtrado de la serie mediante Montecarlo-SSA. Parte Inferior: señal teórica, para contrastar la calidad de la reconstrucción.

de ruido rojo (un proceso AR(1)),

$$X(t) = a_1[X(t-1) - X_0] + \sigma\xi(t) + X_0, \qquad (11)$$

donde $\xi(t)$ es ruido normal de media cero y varianza 1, se estiman en primer lugar los parámetros $\hat{X}_0$, $\hat{\sigma}$ y $\hat{a}_1$ del ruido a partir de $X(t)$, usando un criterio de máxima verosimilitud. Con estas estimaciones se genera un gran número de series artificiales o subrogadas, compatibles con el modelo de $\xi(t)$. A cada una de estas series subrogadas se le calcula su matriz de covarianzas C_R, que es proyectada en la base E correspondiente a los datos originales, obteniéndose un conjunto de matrices

$$\Lambda_R = E' C_R E. \qquad (12)$$

Las matrices Λ_R no serán diagonales, pero la distribución estadística de sus diagonales sirve para producir intervalos de confianza para los valores de λ_k (ver figura 5) de forma que pueda afirmarse que las componentes que están fuera de estos intervalos son significativamente diferentes del ruido, y por tanto candidatas para ser empleadas en la reconstrucción. En la figura 6 se muestra el resultado de reconstruir la señal continua a partir de la serie temporal de valores promedio con estas premisas. En esta misma figura se muestra que, dado que las EOFs están localizadas en el tiempo, la señal reconstruida contiene información acerca del instante en que ocurrió cada perturbación.

Obsérvese también que no hay traza de los impulsos mostrados en la figura 1, ya que hasta este momento no se ha incluido información acerca del máximo y del mínimo de la señal. La generalización de SSA a este caso se introduce en la siguiente sección.

3 SSA para valores borrosos

Suponga que la serie $X(t)$ se percibe mediante una secuencia de intervalos borrosos normales

$$\{\widetilde{X}(t) : t = 1, \ldots, N\}, \qquad (13)$$

con cortes $[\widetilde{X}(t)]_\alpha = [X^l_\alpha(t), X^u_\alpha(t)]$. Suponga también que los 1-cortes de $\widetilde{X}(t)$ cumplen que $X^l_1(t) = X^u_1(t) = X_c(t)$.

Siguiendo la interpretación de la función de pertenencia mencionada en la introducción, se supondrá que cada $[X^l_\alpha(t), X^u_\alpha(t)]$ cubre al valor desconocido $X(t)$ con probabilidad mayor o igual que $1-\alpha$.

3.1 Cálculo del espectro SSA a partir de datos borrosos

Introduzcamos en primer lugar el concepto de *selección* de una serie temporal borrosa. Sea $\{S(t) : t = 1, \ldots, N\}$ una serie temporal arbitraria, y sea

$$\mu(\{S(t)\}) = \min_{t=1}^{N} \widetilde{X}(t)(S(t)) \qquad (14)$$

el grado de compatibilidad entre $\{S(t)\}$ y $\{\widetilde{X}(t)\}$. Diremos que $S = \{S(t)\}$ es una selección de $\{\widetilde{X}(t)\}$ cuando $\mu(S) > 0$ o, en otras palabras, cuando existe un valor α_0 para el que $X^l_{\alpha_0}(t) \leq S(t) \leq X^u_{\alpha_0}(t)$ para todo t. Obsérvese que $\mu(\{X_c\}) = 1$. C_S es la matriz de covarianza muestral de la serie S (ver ec. 4) y $E_S C_S E'_S$ es una matriz diagonal.

Al igual que se ha hecho en el caso nítido, la serie temporal borrosa se incluye en un espacio vectorial de dimensión M,

$$\widetilde{\mathbf{X}}(t) = (\widetilde{X}(t), \widetilde{X}(t+1), \ldots, \widetilde{X}(t+M-1)) \qquad (15)$$

para formar una serie borrosa multivariante

$$\{\widetilde{\mathbf{X}}(t) : t = 1,\ldots,N-M+1\}. \quad (16)$$

La la matriz de covarianzas de la serie aumentada no es conocida, salvo por pertenecer al conjunto de las covarianzas de las selecciones de $\{\widetilde{\mathbf{X}}(t)\}$,

$$C_{\widetilde{X}}(C) = \quad \text{máx}\{\mu(S) : C = C_S\}. \quad (17)$$

Alternativamente, puede decirse que $C_{\widetilde{X}}$ es una "matriz de covarianzas borrosa", entendida como una subconjunto borroso del espacio de las matrices reales de dimensión M.

3.1.1 Definición basada en aritmética borrosa

Para aproximar $C_{\widetilde{X}}$ se propone extender la definición de Vautard and Ghil's (eq. 4) mediante aritmética borrosa, y definir la covarianza borrosa muestral como

$$C^{(1)}_{\widetilde{X}} = \frac{1}{N-|i-j|} \bigoplus_{t=1}^{N-|i-j|} \widetilde{\mathbf{X}}(t) \odot \widetilde{\mathbf{X}}(t+|i-j|) \quad (18)$$

donde

$$(\widetilde{A} \oplus \widetilde{B})(x) = \text{máx}\{\text{mín}(\widetilde{A}(a), \widetilde{B}(b)) : a+b = x\} \quad (19)$$

$$(\widetilde{A} \odot \widetilde{B})(x) = \text{máx}\{\text{mín}(\widetilde{A}(a), \widetilde{B}(b)) : ab = x\}. \quad (20)$$

Suponiendo ahora que las EOFs de la serie temporal son aproximadamente iguales a las de X_c (esto es, que $E'_{X_c} C_X E_{X_c} = \Lambda_X$ con Λ_X diagonal) entonces puede hacerse

$$\widetilde{\Lambda} \approx E'_{X_c} C^{(1)}_{\widetilde{X}} E_{X_c} \quad (21)$$

y usar los elementos de la diagonal de $\widetilde{\Lambda}$ como aproximación borrosa de la transformada SSA de la serie $X(t)$.

Desafortunadamente, las cotas obtenidas así son excesivamente conservadoras, y esta definición sólo es útil en el caso de que los conjuntos $\widetilde{X}(t)$ sean muy específicos. En caso contrario, se debe recurrir al segundo procedimiento que se propone a continuación.

3.1.2 Definición basada en la transformada de Karhunen-Loewe

En este segundo método, en primer lugar se preprocesan (rotación y escalado) los vectores de la serie aumentada, para que su matriz de covarianzas sea aproximadamente la unidad, lo cual permite estimar de forma más eficiente la matriz de covarianzas en este espacio transformado (despreciando los términos fuera de la diagonal) y por último se aplica la transformación inversa a la covarianza muestral para obtener los coeficientes borrosos de la transformada SSA. Se remarca que en este segundo procedimiento no se supondrá que las EOFs de $X_c(t)$ coinciden con las de $X(t)$. Las EOFs más adecuadas se determinarán mediante un algoritmo numérico de optimización, de forma que éstas maximicen la especificidad de la matriz de covarianzas muestral.

Sea $\mathbf{S}$ la extensión vectorial de una selección S,

$$\mathbf{S}(t) = (S(t),\ldots,S(t+M-1)), \quad (22)$$

y sea

$$\mathbf{Z}_S(t) = E_S \Lambda_S^{-\frac{1}{2}} \mathbf{S}(t), \quad (23)$$

la transformada de Karhunen-Loewe (KH) de $\mathbf{S}$, de modo que la covarianza de $\mathbf{Z}_S$ es la matriz identidad. Como primera hipótesis, se supondrá que existe una selección $\mathbf{S}^*$ suficientemente parecida a la serie $\mathbf{X}$ como para que el resultado de girar y escalar la serie multivariada borrosa $\widetilde{\mathbf{X}}(t)$ mediante las matrices E_{S^*} y Λ_{S^*} produzca una matriz de covarianzas

$$\widetilde{\mathbf{Z}}_{S^*}(t) = E_{S^*} \Lambda_{S^*}^{-\frac{1}{2}} \widetilde{\mathbf{X}}(t) \quad (24)$$

aproximadamente diagonal. El producto entre una matriz nítida y un vector borroso se define de la forma siguiente:

$$[E_S \Lambda_S^{-\frac{1}{2}} \widetilde{\mathbf{X}}(t)](V) = \text{máx}\{\widetilde{\mathbf{X}}(t)(\mathbf{X}) : V = E_S \Lambda_S^{-\frac{1}{2}} \mathbf{X}\}. \quad (25)$$

Si la imprecisión de $\widetilde{\mathbf{X}}$ no es excesivamente alta, $C_{\widetilde{\mathbf{Z}}_S}$ (ver eq. 17) es aproximadamente la matriz identidad para cualquier S, pero su especificidad dependerá de la matriz E_S elegida. Como segunda hipótesis, se admitirá que la selección S^* más adecuada es aquella para la que el determinante de su matriz asociada $C_{\widetilde{\mathbf{Z}}_{S^*}}$ sea el valor más específico, entendido como el más cercano al valor nítido 1.

Dado que $C_{\widetilde{\mathbf{Z}}_{S^*}}$ es aproximadamente la matriz unidad, se supondrá que sus términos fuera de su diagonal son cero, y que el término k-ésimo de su diagonal es

$$(C^{(2)}_{\widetilde{Z}})_{kk} = \frac{1}{N-|i-j|} \bigoplus_{t=1}^{N-|i-j|} (\widetilde{\mathbf{Z}}(t))_k \odot (\widetilde{\mathbf{Z}}(t+|i-j|))_k. \quad (26)$$

El algoritmo numérico de optimización empleado para buscar la selección $\mathbf{S}^*$ en que se alcanza la mínima especificidad del determinante de $C_{\widetilde{\mathbf{Z}}_{S^*}}$ está basado en algoritmos genéticos, y no se incluye en esta comunicación por razones de espacio.

3.2 Detección de oscilaciones en series borrosas

De igual manera que se ha hecho con los datos nítidos (ver figura 5), pueden detectarse oscilaciones en series borrosas comparando su espectro SSA con el característico de diferentes clases de ruido. Dado que en nuestra definición los valores propios de la transformación son borrosos pero las EOFs siguen siendo nítidas, se propone que cada punto del diagrama 4 se convierta en un segmento, el soporte del valor propio correspondiente.

La primera consecuencia de esta generalización está en que existe una cierta ambigüedad en la detección de oscilaciones, ya que la frase "amplitudes similares a la misma frecuencia fundamental" ha de extenderse al caso de que las amplitudes sean borrosas, mediante el uso de operadores de similaridad entre conjuntos borrosos.

Lo mismo ocurre cuando el espectro de los datos se compara con la distribución de los datos subrogados. En el caso nítido, es común definir intervalos de confianza al 95 % para los valores propios de las series subrogadas a las frecuencias dominantes de los EOFs. La hipótesis nula ("la serie es ruido") se rechaza si cualquiera de los valores propios está fuera del intervalo de confianza correspondiente [2]. En el caso contemplado en este estudio, se usa un test estadístico bootstrap para datos imprecisos, teniendo en cuenta la interpretación epistémica de las pertenencias [5].

3.3 Ejemplo ilustrativo

Se han aplicado ambas definiciones para determinar el espectro SSA de una serie temporal borrosa, construida a partir de la señal introducida en la sección anterior. El soporte de cada $\widetilde{X}(t)$ es el intervalo entre el mínimo y el máximo de las 100 muestras a 100KHz correspondientes, y la moda es el promedio de esas muestras. La forma de la función de pertenencia es triangular. Los resultados están en la figura 7. La incertidumbre de los valores propios es notablemente mejor en el segundo enfoque, y en este espectro se aprecia que hay información para detectar las oscilaciones (círculos rojos) mientras que al mismo tiempo el pulso de amplitud $20\mu S$ a 180 Hz es claramente visible (flecha roja).

4 Conclusiones y trabajo futuro

Cuando la tasa de muestreo efectivo de una señal es baja, el contenido en frecuencia de la señal promediada está limitado por la frecuencia de Nyquist de ese muestreo. No obstante, si el promedio se complementa con las señales mínimo y máximo, hay una información parcial acerca del contenido en frecuencias por encima de este umbral, que puede usarse para detectar oscilaciones ocultas para todos los métodos clásicos.

En este estudio se ha propuesto una extensión del análisis espectral singular para datos borrosos que, como caso particular, permite solucionar el problema anterior y que, en general, proporciona una información en frecuencia que explota de forma más eficiente datos que, por limitaciones en el sistema de captura, tienen una baja calidad.

Como trabajo futuro, se avanzará en las técnicas de filtrado y reconstrucción de la señal que permitan obtener una aproximación más específica a la señal de interés, combinando las EOFs nítidas con los valores borrosos de las componentes de interés, definiendo una nueva categoría de algoritmos de filtrado para datos con incertidumbre no estocástica.

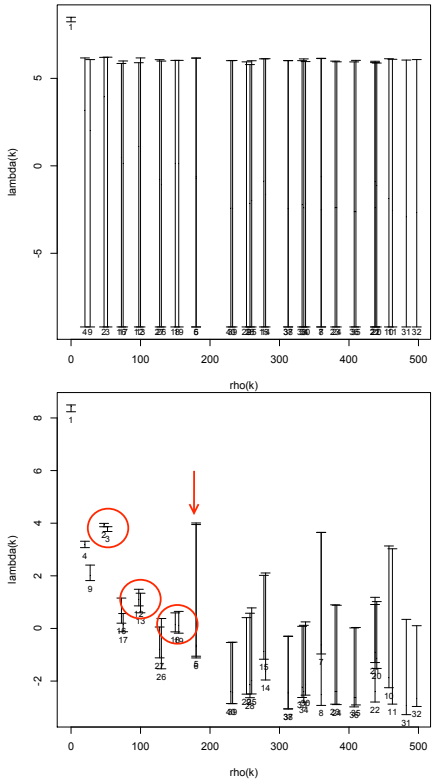


Figura 7: Fuzzy-SSA usando aritmética borrosa (parte superior) and el procedimiento basado en la transformada KH (inferior)

Referencias

[1] Allen, M. R., Robertson, A. W., Distinguishing modulated oscillations from coloured noise in multivariate datasets. Climate Dynamics 12, 775-784, 1996.

[2] Allen, M. R., and L. A. Smith, Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise, Journal of Climatics, 9, 3373-3404, 1996.

[3] Butkiewicz, B.S., Fuzzy approach to Fourier transform. Proceedings of SPIE, 6159-II, Article number 615947. 2006.

[4] Couso, I., Sánchez, L. Higher order models for fuzzy random variables. Fuzzy Sets and Systems 159, 237-258. 2008.

[5] Couso, I. Sánchez, L. Mark-recapture techniques in statistical tests for imprecise data. International Journal of Approximate Reasoning. doi 10.1016/j.ijar.2010.07.009. 2010.

[6] Fang, L.Q, Mi, D:, Xu, X.S., Chen, D. G. Fault diagnosis of fuel injection system of engine based on the singular spectrum analysis. Journal of Combustion Science and Technology 9 (2), 108-111. 1993.

[7] Ghil et al., Advanced spectral methods for climatic time series. Reviews of Geophysics, 40(1), 1-41. 2002.

[8] Ho, D. W. C., Zhang, P.m Xu, J. Fuzzy Wavelet Networks for Function Learning. IEEE Transactions on Fuzzy Systems, 9(1), 200-211. 2001.

INTRODUCING A GENETIC FUZZY LINGUISTIC COMBINATION METHOD FOR BAGGING FUZZY RULE-BASED MULTICLASSIFICATION SYSTEMS

Krzysztof Trawiński[1], **Oscar Cordón**[2], **Arnaud Quirin**[3], **Luciano Sánchez**[4]

[1,2,3]*European Centre for Soft Computing, 33600 - Mieres (Asturias), SPAIN*,
{krzysztof.trawinski,oscar.cordon,arnaud.quirin}softcomputing.es
[2]*the Dept. of Computer Science and Artificial Intelligence (DECSAI) and the Research Center on Information and Communication Technologies (CITIC-UGR), University of Granada, 18071 Granada, Spain*, ocordon@decsai.ugr.es
[4]*Computer Science Department, University of Oviedo, 33203 - Gijón (Asturias), SPAIN*, luciano@uniovi.es

Abstract

By using a fuzzy linguistic rule-based classification system as a combination method, classifier ensembles show a hierarchical structure. This approach allows fuzzy classifiers to deal with high dimensional problems avoiding the curse of dimensionality in fuzzy multiclassification systems. Furthermore, this combination seamlessly extends to vague datasets, being a potential alternative towards producing ensembles of rule-based classifiers from imprecisely perceived data. The present contribution establishes the first basis in this direction by introducing a genetic fuzzy system-based framework to build the fuzzy linguistic combination method for a bagging fuzzy multiclassification system and by proposing a theoretical extension of this procedure to fuzzy data.

Keywords: Fuzzy rule-based multiclassification systems, classifier fusion, Linguistic classifier selection and fusion, genetic fuzzy systems

1 INTRODUCTION

Multiclassification systems (MCSs), also called classifier ensembles, are machine learning tools capable to obtain better performance than a single classifier when dealing with complex classification problems, especially when the number of dimensions or the size of the data are really large [21]. The most common base classifiers are decision trees [17], neural networks [24], and more recently fuzzy classifiers [8, 9, 10].

MCS design is mainly based on two stages [11]: the learning of the component classifiers and the combination mechanism for the individual decisions provided by them into the global MCS output. The overall accuracy of the MCS relies on the performance and the proper integration of these two tasks.

The research area of combination methods is very active. It considers both the direct combination of the results provided by all the initial set of component classifiers to compute the final output (*fusion-based methods*) and the selection of the best single classifier or classifier subset which will be taken into account to provide a decision for each specific input pattern. Besides, hybrid strategies between the two groups have also been introduced [21].

While the weighted majority vote could be considered as the most extended fusion-based combination method [22], many other proposals have been developed in the specialized literature, including several successful procedures based on the use of fuzzy set theory and, specifically, of fuzzy aggregation operators [20]. However, up to our knowledge, there has not been any previous proposal of a MCS combination method considering the use of a fuzzy linguistic system to accomplish this task, which can constitute a very interesting alternative.

In this contribution[1] we introduce a framework to derive a fuzzy rule-based classification system (FRBCS) playing the role of the MCS combination method. This fuzzy linguistic combination method presents an interpretable structure as it is based on the use of a single disjunctive fuzzy classification rule per problem class as well as on the classical single-winner fuzzy reasoning method [6, 18]. The antecedent variables correspond to the component FRBCSs (and thus its number is bounded by the existing number) and each of them has a weight associated representing the certainty degree of each ensemble member in the classification of each class. Furthermore, we will show that this combination seamlessly extends to vague datasets, being an interesting candidate for deriving ensembles of rule-based classifiers from data with missing, censored or otherwise imprecisely known values.

[1]This work has been supported by the Spanish Ministerio de Ciencia e Innovación under projects TIN2009-07727 and TIN2011-24302, both including EDRF fundings.

A specific genetic algorithm (GA) to design such FRBCS-based combination method (FRBCS-CM) will be proposed with the ability of selecting features and linguistic terms in the antecedent parts of the rules. In such way, it will perform both classifier fusion and classifier selection at class level. The resulting system is thus a genetic fuzzy system (GFS) [7] (in particular, a genetic fuzzy rule-based classification system (GFRBCS)) dealing with the interpretability-accuracy trade-off in a proper way [4].

In particular, in the current work the novel FRBCS-CM will be applied on fuzzy rule-based multiclassification systems (FRBMCSs) generated from the bagging and feature selection methodology we proposed in [8, 9, 10]. Therefore, the resulting FRBMCS will show a clear hierarchical structure composed of two levels of FRBCSs allowing it to deal with high dimensional problems. A preliminary study will be conducted on 9 datasets from the UCI machine learning repository to test the accuracy and complexity of the derived FRBMCSs in comparison to the original FRBMCS.

This paper is set up as follows. In the next section, the preliminaries required for a good understanding of our work (MCS combination methods, fuzzy MCS combination methods, and our approach for designing FRBMCSs considering bagging and feature selection) are reviewed. Section 3 describes the proposed FRBCS-CM framework, structure, and the GA considered to design it. The experiments developed and their analysis are shown in Sec. 4. Finally, Sec. 5 collects some concluding remarks and future research lines.

2 PRELIMINARIES

2.1 Multiclassification System Combination Methods

Two main approaches arise in the literature for the combination of the outputs provided by a previously generated set of individual classifiers into a single MCS output [27]: *classifier fusion* and *classifier selection*.

Classifier fusion relies on the assumption that all ensemble members make independent errors. Thus, combining the decisions of the ensemble members may lead to increasing the overall performance of the system. Majority voting, sum, product, maximum and minimum are examples of functions used to combine their decisions [19]. The most extended one is the weighted majority voting, which allows to weight the contribution of each individual classifier to the final decision according to its "classification competence" using coefficients of importance [22].

Alternatively, classifier selection is based on the fact that not all the individual classifiers but only a subset of them will influence on the final decision for each input pattern. Two categories of classifier selection techniques exist: static and dynamic [15, 27]. In the first case, regions of competence are defined during the training phase, while in the second case, they are defined during the classification phase taking into account the characteristics of the sample to be classified. There is also another family of static classifier selection methods, called overproduce-and-choose strategy (OCS) [25], wich is based on the fact that a large set of candidate classifiers is generated and then selected to extract the best performing subset (removing duplicates and poor-performing candidate classifiers), which composes the final MCS used to classify the whole test set.

The FRBCS-CM proposed in the current contribution will belong to the static OCS group and it will be able to either completely remove a whole candidate classifier or to reduce its contribution to only some specific classes with a specific weight measuring our confidence in the individual classifier for that specific class (as done in [14]). All the latter will be performed using a human-interpretable structure generated by means of a GFRBCS.

2.2 Multiclassification System Fuzzy Combination Methods

Fuzzy set theory has been extensively and successfuly considered for classifier fusion. The use of fuzzy connectives to combine the outputs of the component classifiers of an ensemble was first proposed in [5]. Since then, many different fuzzy aggregation operators have been considered in the specialized literature [20]. In [20] the accuracy of some of them was compared to that of seven of the usual crisp (i.e., non-fuzzy) aggregation operators when considered as combination operators for Boosting classifier ensembles. The conclusions drawn from that experimentation were that fuzzy combination methods clearly outperformed non-fuzzy ones. Besides, some other works have extended the classifier fusion scope and have proposed some techniques which show some similarities with our new proposal, such as [3].

In previous studies [9, 10], we described a methodology based on classical MCS design approaches such as bagging [2], random subspace [17], and mutual information-based feature selection [1] to generate FRBMCSs. The approach uses the basic heuristic fuzzy classification rule generation method proposed by Ishibuchi [18] as well as a GA-based classifier selection technique driven by a multicriteria fitness function either composed of only an error measure [8] or of its combination with a diversity measure [26]. For a more detailed description on the methodology, the interested reader is referred to the provided references.

3 GFRBCS - A FUZZY LINGUISTIC COMBINATION METHOD

3.1 Fuzzy linguistic combination

The FRBCSs considered in the ensemble will be based on fuzzy classification rules with a class and a certainty de-

gree in the consequent. Let R_j^k be the j-th rule of the k-th member of an ensemble of K components,

$$\text{if } x \text{ is } A_j^k \text{ then Class } C_j^k \text{ with } CF_j^k,$$

where $C_j^k \in \{1,\ldots,n_c\}$ and n_c is the number of classes.

We will use the expression $\mathscr{G}^k = \{R_1^k,\ldots,R_{N_k}^k\}$ to denote the list of fuzzy rules comprising the k-th ensemble member. Let us partition each one of these lists into so many sublists $\mathscr{G}_c^k$ as classes. $\mathscr{G}_c^k$ contains the rules of $\mathscr{G}^k$ whose consequent is the class c.

Let us also define $R^k(x)$ to be the intermediate output of the k-th member of the ensemble, which is the fuzzy subset of the set of class labels computed as follows:

$$R^k(x)(c) = \bigvee_{\{j|C_j^k=c\}} CF_j^k \cdot A_j^k(x). \quad (1)$$

Each component FRBCS maps an input value x to so many degrees of membership as the number of classes in the problem. The highest of these memberships determines the classification of the pattern. That is to say, the k-th FRBCS classifies an object x as being of class $\text{FRBCS}^k(x) = \arg\max_{c\in\{1,\ldots,n_c\}} R^k(x)(c)$. Observe also that $R^k(x)(c)$ is the result of applying the fuzzy reasoning mechanism to the knowledge base defined by the sublist $\mathscr{G}_c^k$.

The simplest linguistic combination of the component FRBCSs consists in stacking a selection of some of the rules R_j^k into a single large rule base. Let us define a binary matrix $[b_{ck}] \in \{0,1\}^{n_c\times K}$, and let us agree that, if b_{ck} is zero, then $\mathscr{G}_c^k$ is removed from the ensemble and $R^k(x)(c) = 0$. This selection is equivalent to the hierarchical FRBCS comprising n_c expressions of the form:

$$\text{if (member}_1 \text{ says that class is c) or } \ldots \text{ or (member}_K \text{ says that class is c) then class is c,}$$

where the asserts "(member$_k$ says that class is c)" have a degree of certainty b_{ck} determined by the rules in the sublist $\mathscr{G}_c^k$, and those asserts for which b_{ck} is zero are omitted. The fuzzy output of this selected ensemble is

$$R^I(x)(c) = \bigvee_{\{(j,k)|C_j^k=c\}} b_{ck} \cdot CF_j^k \cdot A_j^k(x). \quad (2)$$

We can define more powerful linguistic selections which extend this basic fuzzy reasoning schema. In this paper we will use a *sparse* matrix of weights $[w_{ck}] \in [0,1]^{n_c\times K}$ and operate as follows:

$$R^{II}(x)(c) = \bigvee_{\{(j,k)|C_j^k=c\}} w_{ck} \cdot CF_j^k \cdot A_j^k(x). \quad (3)$$

Thus, the selected ensemble can be seen as a hierarchical knowledge base with n_c fuzzy classification rules with weights in the antecedent part

$$\text{if (member}_1(w_{c1}) \text{ says that class is c) or } \ldots \text{ (member}_K(w_{cK}) \text{ says that class is c) then class is c,}$$

where the asserts "member$_k(w_{ck})$ says that class is c" have a certainty determined by the rules in the sublist $\mathscr{G}_c^k$, after multiplying their confidence degrees by the same factor w_{ck}:

$$\text{if } x \text{ is } A_j^k \text{ then Class } C_j^k \text{ with } w_{C_j^k k} \cdot CF_j^k.$$

Again, those rules where $w_{C_j^k k} = 0$ are omitted.

In this case, any of these hierarchical rule bases we have introduced is univocally determined by a matrix $[w_{ck}]$. Therefore, the genetic search of the best selection involves finding the best matrix $[w_{ck}]$, according to certain criteria that will be defined next. Notice that, this search is a selection, because $[w_{ck}]$ is a sparse matrix. As we will detail later, in this contribution the number of terms w_{ck} different than zero is a design parameter.

3.2 Extension to fuzzy data

Let us relax the assumption that the input to the ensemble is a vector of numbers, and extend the combination to the case where the data is described by means of a fuzzy membership function. Applying the extension principle, the intermediate output of the k-th member of the ensemble is a type-II fuzzy set, whose α-cuts are

$$[\widetilde{R}^k(\widetilde{X})(c)]_\alpha = \left\{\bigvee_{\{j|C_j^k=c\}} CF_j^k \cdot A_j^k(x) \mid x \in [\widetilde{X}]_\alpha\right\}. \quad (4)$$

Each component FRBCS maps an input value x to so many fuzzy degrees of membership as the number of classes in the problem. An imprecise pattern cannot always be assigned a unique class by the rule base, however we can determine a fuzzy set of classes for the object being perceived through $\widetilde{X}$,

$$[\text{FRBCS}^k(\widetilde{X})]_\alpha = \left\{\arg\max_{c\in\{1,\ldots,n_c\}} R^k(x)(c) \mid x \in [\widetilde{X}]_\alpha\right\} \quad (5)$$

This extended reasoning method is still compatible with all the different linguistic combination mentioned before, either the binary matrix $[b_{ck}] \in \{0,1\}^{n_c\times K}$ or the sparse matrix of weights $[w_{ck}] \in [0,1]^{n_c\times K}$. In this last case, the output of the ensemble is, again, a type-II fuzzy set,

$$[R^{II}(\widetilde{X})(c)]_\alpha = \left\{\bigvee_{\{(j,k)|C_j^k=c\}} w_{ck} \cdot CF_j^k \cdot A_j^k(x), x \in [\widetilde{X}]_\alpha\right\}, \quad (6)$$

and the meaning of this expression fits the same hierachical knowledge base described in the preceding subsection.

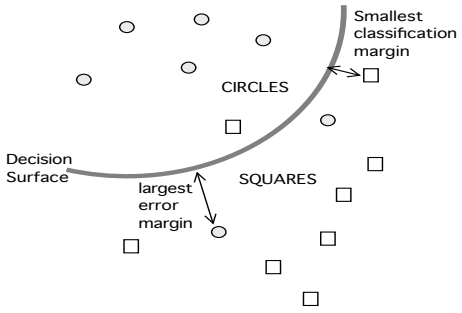


Figure 1: Three components of the fitness: (a) a quantile of the bootstrap estimation of the training error, (b) the largest distance between a misclassified example and the decision surface, and (c) the smallest distance between a correctly classified example and the decision surface.

3.3 Fitness function

We propose that the quality of a selected and combined fuzzy ensemble is defined by three components (e, m_1, m_2) (see Figure 1), thus the fitness of a possible FRBCS-CM design is a triplet comprising three real numbers:

1. Training error e: we compute the error of each ensemble for a large number of bootstrapped resamples of the training set, and use a quantile of the distribution of these errors as the first term of the fitness. This is intended to avoid overfitting when there are outliers in the training set, and also to detect the most robust selections, which are expected to generalize better.

2. Error margin m_1 : the second component of the fitness function depends on the distance between the misclassified examples and their nearest decision surface. Given an example x, we have approximated this value by the difference between the highest and the second highest term of $R^{II}(x)(c)$, and defined that the error margin of an ensemble is the worst (the highest) value of this difference for any example x in the training set.

3. Classification margin m_2: the third component depends on the distance between the correctly classified instances and their nearest decision surface, which is approximated, by the difference between the highest and the second highest terms in $R^{II}(x)(c)$. Here, the margin of an ensemble is the lowest value of this difference for all the examples of the training set; we seek a decision surface with the highest margin.

A lexicographical ordering is defined between two triplets:

$$(e, m_1, m_2) \prec (e', m_1', m_2') \iff \begin{cases} (e < e') \\ (e = e') \text{ and } m_1 < m_1' \\ (e = e') \text{ and } (m_1 = m_1') \text{ and } (m_2 > m_2') \end{cases} \quad (7)$$

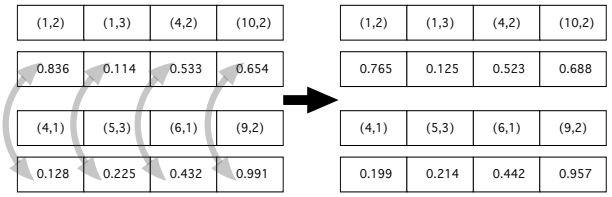


Figure 2: Coding scheme and crossover operation: an individual is a sparse matrix, which is represented by a list of indices and a list of values.

The extension of this criteria to fuzzy values is not included here because of space reasons; in short, the three numbers e, m_1 and m_2 become fuzzy numbers and the lexicographic ordering is no longer possible, thus a true multicriteria optimization is needed.

3.4 Coding scheme, genetic operators and evolutionary scheme

An individual is an sparse matrix $[w_{ck}]$, which will be stored as two fixed-length ordered lists of indexes (c,k) and their corresponding values w_{ck}, as displayed in Figure 2. The chromosome length is defined according to the maximum percentage of non-zero values in the matrix, which a parameter whose value is set by the user in advance. The initial population is randomly generated. We have decided to apply an arithmetic crossover [16] between the lists of values of both individuals, leaving the lists of indices unchanged. The mutation operator randomly alternates a nonuniform mutation of an element of the list of values [23] and the random generation of a completely new individual.

Lastly, since the fitness function is not scalar, we have decided to implement a tournament-based steady state GA [13], where at each generation the two last elements in each tournament are replaced by the offspring of the two winners. This offspring is the result of the application of the crossover operator mentioned before, followed by a mutation with certain probability.

4 EXPERIMENTS AND ANALYSIS OF RESULTS

4.1 Experimental setup

A preliminary evaluation of the performance of FRBCS-CM over crisp data is included in this section. Although is by no means a complete experimental assessment of the quality of this approach (fuzzy data has not been included yet), it allows the reader to grasp the behaviour of the proposed ensembles. Nine popular data sets have been selected (see Table 1), whose number of features ranges from 5 to 60, and the number of examples does so from 208 to 5,477.

Table 1: Data sets considered

Data set	#attr.	#examples	#classes
Sonar	60	208	2
Vehicle	18	846	4
P-Blocks	10	5,477	5
Pima	8	768	2
Glass	9	214	7
Breast	9	699	2
Yeast	8	1484	10
Heart	13	270	2
Phoneme	5	5,404	2

The experiments are based on Dietterich's 5×2-cv design [12]. The bagging FRBMCS generated are initially comprised by 50 classifiers. The granularity and the number of features used to derive them are both equal to 5. The GA for the FRBCS-CM derivation works with a population of 100 individuals and runs for 2,000 generations The tournament size is 5 and the mutation probability considered is 0.1. Four different values have been considered for the chromosome size: 10, 25, 50 and 90% of the terms of $[w_{ck}]$ matrix were allowed to be non-zero, reporting the best choice in each case.

4.2 Experiments developed

The mean values and standard deviations of the test error are shown in Table 2. Observe that the fuzzy ensembles using new fuzzy linguistic combination method have improved the initial ensembles considering the classical voting combination in both accuracy and complexity in all cases but one. The new FRBCS-CM method outperforms the voting one in all considered datasets apart from Breast.

Table 2: Compared test error

Data set	Bagging	Linguistic selection	%red.
Sonar	0.2048 ± 0.030	$\mathbf{0.1932 \pm 0.028}$	75%
Vehicle	0.3570 ± 0.022	$\mathbf{0.3530 \pm 0.020}$	75%
P-Blocks	0.0896 ± 0.005	$\mathbf{0.0583 \pm 0.006}$	90%
Pima	0.2628 ± 0.021	$\mathbf{0.2500 \pm 0.019}$	90%
Glass	0.3879 ± 0.082	$\mathbf{0.3486 \pm 0.049}$	75%
Breast	$\mathbf{0.0323 \pm 0.005}$	0.0349 ± 0.006	10%
Yeast	0.4341 ± 0.022	$\mathbf{0.4334 \pm 0.021}$	75%
Heart	0.1719 ± 0.034	$\mathbf{0.1607 \pm 0.031}$	50%
Phoneme	0.2137 ± 0.006	$\mathbf{0.2042 \pm 0.008}$	90%

The statistical relevance of the differences is graphically shown in Figure 3, where we have displayed the 9 boxplots of the paired differences between the results the linguistic selection we have introduced in this paper and those of the bagging-based ensemble. Negative values signal an advantage to our method. The median of the difference has always been negative, but in Breast, where there is a tie between both methods.

The highest statistical significances are in Phoneme (p-value of 0.01 with a Wilcoxon test) and P-Blocks (p-value $< 10^{-5}$). Coincidentally, these are the hardest problems

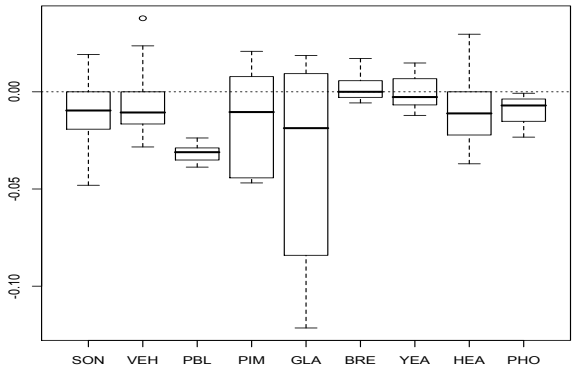


Figure 3: Results obtained by the FRBCS-CM.

of our benchmark. This may indicate that our FRBCS-CM will be able to improve the accuracy of the original fuzzy ensemble in those problems where bagging a random selection of fuzzy classifiers does not produce optimal results. Our bootstrap-based estimation of the training error has prevented an excessive overfitting and has also proven effective detecting those individual FRBCSs that should be removed, in those cases there is actually room for an improvement. Not being able to improve the results of Breast points in the same direction: in this dataset, the GFRBCS converged to ensembles with a null training error, thus all the resamples of the training sets will also have null error and the quantile of the bootstrap distribution does not give more information than a central moment of the training data. The error margin is null, either, so further improvements in the learning are guided by the third term of our fitness function, the classification margin. In this case, the decreased performance attributable to the linguistic interpretability of our ensemble (i.e., the use of a t-norm instead of a sum for combining the classifiers) cannot be compensated by removing the worst ensemble members.

5 CONCLUDING REMARKS AND FUTURE WORKS

We have proposed a novel MCS fuzzy combination method based on the use of a FRBCS automatically derived by means a GA. The new fuzzy linguistic combination method shows interesting characteristics such as transparency and capability to jointly perform classifier fusion and selection. The hierarchical structure of the system, when combined with a fuzzy classifier ensemble, makes FRBCSs able to deal with high dimensional problems.

Our next steps will include comparing the introduced FRBCS-CM with some other crisp and fuzzy combination methods, and with other OCS classifier selection methods.

References

[1] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Trans. on Neural Networks*, 5(4):537–550, 1994.

[2] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.

[3] P. Bulacio, S. Guillaume, E. Tapia, and L. Magdalena. A selection approach for scalable fuzzy integral combination. *Information Fusion*, 2010. in press.

[4] J. Casillas, F. Herrera, R. Pérez, M.J. del Jesus, and P. Villar. Special issue on genetic fuzzy systems and the interpretability-accuracy trade-off. *Int. J. of Approx. Reason.*, 44(1), January 2007.

[5] S.-B. Cho and J.H. Kim. Multiple network fusion using fuzzy logic. *IEEE Trans. on Neural Networks*, 6(2):497–501, 1995.

[6] O. Cordón, M.J. del Jesus, and F. Herrera. A proposal on reasoning methods in fuzzy rule-based classification systems. *Int. J. of Approx. Reason.*, 20:21–45, 1999.

[7] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena. Ten years of genetic fuzzy systems: Current framework and new trends. *Fuzzy Sets and Syst.*, 141(1):5–31, 2004.

[8] O. Cordón and A. Quirin. Comparing two genetic overproduce-and-choose strategies for fuzzy rule-based multiclassification systems generated by bagging and mutual information-based feature selection. *Int. J. of Hybrid Intell. Syst.*, 7(1):45–64, 2010.

[9] O. Cordón, A. Quirin, and L. Sánchez. A first study on bagging fuzzy rule-based classification systems with multicriteria genetic selection of the component classifiers. In *3rd IEEE Int. Workshop on Genetic and Evolving Fuzzy Syst. (GEFS)*, pages 11–16, Witten-Bommerholz, 2008.

[10] O. Cordón, A. Quirin, and L. Sánchez. On the use of bagging, mutual information-based feature selection and multicriteria genetic algorithms to design fuzzy rule-based classification ensembles. In *Eighth Int. Conf. on Hybrid Intell. Syst. (HIS)*, pages 549–554, Barcelona, 2008.

[11] B.V. Dasarathy and B.V. Sheela. A composite classifier system design: Concepts and methodology. *Proceedings of IEEE*, 67(5):708–713, 1979.

[12] T.G. Dietterich. Approximate statistical test for comparing supervised classification learning algorithms. *Neural Computation*, 10(7):1895–1923, 1998.

[13] A.E. Eiben and J.E. Smith. *Introduction to Evolutionary Computing*. Natural Comput. Series. Springer, 2nd edition edition, 2007.

[14] B. Gabrys and D. Ruta. Genetic algorithms in classifier fusion. *Appl. Soft Comput.*, 6(4):337–347, 2006.

[15] G. Giacinto and F. Roli. Dynamic classifier selection based on multiple classifier behaviour. *Pattern Recogn.*, 34(9):1879–1881, 2001.

[16] F. Herrera, M. Lozano, and J.L. Verdegay. Tackling real-coded genetic algorithms: Operators and tools for the behaviour analysis. *Artif. Intell. Review*, 12:265–319, 1998.

[17] T. Ho. The random subspace method for constructing decision forests. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 20(8):832–844, 1998.

[18] H. Ishibuchi, T. Nakashima, and M. Nii. *Classification and Modeling With Linguistic Information Granules*. Springer, 2005.

[19] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas. On combining classifiers. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 20(3):226–238, 1998.

[20] L. Kuncheva. "Fuzzy" versus "nonfuzzy" in combining classifiers designed by boosting. *IEEE Trans. on Fuzzy Syst.*, 11(6):729–741, 2003.

[21] L. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley, 2004.

[22] L. Lam and C.Y. Suen. Application of majority voting to pattern recogn.: An analysis of its behavior and performance. *IEEE Trans. on Syst., Man, and Cybernetics*, 27:553–568, 1997.

[23] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, 1996.

[24] D. Optiz and R. Maclin. Popular ensemble methods: An empirical study. *J. of Artif. Intell. Research*, 11:169–198, 1999.

[25] D. Partridge and W.B. Yates. Engineering multiversion neural-net systems. *Neural Computation*, 8(4):869–893, 1996.

[26] K. Trawinski, O. Cordón, and A. Quirin. On the combination of accuracy and diversity measures for genetic selection of bagging fuzzy rule-based multiclassification systems. In *9th Int. Conf. on Intell. Syst. Design and Applications (ISDA)*, Pisa, 2009.

[27] K. Woods, W.P. Kegelmeyer, and K. Bowyer. Combination of multiple classifiers using local accuracy estimates. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 19(4):405–410, 1997.

COMPARACIÓN DE MÉTODOS DE EVALUACIÓN EN ALGORITMOS GAP CON REPRESENTACIÓN DE IMPRECISIÓN

José R. Villar[1], **Enrique de la Cal**[1], **Marco García-Tamargo**[1], **Javier Sedano**[2]

[1]*Departamento de Informática, Universidad de Oviedo*, {villarjose,delacal,marco}@uniovi.es
[2]*Instituto Tecnológico de Castilla y León*, javier.sedano@itcl.es

Resumen

En entornos reales, el ruido proveniente de la propia instalación incrementa el nivel de imprecisión y los efectos indeseados en los sensores. Los datos obtenidos de esta manera se denominan Datos de Baja Calidad. Es necesario disponer de una herramientas de representación de la imprecisión si se desea aprender modelos en este tipo de situaciones. Este trabajo describe una representación de la imprecisión y cómo aprender modelos de caja blanca basados en ecuaciones. Se comparan las prestaciones ante diferentes niveles de ruido y con diferentes tipologías de funciones de evaluación. Los resultados permiten afirmar que el uso de los tres objetivos descritos mejora la eficiencia de los algoritmos. Igualmente, se ha comprobado que cada meta-heurística condiciona la tipología de la función de evaluación de los modelos.

Palabras Clave: Programación Genética, Algoritmos Genéticos, Datos de Baja Calidad.

1 INTRODUCCIÓN

Es bien conocido que algunos tipos de sensores comerciales no mantienen una relación exacta con la magnitud a medir; por el contrario, presentan una serie de diferencias, entre las que se encuentran la histéresis, rangos de no linealidad, etc. En entornos reales, el ruido proveniente de la propia instalación incrementa el nivel de imprecisión y los efectos indeseados en los sensores. Los datos obtenidos de esta manera se denominan Datos de Baja Calidad (Low Quality Data, LQD).

La necesidad de algoritmos capaces de aprender ante LQD es bien conocida en la literatura [5]: cuanta más imprecisión e incertidumbre presentan los datos, peores serán las prestaciones de los algoritmos de aprendizaje diseñados para datos discretos sin imprecisión.

Los sistemas basados en reglas difusas (SBRD) se han utilizado como modelos apropiados para representar la incertidumbre propia del proceso. Sin embargo, se requiere un análisis de la complejidad intrínseca de los data sets para seleccionar el mejor SBRD, esto es, para su aprendizaje y ajuste. En [9] se analizan varias de estas medidas utilizando un mismo método de aprendizaje para validar los resultados. Estas medidas deberían extenderse para LQD.

En nuestra opinión, uno de los estudios más interesantes en Soft Computing en el aprendizaje con datos imprecisos se presenta en [3]. En este estudio se muestran las bases matemáticas para el aprendizaje de sistemas genético difusos ante datos imprecisos asumiendo éstos como números difusos; es decir, para cada α-cut se obtiene un valor intervalar de cada uno de los datos. Debido al uso de esta representación de la imprecisión en los datos, las funciones de evaluación de los diferentes modelos son también difusas y, por lo tanto, es necesario adaptar los algoritmos a este tipo de función de evaluación o fitness [11]. Estas ideas se han utilizado en varias aplicaciones con data sets realísticos basados en problemas reales [12, 16].

En este estudio se resume una propuesta de representación de imprecisión en modelos de caja blanca basados en ecuaciones (MBE). Esta propuesta está basada en la evolución mediante la variante multi-objetivo de la metaheurística Simulated Annealing (MOSA) [13], de estructuras genéticas hibridas de algoritmos genéticos y algoritmos de programación genética (GAP) [6], que representan en si el modelo de imprecisión. Se pretende comparar dos métodos diferentes de evaluar los modelos, una basada en el uso de valores α-cut y otro basado en la obtención de las cotas superior e inferior del error, ambos sin una relación de orden total. Se analizará el comportamiento del aprendizaje de estos modelos ante funciones bien conocidas a las que se les añadirá diferentes niveles de ruido.

La organización de este trabajo es como sigue. Primeramente, en la siguiente sección, se introducirá la representa-

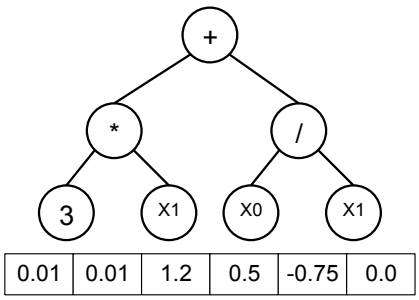


Figura 1: Representación de un MBE. El árbol se corresponde con la ecuación $y = 0.5 * x_1 + x_0/x_1$. Cada modelo contiene el árbol de la estructura y el conjunto de constantes propias.

ción de imprecisión en los MBE, evolucionando mediante la técnica MOSA estructuras GAP imprecisas. Posteriormente, en al sección 3 se describirán los dos métodos de evaluación analizados. En la Sección 4 se presentarán los data sets a utilizar en la comparación y se comentarán los resultados obtenidos. Se finaliza este estudio con las conclusiones y trabajos futuros.

2 REPRESENTACIÓN DE LA IMPRECISIÓN EN MODELOS GAP

Si bien es posible representar la imprecisión dentro del conjunto de datos [16], esta opción implica mayor coste computacional y complejidad en los procesos de aprendizaje.

En este estudio se plantea representar la imprecisión en el proceso como algo inherente a cada modelo, considerando los datos como discretos. Así, se plantea el uso de MBE del tipo descrito en la Figura 1, donde una ecuación relaciona la variable de salida con las diferentes variables de entrada.

La programación genética es típicamente utilizada en aprendizaje de MBE [7], de forma que las constantes están incrustadas en el árbol de estructura. Es posible, sin embargo, hibridar GP's con algoritmos genéticos, de forma que las constantes se mantienen independientemente del árbol de estructura, quien indexa sobre el vector de constantes. Este esquema de evolución es denominado GAP, alternando procesos evolutivos de ambos tipos en el tiempo [13]. Esta última es la opción utilizada en este estudio y reflejada en la Figura 1.

La representación de la imprecisión en MBE ha sido, en nuestra opinión, un tema poco estudiado en la literatura. Uno de los trabajos realizados se presenta en [10], donde se estudia la inclusión de constantes imprecisas. Análogamente, en [15] se planteó representar la imprecisión dentro del vectores de constantes. Siendo así, a cada variable catalogada como imprecisa se le asigna dos constantes $\{C^-, C^+\}$ de forma que al evaluar una de estas variables X_i en el dato d^i_j se genera un número difuso triangular $\{d^i_j - C^-, d^i_j, d^i_j + C^+\}$. Si se acepta una representación de la imprecisión simétrica entonces solo es necesaria una constante por variable imprecisa.

En este estudio se utiliza esta segunda representación de la imprecisión en los MBE, quedando para futuros trabajos la implementación de la imprecisión también para las constantes. Consecuentemente, el número de constantes disponibles para ser utilizadas dentro de una ecuación se reduce correspondientemente.

Una consecuencia importante de esta representación de imprecisión es que la evaluación de un modelo ante datos discretos no genera una salida discreta, sino que la salida es un número difuso. Es por ello que la evolución de estos modelos requiere de algoritmos de aprendizaje que gestionen funciones de evaluación difusas.

2.1 REPRESENTACIÓN DE LOS INDIVIDUOS GAP

Cada individuo se representa como un árbol de nodos siguiendo una gramática determinada conjuntamente con el vector de constantes, el árbol indexa sobre el vector de constantes para referirse a un valor numérico constante dentro de la ecuación (ver Figura 2).

Para cada una de las variables especificadas a-priori como imprecisas se reserva 1 ó 2 constantes para representar la imprecisión. El número de constantes por variable imprecisa es conocido también a-priori.

Se reservan, por tanto, las primeras constantes para la representación de la incertidumbre, dejando el resto para su uso por parte de la ecuación. Los valores para representación de la imprecisión se acotan a un porcentaje del rango de la correspondiente variable.

Tal y como fue analizado en [1], la generación aleatoria de individuos supera incluso a los métodos de generación de árboles con conocimiento a-priori, por lo que es el tipo de generación de individuos utilizado en este estudio. Para la generación de árboles se utilizan las reglas gramática que se describen seguidamente.

EXP $\mapsto$ OperadorUnario | OperadorBinario | Constante | Variable
OperadorUnario $\mapsto$ (sen EXP) | (cos EXP) | (**retardo** EXP)
OperadorBinario $\mapsto$ (+ EXP EXP) | (− EXP EXP) | (∗ EXP EXP) |
(/ EXP EXP) | (mín EXP EXP) | (máx EXP EXP)

2.2 OPERADORES GENÉTICOS

Se han incluido dos tipos de operaciones genéticas -cruce y mutación-, bien correspondientes a las operaciones GP,

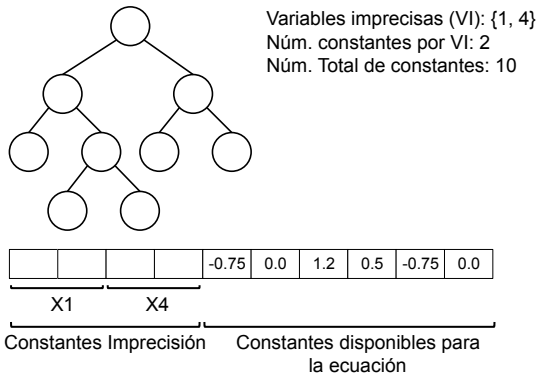


Figura 2: Representación de un individuo: el árbol representa la estructura y el vector de constantes contiene la representación de la imprecisión mas las constantes propias de la ecuación.

bien correspondientes a las operaciones GA. Las operaciones GP son estructurales y afectan exclusivamente a los árboles, mientras que las GA afectan exclusivamente a las constantes. Tanto unas como otras se aplican de acuerdo a las probabilidades definidas a-priori.

El cruce GP se realiza mediante un torneo binario por el que se seleccionan los dos individuos a participar. Para cada uno de ellos se selecciona aleatoriamente un nodo, procediéndose a intercambiar los nodos entre los individuos y sin intercambio de constantes. La mutación GP modifica el tipo de nodo siempre dentro de los nodos del mismo tipo: en caso de ser un operador, entre los tipos de nodo de igual aridad, en caso de ser una variable o una constante, entre los índices posibles en cada caso.

Por su parte, el cruce GA es el clásico cruce de dos puntos que intercambia las correspondientes partes de los vectores de constantes. Los individuos son seleccionados también por torneo binario. En cuanto a la mutación GA, para cada constante, y acorde a la probabilidad de mutación, se asigna un valor aleatorio dentro del rango de valores admisibles para la misma.

3 FUNCIONES FITNESS DE EVALUACIÓN DE LOS MODELOS

Dada la representación de imprecisión utilizada, la evaluación de un MBE con unos datos discretos como entrada genera un número difuso. Consecuentemente, si se utiliza una medida de error o de bondad como pueda ser el error cuadrático o el error cuadrático medio, ésta también será un número difuso. Si se trata de buscar el mejor modelo respecto a una medida de error como ésta debemos ser capaces de ordenar las evaluaciones obtenidas, bien mediante un orden total, bien mediante un orden parcial.

Si bien es posible establecer relaciones de orden total en números difusos [8], éstas conllevan un conocimiento a-priori del problema. Este hecho implica que para cada problema se ha de utilizar la correcta relación de orden para obtener los resultados óptimos. Por contra, en este estudio se plantea comparar métodos de evaluar los MBE, de forma que podamos establecer un orden parcial.

La primera relación de orden es la utilizada en [15], que se denominará αMSE. En este trabajo se propone el error cuadrático medio como medida de fitness de un individuo, objetivo a minimizar. Esta medida se simplifica mediante la definición de un α-cut, empleando el correspondiente intervalo como error y acumulando los errores cuadráticos intervalares para todos los ejemplos del data set. En este caso, se realiza una comparación intervalar, de forma que para afirmar que un valor es menor (o mayor) que otro la intersección debe ser nula; en caso contrario no se puede establecer un orden.

Igualmente, esta propuesta incluye el uso de otros dos objetivos complementarios: el *ancho medio del intervalo* (ANCHO) de salida del modelo para el α-cut definido, objetivo a minimizar, y el *porcentaje de ejemplos del data set cubiertos* (PCT), esto es, aquellos para los que la variable de salida está contenida en el intervalo de salida del modelo para el α-cut definido, objetivo a maximizar.

La segunda relación de orden se describe en [11], utiliza el error cuadrático medio para cada ejemplo del data set y se denominará *boundedMSE*. En el caso que el error cuadrático sea un número difuso, en caso de incluir el origen, se pliega sobre el semi-eje positivo. Posteriormente, se integran las áreas de las funciones de pertenencia por encima y por debajo del punto de máxima pertenencia, generando así las cotas superior e inferior del error cuadrático. Acumulando los valores de error para todos los ejemplos del data set se obtiene un valor intervalar. La comparación intervalar es la misma que la indicada para el caso anterior.

4 EXPERIMENTOS Y RESULTADOS

4.1 EXPERIMENTOS A REALIZAR

Para evidenciar problemas en el aprendizaje de MBE con imprecisión se ha planteado una batería de experimentos con data sets generados con fórmulas documentadas en la literatura, las cuales se describen en el siguiente apartado. Todas las pruebas se han evaluado para 3 diferentes niveles de imprecisión en los datos de entrada: 0 %, 1 % y 5 % del rango de las variables. Para todos ellos se utiliza el mismo valor de la variable de salida calculado para el caso sin ruido.

Los experimentos diseñados incluyen la evaluación del aprendizaje de MBE tal y como fue descrito en [15], utilizando como metaheurística el algoritmo MOSA descrito en [13]. Esta propuesta se ha analizado para los casos de

utilizar 1 solo objetivo (el error cuadrático a minimizar), dos objetivos (incluyendo además el PCT a maximizar) y tres objetivos (error cuadrático, PCT y ANCHO).

Para este último caso, y para los tres niveles de imprecisión indicados, se comparan resultados con la aproximación indicada en [11], utilizando igualmente el MOSA como metaheurística. No se incluye la comparación de esta opción con uno o dos objetivos debido a que, dado el esquema de imprecisión representado y la medida de error más completa que en este estudio se define, el MSE se mostraría ávido y se produciría una convergencia prematura del método. No obstante, se realizaron pruebas que demostraron este comportamiento y que, por motivos de espacio, no se incluyen en el presente estudio.

En todos los casos se utilizaron los parámetros que se indican seguidamente. Todas las variables se consideraron imprecisas, se utilizó una constante por variable para representar la imprecisión, con un margen de variación del 1% del margen de variación de la correspondiente variable. Se utilizaron modelos con un máximo de 10 constantes, con un valor absoluto máximo de 10.0. La profundidad máxima de los árboles se limitó a 7, el número máximo de nodos a 15. El tamaño máximo de población es de 100 individuos, con 1000 generaciones; como probabilidades de cruce y mutación se utilizaron 0.75, 0.25, tanto en GP como en GA. Los parámetros del MOSA relacionados con el simulated annealing se fijaron en $\Delta = 0.91629$, $T_{init} = 10$ y $T_{fin} = 1$.

4.2 DATA SETS ESTÁNDAR UTILIZADOS

Se han generado data sets con valores obtenidos de aplicar las fórmulas bien conocidas en la literatura de GP's. Todos los data sets incluyen 5 variables de entrada y 250 ejemplos. Para cada ejemplo, los valores de las variables de entrada se han generado aleatoriamente dentro de los rangos de las variables de entrada indicados en la correspondiente referencia. Las formulaciones utilizadas son las que se muestran a continuación.

T2: extraida de [2], la salida del data set se calcula como $f(x_1) = \sin(x_1) \cdot x_2 + 5$, con $x_i \in [-5,5]$, $\forall i = 1,\ldots,5$. Solo se utiliza la primera variable para el cálculo de la variable de salida.

P3: función polinomial extraída de [4], la salida del data set se calcula como $f(x_1) = 1.5 + 24.3 \cdot x_1^2 - 15 \cdot x_1^3 + 3.2 \cdot x_1^4$, con $x_i \in [-0,1]$, $\forall i = 1,\ldots,5$. Solo se utiliza la primera variable para el cálculo de la variable de salida.

F2: la función de Ackley tal y como se muestra en [14], la salida del data set se calcula como $f(\{x_i\}) = 20 - 20 \cdot \exp(-0.2 \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} x_i^2}) + \exp(1) - \exp(\frac{1}{N} \cdot \sum_{i=1}^{N} \cos(2 \cdot \pi \cdot x_i))$, con $x_i \in [-100,100]$, $\forall i = 1,\ldots,5$. Las primeras dos variables de entrada se utilizan en el cálculo de la variable de salida.

F5: la función de Schwefel escalada [14], la salida se calcula como $f(\{x_i\}) = 4.189828872724339 \cdot N - 0.01 \cdot \sum_{i=1}^{N} (x_i \cdot \sin(\sqrt{|x_i|})$, con $x_i \in [-500,500]$, $\forall i = 1,\ldots,5$. De nuevo, las dos primeras variables de entrada se usan en el cálculo de la variable de salida.

F8: la función discontinua de Schaffer expandida [14], calculando la variable de salida como $f(\{x_i\}) = \sum_{i=1}^{N} F(y_i, y_{(i+1)\%N})$, con $x_i \in [-500,500]$, donde $y_i = x_i$ cuando $|x_i| \leq 0.5$, $y_i = round(2 \cdot x_i)/2$ en otro caso, $\%$ es el operador módulo, $F(x,y) = 0.5 + \frac{\sin^2(\sqrt{x^2+y^2}]) - 0.5}{1+0.001 \cdot \sqrt{x^2+y^2}}$ and $\forall i = 1,\ldots,5$. La variable de salida se calcula como función de las dos primeras variables de entrada.

F9: la función de Schaffer expandida y rotada [14], siendo la variable de salida calculada según $f(\{x_i\}) = \sum_{i=1}^{N} F(x_i, x_{(i+1)\%N})$, con $x_i \in [-500,500]$, donde $\%$ es el operador módulo, $F(x,y) = 0.5 + \frac{\sin^2(\sqrt{x^2+y^2}]) - 0.5}{1+0.001 \cdot \sqrt{x^2+y^2}}$ y $\forall i = 1,\ldots,5$. La variable de salida se calcula como función de las dos primeras variables de entrada.

Tabla 1: Resultados numéricos obtenidos con un objetivo (MSE) y dos objetivos (PSE+PCT) utilizando la evaluación α-MSE. DS es el data set, NR es el nivel de ruido introducido, MMN y MND son la media y mediana de los puntos medios de las 10 diferentes ejecuciones

Data set		α-MSE solo MSE		α-MSE MSE + PCT	
DS	**NR**	**MMN**	**MMD**	**MMN**	**MMD**
T2	0.00	0.0017	0.0004	2.2444	0.0555
	0.01	0.0014	0.0004	5.0000	0.0717
	0.05	0.0009	0.0005	0.0667	0.0001
P3	0.00	0.0006	0.0004	3.6442	0.0015
	0.01	0.0018	0.0008	5.0000	0.0023
	0.05	0.0033	0.0006	5.0000	0.0284
F2	0.00	0.0212	0.0113	1.2402	0.0476
	0.01	0.0236	0.0126	5.0000	1.5854
	0.05	0.0078	0.0036	4.9459	1.2482
F5	0.00	0.0823	0.0276	5.0000	5.0000
	0.01	0.0233	0.0174	5.0000	5.0000
	0.05	0.0354	0.0221	5.0000	0.4953
F8	0.00	0.0026	0.0006	5.0000	0.6258
	0.01	0.0234	0.0017	5.0000	0.3170
	0.05	0.0016	0.0012	5.0000	0.3557
F9	0.00	0.0298	0.0020	5.0000	0.2589
	0.01	0.0516	0.0002	1.6353	0.1562
	0.05	0.1141	0.0104	5.0000	0.1011

4.3 RESULTADOS

Para cada data set y nivel de ruido se han realizado 10 ejecuciones independientes, seleccionando en cada una de ellas el individuo que presentó el menor valor de punto central de error cuadrático medio. En la Tabla 1 se visualizan los resultados empleando la evaluación α-MSE [15] con un solo objetivo MSE y con dos objetivos (MSE y el porcentaje de puntos cubiertos -PCT-).

Tanto este problema de convergencia como el mantenimiento de la diversidad son aspectos de trabajo futuro inmediato.

En Tabla 2 se muestran los resultados obtenidos empleando tres objetivos: el MSE, el PCT y el ancho medio del intervalo (ANC). Las dos primeras columnas se obtuvieron empleando la evaluación α-MSE [15], mientras que las dos últimas corresponden a la evaluación *bounded*-MSE [11].

En la Figura 3 se muestra el boxplot comparativo de cómo evoluciona el aprendizaje de MBE utilizando la evaluación α-MSE ante un aumento del ruido en los datos.

De los resultados mostrados se desprende que el uso de dos objetivos (MSE+PCT) empeora claramente las prestaciones del algoritmo. Por otro lado, y como es de esperar, el uso exclusivo del MSE no implica incluir o contener los ejemplos del data set.

Por otra parte, el uso de los tres objetivos permite obtener MBE de menores cotas de imprecisión y que a la vez incluyan más ejemplos del data set. Se puede concluir que los tres objetivos son necesarios para que los MBE evolucionen adecuadamente cuando se utiliza la representación de la imprecisión descrita.

De la comparación de los resultados entre los dos métodos de evaluación se puede concluir que el α-MSE es más apropiado para el uso de MOSA. Por contra, el *bounded*-MSE necesita mayor número de generaciones y grandes poblaciones para que presente unos altos valores de eficiencia. Debido a las limitaciones de espacio no se pueden mostrar los resultados que soportan esta aseveración. También se comprueba que este segundo método es más sensible al problema. Parece evidente que este segundo método de evaluación es más apropiado para metaheurísticas del tipo NSGA-II que del tipo MOSA.

Del análisis de los boxplots, especialmente para los data sets F5 y F9, se observan peores resultados para los data sets con ruido al 1 % que para los que contienen un nivel de ruido del 5 %. Esto puede ser debido a dos factores, el primero de ellos viene dado porque al utilizar como criterio de selección uno de los objetivos en un problema multiobjetivo, estamos seleccionando individuos extremos en la aproximación al frente de Pareto. Por otra parte, este hecho también puede deberse a la aparición de convergencia prematura del algoritmo.

Tabla 2: Resultados numéricos obtenidos con tres objetivos (MSE+PCT+ANCHO) cuando se utiliza la evaluación α-MSE y la evaluación *bounded-MSE*. DS es el data set, NR es el nivel de ruido introducido, MMN y MND son la media y mediana de los puntos medios de las 10 diferentes ejecuciones

Data set		α-MSE MSE+PCT+ANC		bounded-MSE MSE+PCT+ANC	
DS	**NR**	**MMN**	**MMD**	**MMN**	**MMD**
T2	0.00	0.0084	0.0002	0.0209	0.0120
	0.01	0.0018	0.0002	0.0025	0.0001
	0.05	0.0052	0.0002	0.0401	0.0055
P3	0.00	0.0013	0.0004	0.0172	0.0000
	0.01	0.0209	0.0007	0.0178	0.0076
	0.05	0.0022	0.0009	0.0073	0.0053
F2	0.00	0.0168	0.0056	0.1430	0.1031
	0.01	0.0094	0.0030	0.2679	0.1167
	0.05	0.0131	0.0062	0.1140	0.0919
F5	0.00	0.0092	0.0012	0.9504	0.5703
	0.01	0.0049	0.0002	0.4204	0.4243
	0.05	0.0483	0.0036	0.4196	0.0960
F8	0.00	0.0367	0.0003	0.2603	0.0010
	0.01	0.0025	0.0001	0.5675	0.0331
	0.05	0.0413	0.0002	0.1941	0.0131
F9	0.00	0.0346	0.0002	0.0083	0.0044
	0.01	0.0269	0.0001	0.0181	0.0082
	0.05	0.0683	0.0007	0.0089	0.0070

5 CONCLUSIONES Y TRABAJOS FUTUROS

En este estudio se ha presentado una representación de la imprecisión en MBE a ser aprendidos mediante técnicas evolutivas. Se han comparado tanto la influencia del ruido en los datos como diferentes métodos de evaluación de los modelos. Los resultados permiten afirmar que el uso de los tres objetivos descritos mejora la eficiencia de los algoritmos. Igualmente, se ha comprobado que cada metaheurística condiciona la tipología de la función de evaluación de los modelos.

Como trabajos futuros cabe destacar la inclusión de constantes con imprecisión descrita en la literatura, la unificación de la imprecisión para todos los modelos y el estudio de la diversidad en las poblaciones que permita ampliar el espacio de búsqueda.

Agradecimientos

Este estudio ha sido financiado por el por el proyecto TIN2008-06681-C06-04 del Ministerio de Ciencia e Innovación, y por el proyecto del ITCL y FICYT denominado CONSOCO.

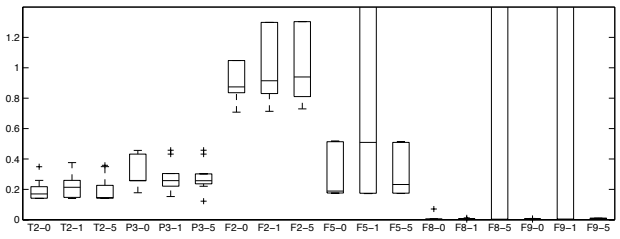


Figura 3: Boxplot de los resultados obtenidos para la evaluación α-MSE los diferentes data sets y niveles de ruido considerados(0, 1, y 5 se corresponden con niveles de ruido 0.00 %, 0.01 % y 0.05 %.

Referencias

[1] Alba Berzosa, José R. Villar, Javier Sedano, and Marco García-Tamargo. An study of the tree generation algorithms in equation based model learning with low quality data. In Emilio Corchado, Marek Kurzynski, and Michal Wozniak, editors, *Hybrid Artificial Intelligent Systems*, volume 6679 of *Lecture Notes in Computer Science*, pages 84–91. Springer Berlin / Heidelberg, 2011.

[2] Markus Brameier and Wolfgang Banzhaf. Explicit control of diversity and effective variation distance in linear genetic programming. In J.A. Foster et al., editor, *EuroGP 2002*, volume 2278 of *Lecture Notes in Computer Science*, pages 37–49. Springer-Verlag Berlin Heidelberg, 2002.

[3] Inés Couso and Luciano Sánchez. Higher order models for fuzzy random variables. *Fuzzy Sets Syst.*, 159:237–258, February 2008.

[4] Anikó Ekárt and S. Z. Németh. A metric for genetic programs and fitness sharing. In Riccardo Poli, Wolfgang Banzhaf, William Langdon, Julian Miller, Peter Nordin, and Terence Fogarty, editors, *Genetic Programming*, volume 1802 of *Lecture Notes in Computer Science*, pages 259–270. Springer Berlin / Heidelberg, 2000.

[5] Andres Folleco, Taghi M. Khoshgoftaar, Jason Van Hulse, and Amri Napolitano. Identifying learners robust to low quality data. *Informatica (Slovenia)*, 33(3):245–259, 2009.

[6] L. Howard and D. D' Angelo. The ga-p: a genetic algorithm and genetic programming hybrid. *IEEE Expert*, 10:11–15, 1995.

[7] John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA, 1992.

[8] H. Lee-Kwang and Jee-Hyong Lee. Method for ranking fuzzy numbers and its application to decision-making. *IEEE Transactions on Fuzzy Systems*, 7(6):677–685, 1999.

[9] Julián Luengo and Francisco Herrera. Domains of competence of fuzzy rule based classification systems with data complexity measures: A case of study using a fuzzy hybrid genetic based machine learning method. *Fuzzy Sets and Systems*, 161(1):3 – 19, 2010. Special section: New Trends on Pattern Recognition with Fuzzy Models.

[10] Luciano Sánchez. Interval-valued ga-p algorithms. *IEEE Transactions on Evolutionary Computation*, 4:64–72, 2000.

[11] Luciano Sánchez, Inés Couso, and Jorge Casillas. Genetic learning of fuzzy rules based on low quality data. *Fuzzy Sets and Systems*, 160(17):2524 – 2552, 2009. Theme: Learning.

[12] Luciano Sánchez, M. Rosario Suárez, J. R. Villar, and I. Couso. Mutual information-based feature selection and partition design in fuzzy rule-based classifiers from vague data. *Int. J. Approx. Reasoning*, 49:607–622, November 2008.

[13] Luciano Sánchez and José R. Villar. Obtaining transparent models of chaotic systems with multi-objective simulated annealing algorithms. *Inf. Sci.*, 178:952–970, February 2008.

[14] Adam Slowik. Fuzzy control of trade-off between exploration and exploitation properties of evolutionary algorithms. In Emilio Corchado, Marek Kurzynski, and Michal Wozniak, editors, *Hybrid Artificial Intelligent Systems*, volume 6678 of *Lecture Notes in Computer Science*, pages 59–66. Springer Berlin / Heidelberg, 2011.

[15] José R. Villar, Alba Berzosa, Enrique de la Cal, Javier Sedano, and Marco García-Tamargo. Multi-objective learning of white box models with low quality data. *Neurocomputing*, in Press.

[16] José R. Villar, Adolfo Otero, José Otero, and Luciano Sánchez. Taximeter verification using imprecise data from gps. *Eng. Appl. Artif. Intell.*, 22:250–260, March 2009.

Estimador recursivo subóptimo para el ajuste de modelos borrosos TS

Antonio Javier Barragán Piña[1], **Mariano José Aznar Torres**[2], **José Manuel Andújar Márquez**[1]

[1] *Departamento de Ingeniería Electrónica, de Sistemas Informáticos y Automática, Universidad de Huelva*, {antonio.barragan,andujar}@diesia.uhu.es

[2] *Máster de de Ingeniería de Control, Sistemas Electrónicos e Informática Industrial, Universidad de Huelva*, marianojose.aznar@alu.uhu.es

Resumen

Este trabajo presenta una metodología de modelado borroso completamente general y basada en el filtro de Kalman extendido, mediante la cual es posible afrontar el modelado en línea de un sistema no lineal de múltiples entradas y múltiples salidas (MIMO) sin restricciones en el tipo o distribución de las funciones de pertenencia utilizadas, y en presencia de ruido. El filtro de Kalman extendido permite obtener modelos de gran precisión en línea con el funcionamiento del sistema, y sin renunciar a la eficiencia computacional que lo caracteriza.

Palabras Clave: Filtro de Kalman, estimación, modelado, sistema borroso.

1 INTRODUCCIÓN

El filtro de Kalman se ha utilizado junto con la lógica borrosa en diversas aplicaciones, como en la extracción de reglas a partir de una base de reglas dada [10] o en optimización de los parámetros de mecanismos de defuzificación basados tanto en distribuciones Gaussianas como polinómicas [7].

La consecución de un modelo adecuado es un paso fundamental para la posterior aplicación de distintas técnicas tanto de análisis [3] como de diseño [2].

En este trabajo se presenta un algoritmo para el uso del filtro de Kalman extendido (*EKF*, *Extended Kalman Filter*) en la estimación de los parámetros adaptables de un modelo borroso general, esto es, sin restricciones ni en el tamaño de los vectores de entrada o salida, ni en el tipo o distribución de las funciones de pertenencia utilizadas al definir los conjuntos borrosos del modelo. De esta forma, los autores pretenden emplear las excelentes características del filtro de Kalman en la obtención de modelos borrosos pertenecientes a sistemas en principio desconocidos a partir de datos de entrada/salida provenientes de estos sistemas, y permitiendo su aplicación en tiempo real [8].

El presente artículo se organiza como sigue: en la sección 2 se presenta el problema del modelado borroso de forma completamente general. En la sección 3 se presenta el filtro de Kalman extendido y su uso en el modelado de sistemas. Posteriormente, se propone la aplicación del filtro de Kalman extendido al problema del modelado borroso. Finalmente, se presenta un ejemplo práctico y algunas conclusiones.

2 FORMULACIÓN DEL PROBLEMA

Sea n el número de variables de entrada y m el salidas de un sistema a modelar (dinámico o no), un modelo borroso discreto de múltiple entrada y múltiple salida –MIMO (*Multiple Input Multiple Output*)– general puede representarse mediante el siguiente conjunto de reglas [4, 12]:

$$\begin{aligned} &R^{(l,i)}\text{: Si } u_1(k) \text{ es } A^l_{1i} \text{ y} \ldots \text{y } u_n(k) \text{ es } A^l_{ni} \\ &\text{Entonces } y^l_i(k) = a^l_{0i} + \sum_{j=1}^{n} a^l_{ji} u_j(k), \end{aligned} \tag{1}$$

donde u_j son las entradas del modelo, $l = 1..M_i$ es el índice de la regla y M_i el número de reglas que modelan la evolución de la i-ésima salida del sistema, y_i.

Nótese que empleando la representación anterior, cada una de las salidas se puede modelar mediante un número distinto de reglas, lo que facilita la reducción de la cantidad total de reglas necesarias para modelar correctamente un sistema complejo, y, por lo tanto, facilita el proceso de modelado al reducir el número de parámetros del modelo.

Si el vector de entradas se extiende en una coordenada

$\tilde{u}_0 = 1$, el vector extendido, $\tilde{\mathbf{u}}$ adopta la forma [1, 2]:

$$\tilde{\mathbf{u}} = (\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_n)^{\mathrm{T}} = (1, u_1, \ldots, u_n)^{\mathrm{T}} \quad (2)$$

Utilizando el promedio ponderado como método de agregación y la extensión dada en (2), la salida y_i puede calcularse mediante [13, 14]:

$$y_i(k) = \sum_{j=0}^{n} a_{ji}(\mathbf{u})\tilde{u}_j(k), \quad (3)$$

siendo $a_{ji}(\mathbf{u})$ coeficientes variables [15] definidos por (4), donde $w_i^l(\mathbf{u})$ se calcula mediante (5) y representa el grado de activación de las reglas del modelo borroso.

$$a_{ji}(\mathbf{u}) = \frac{\sum_{l=1}^{M_i} w_i^l(\mathbf{u}) a_{ji}^l}{\sum_{l=1}^{M_i} w_i^l(\mathbf{u})} \quad (4)$$

$$w_i^l(\mathbf{u}) = \prod_{j=1}^{n} \mu_{ji}^l(u_j(k), \boldsymbol{\sigma}_{ji}^l) \quad (5)$$

$\mu_{ji}^l(u_j(k), \boldsymbol{\sigma}_{ji}^l)$ representa la j-ésima función de pertenencia de la regla l para la i-ésima salida del modelo, la cual determina el conjunto borroso A_{ji}^l. Los elementos $\boldsymbol{\sigma}_{ji}^l$ representan el conjunto de parámetros adaptables de dicha función de pertenencia, por lo que dichos valores, junto con los parámetros adaptables de los consecuentes de las reglas a_{ji}^l, han de ser determinados mediante el algoritmo de estimación para la obtención de un modelo adecuado del sistema.

3 FILTRO DE KALMAN EXTENDIDO

El filtro de Kalman fue desarrollado por Rudolph E. Kalman en 1960 [9], y permite construir un observador óptimo para sistemas lineales en presencia de ruido blanco tanto en el modelo como en las medidas. Posteriormente, el filtro de Kalman fue adaptado para su utilización en sistemas no lineales a través del filtro de Kalman extendido (EKF) [11], siempre que el sistema admita modelos linealizados en torno a cualquier punto de trabajo. Aunque el filtro de Kalman extendido no resulta óptimo, ya que se basa en una aproximación lineal del modelo y su precisión depende estrechamente de la bondad de dicha aproximación, es una herramienta extremadamente útil para la estimación en entornos con presencia de ruidos.

Si se considera un sistema discreto no lineal de la forma:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k)), \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{g}(\mathbf{x}(k)) + \mathbf{e}(k), \end{aligned} \quad (6)$$

donde $\mathbf{v}(k)$ y $\mathbf{e}(k)$ son ruidos blancos que representan la incertidumbre en el modelo de la ecuación de estado y salida, respectivamente, y siendo las matrices jacobianas del sistema:

$$\boldsymbol{\Phi}(\mathbf{x}(k), \mathbf{u}(k)) = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}(k), \mathbf{u}=\mathbf{u}(k)} \quad (7)$$

$$\boldsymbol{\Gamma}(\mathbf{x}(k), \mathbf{u}(k)) = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\mathbf{x}=\mathbf{x}(k), \mathbf{u}=\mathbf{u}(k)} \quad (8)$$

y

$$\mathbf{C}(\mathbf{x}(k)) = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}(k)} \quad (9)$$

el filtro de Kalman extendido puede resolverse mediante la aplicación iterativa del siguiente conjunto de ecuaciones [6]:

Predicción:

$$\hat{\mathbf{x}}(k|k-1) = \boldsymbol{\Phi}(k)\hat{\mathbf{x}}(k-1|k-1) + \boldsymbol{\Gamma}(k)\mathbf{u}(k) \quad (10)$$

$$\mathbf{P}(k|k-1) = \boldsymbol{\Phi}(k)\mathbf{P}(k-1|k-1)\boldsymbol{\Phi}^{\mathrm{T}}(k) + \mathbf{R}_v \quad (11)$$

Corrección:

$$\begin{aligned} \mathbf{K}(k) &= \big(\boldsymbol{\Phi}(k)\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}(k) + \mathbf{R}_{ve}\big) \\ &\quad \big(\mathbf{C}(k)\mathbf{P}(k|k-1)\mathbf{C}^{\mathrm{T}}(k) + \mathbf{R}_e\big)^{-1} \end{aligned} \quad (12)$$

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left(\mathbf{y}(k) - \hat{\mathbf{y}}(k)\right) \quad (13)$$

$$\begin{aligned} \mathbf{P}(k|k) &= \boldsymbol{\Phi}(k)\mathbf{P}(k|k-1)\boldsymbol{\Phi}^{\mathrm{T}}(k) + \mathbf{R}_v \\ &\quad - \mathbf{K}(k)\big(\mathbf{C}(k)\mathbf{P}(k|k-1)\boldsymbol{\Phi}^{\mathrm{T}}(k) + \mathbf{R}_{ve}^{\mathrm{T}}\big), \end{aligned} \quad (14)$$

donde $\hat{\mathbf{x}}(k|k-1)$ y $\hat{\mathbf{x}}(k|k)$ representan respectivamente la estimación del estado $\mathbf{x}$ *a priori*, y la estimación actualizada *a posteriori* dadas las observaciones actuales, e incluyendo el instante k. $\mathbf{P}(k|k-1)$ es la estimación de la matriz de covarianza *a priori*, y $\mathbf{P}(k|k)$ es la estimación actualizada de la matriz de covarianza *a posteriori*. $\mathbf{K}(k)$ es la ganancia de Kalman, $\hat{\mathbf{y}}(k)$ la estimación de la salida, y $\mathbf{R}_v$, $\mathbf{R}_{ve}$ y $\mathbf{R}_e$ son las matrices de covarianza del ruido, calculadas a partir del operador esperanza.

Una vez inicializado el filtro, éste evoluciona en línea con el sistema, obteniendo una solución que minimiza tanto el error de estimación como la matriz de covarianza del mismo para la linealización obtenida en cada instante. Si el sistema es completamente desconocido y no se posee una estimación inicial de $\hat{\mathbf{x}}(0)$, éste suele inicializarse con un valor arbitrario, en cuyo caso, la matriz de covarianza debe inicializarse como la matriz identidad multiplicada por un factor positivo elevado, que indica la poca certeza que se posee sobre el valor inicial de $\hat{\mathbf{x}}(0)$.

Una aplicación del filtro de Kalman extendido consiste en la identificación adaptativa de parámetros en sistemas no lineales, lo cual permite la obtención en línea

del conjunto de parámetros adaptables de un modelo no lineal discreto en presencia de ruido y de forma pseudo-óptima (es óptima para sistemas lineales). La identificación de un modelo borroso TS puede verse como la obtención de los parámetros de un modelo no lineal, por lo que puede ser aplicado el filtro de Kalman a través del algoritmo extendido para la estimación de dichos parámetros. En primer lugar se debe plantear el problema de estimación mediante el EKF, construyendo un sistema cuyos estados sean los parámetros a estimar, para posteriormente aplicar recursivamente las ecuaciones (10) a (14).

Sea $\mathbf{p}(k)$ el conjunto de parámetros a estimar del sistema borroso, e $\mathbf{y}(k)$ el conjunto de salidas del mismo, el sistema representado en (15) permite la obtención de dichos parámetros a partir de la aplicación del filtro de Kalman extendido.

$$\begin{aligned} \mathbf{p}(k+1) &= \mathbf{p}(k) \\ \mathbf{y}(k) &= \mathbf{h}(\mathbf{p}(k)) + \mathbf{e}(k) \end{aligned} \tag{15}$$

El conjunto de parámetros, $\mathbf{p}(k)$, es el vector de estado del sistema (15), y $\mathbf{h}(\mathbf{p}(k))$ es el modelo borroso cuyos parámetros se desean adaptar. $\mathbf{e}(k)$ es la incertidumbre de la medida de la salida del sistema, y viene representada por un ruido blanco de covarianza $\mathbf{R}_e$.

El primer paso a realizar es el cálculo de las matrices jacobianas del sistema, expresiones (7), (8) y (9), sobre el sistema (15), esto es:

$$\mathbf{\Phi}\left(\mathbf{p}(k)\right) = \mathbf{I}, \tag{16}$$

$$\mathbf{\Gamma}\left(\mathbf{p}(k)\right) = \mathbf{0}, \tag{17}$$

$$\mathbf{C}\left(\mathbf{p}(k)\right) = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{p}}\right|_{\mathbf{p}=\hat{\mathbf{p}}(k)}, \tag{18}$$

siendo $\hat{\mathbf{p}}(k)$ la estimación actual del vector de parámetros del modelo borroso.

Nótese que dada la formulación planteada en la sección 2, el problema de estimación consiste en determinar los valores adecuados de los parámetros adaptables tanto de los antecedentes, $\boldsymbol{\sigma}_{ji}^l$, véase (5), como de los consecuentes de las reglas, a_{ji}^l, véase (1). Por lo tanto, para un modelo borroso TS, la expresión $\mathbf{h}(\mathbf{p}(k))$ se corresponde con la ecuación (3), y la expresión (18) debe obtenerse a partir de la derivada de dicha expresión con respecto a cada uno de los parámetros adaptables del modelo borroso.

Como puede comprobarse en (3) y (4), la función $\mathbf{h}(\cdot)$ es lineal respecto del conjunto de parámetros adaptables de los consecuentes, a_{ji}^l, de forma que:

$$\frac{\partial h_i}{\partial a_{JI}^L} = \begin{cases} \dfrac{w_I^L x_J}{\sum_{l=1}^{M_i} w_i^l} & \text{si } i = I \\ 0 & \text{si } i \neq I, \end{cases} \tag{19}$$

donde L, J e I determinan el parámetro concreto a_{JI}^L del conjunto a_{ji}^l posible de parámetros del consecuente. Por otro lado, para cada conjunto de parámetros de las funciones de pertenencia del antecedente, se obtiene:

$$\frac{\partial h_i}{\partial \boldsymbol{\sigma}_{JI}^L} = \sum_{j=0}^{n} \frac{\partial \left(\dfrac{\sum_{l=1}^{M_i} w_i^l a_{ji}^l}{\sum_{l=1}^{M_i} w_i^l} \right)}{\partial \boldsymbol{\sigma}_{JI}^L} \tilde{x}_j \tag{20}$$

Sólo la salida I-ésima dependerá del parámetro $\boldsymbol{\sigma}_{JI}^L$, por lo tanto,

$$\frac{\partial h_i}{\partial \boldsymbol{\sigma}_{JI}^L} = 0 \text{ si } i \neq I. \tag{21}$$

Desarrollando la derivada parcial de (20), y teniendo en cuenta que el parámetro $\boldsymbol{\sigma}_{JI}^L$ sólo está presente en la regla L de la I-ésima salida, es fácil deducir que

$$\frac{\partial \left(\dfrac{\sum_{l=1}^{M_I} w_I^l a_{jI}^l}{\sum_{l=1}^{M_I} w_I^l} \right)}{\partial \boldsymbol{\sigma}_{JI}^L} = \frac{\partial w_I^L}{\partial \boldsymbol{\sigma}_{JI}^L} \left(\frac{\sum_{l=1}^{M_I} w_I^l (a_{jI}^L - a_{jI}^l)}{\left(\sum_{l=1}^{M_I} w_I^l \right)^2} \right) \tag{22}$$

Sustituyendo (22) en (20), se obtiene la expresión final para el cálculo de $\dfrac{\partial h_i}{\partial \boldsymbol{\sigma}_{JI}^L}$ cuando $i = I$:

$$\frac{\partial h_I}{\partial \boldsymbol{\sigma}_{JI}^L} = \frac{\partial w_I^L}{\partial \boldsymbol{\sigma}_{JI}^L} \sum_{j=0}^{n} \left(\frac{\sum_{l=1}^{M_I} w_I^l (a_{jI}^L - a_{jI}^l)}{\left(\sum_{l=1}^{M_I} w_I^l \right)^2} \right) \tilde{x}_j \tag{23}$$

Finalmente, para concluir el cálculo de la ecuación (23) es necesario determinar la derivada del grado de activación de las reglas del modelo borroso, w_i^l, con respecto a cada conjunto de parámetros de los antecedentes. Evidentemente este cálculo es dependiente del tipo de función de pertenencia que se utilice para cada antecedente. No obstante, es posible expresar de forma general:

$$\frac{\partial w_I^L}{\partial \boldsymbol{\sigma}_{JI}^L} = \frac{\partial}{\partial \boldsymbol{\sigma}_{JI}^L} \left(\prod_{q=1}^{n} \mu_{qI}^L (x_q(k), \boldsymbol{\sigma}_{qI}^L) \right), \tag{24}$$

o de forma más desarrollada,

$$\frac{\partial w_I^L}{\partial \boldsymbol{\sigma}_{JI}^L} = \frac{\partial \mu_{JI}^L (x_J(k), \boldsymbol{\sigma}_{JI}^L)}{\partial \boldsymbol{\sigma}_{JI}^L} \prod_{q=1, q\neq J}^{n} \mu_{qI}^L (x_q(k), \boldsymbol{\sigma}_{qI}^L) \tag{25}$$

Nótese que $\dfrac{\partial \mu_{JI}^L (x_J(k), \boldsymbol{\sigma}_{JI}^L)}{\partial \boldsymbol{\sigma}_{JI}^L}$ representa la derivada de la función de pertenencia definida mediante el conjunto de parámetros $\boldsymbol{\sigma}_{JI}^L$; por lo tanto, el cálculo de

dicha derivada depende del tipo de función de pertenencia utilizada, y podrá realizarse a partir de la expresión que define dicha función de pertenencia.

En el algoritmo 1 se muestra el procedimiento para la utilización del filtro de Kalman extendido en el modelado borroso, donde α y β son números reales positivos que permiten inicializar la matriz de covarianza del algoritmo.

Algoritmo 1 Algoritmo de adaptación paramétrica mediante EKF.

1: $\mathbf{p}_a$ = parámetros antecedentes iniciales
2: $\mathbf{p}_c = \mathbf{0}$
3: $\hat{\mathbf{p}}(0|-1) = [\mathbf{p}_a, \mathbf{p}_c]$
4: $\mathbf{P}(0|-1) = \begin{pmatrix} \mathbf{I}\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\alpha \end{pmatrix}$
5: **para** $k = 0..k_{end}$ **hacer**
6: Calcular $\mathbf{P}(k|k-1)$ mediante (11)
7: Estimar $\hat{\mathbf{y}}(k)$ usando el modelo borroso
8: Calcular $\partial\mathbf{w}/\partial\boldsymbol{\sigma}$ mediante (25)
9: Calcular $\mathbf{C}_a(k)$ mediante (23)
10: Calcular $\mathbf{C}_c(k)$ mediante (19)
11: $\mathbf{C}(k) = [\mathbf{C}_a(k), \mathbf{C}_c(k)]$
12: Obtener $\mathbf{K}(k)$ mediante (12)
13: Actualizar $\hat{\mathbf{p}}(k|k)$ mediante (10) y (13)
14: Actualizar $\mathbf{P}(k|k)$ mediante (14)
15: **fin para** k

4 Ejemplo

A continuación se muestra un ejemplo de aplicación práctica del filtro extendido de Kalman al modelado borroso.

Sea el circuito diodo-túnel mostrado en la Figura 1, donde $R = 1.5$ kΩ, $C = 2$ pF y $L = 5$ μH, y cuya dinámica puede ser representada por el modelo en espacio de estado MIMO (26), donde $x_1(t) = v_C(t)$, $x_2(t) = i_L(t)$, $T = 0.1$ns es el tiempo de muestreo, y $h(v)$ la característica $v_R - i_R$ no lineal del diodo-túnel [5].

$$\begin{aligned} \dot{x}_1(t) &= \frac{x_2(t) - h(x_1(t))}{C} \\ \dot{x}_2(t) &= \frac{u(t) - x_1(t) - Rx_2(t)}{L}. \end{aligned} \tag{26}$$

Se considera que el sistema está afectado por un ruido blanco $\mathbf{e}(k)$ de covarianza:

$$\mathbf{R}_e = \begin{pmatrix} 10^{-4} & 0 \\ 0 & 0.01 \end{pmatrix}, \tag{27}$$

y el estado inicial es $\mathbf{x}(k) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. La señal de entrada aplicada al sistema es

$$u_{tr}(k) = \frac{20\sin(kT/10)}{(kT+10)},$$

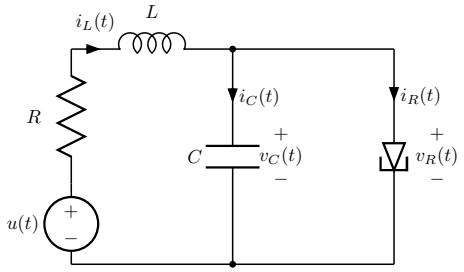


Figura 1: Circuito diodo túnel.

para el proceso de entrenamiento, y

$$u_{te}(k) = \frac{10\sin((kT+10)/5)}{(kT+15)},$$

para la validación.

Para probar la respuesta del algoritmo en función de los valores asignados a los parámetros α y β, se han realizado varias ejecuciones para distintos valores de dichos parámetros (ver Tabla 1). Con la intención de mitigar el efecto que el ruído puede causar en los errores de modelado, cada una de las pruebas se ha realizado 10 veces. Al ser el sistema (26) dinámico, las entradas del modelo son el estado actual, $\mathbf{x}(k)$, y la señal de entrada, $u(k)$; y sus salidas son el siguiente valor del vector de estado, $\mathbf{x}(k+1)$. Para la inicialización del modelo borroso se ha empleado: para $x_1(k+1)$, dos funciones de pertenencia gaussianas uniformemente distribuidas para cada una de las variables, generando un total de 8 reglas; y para $x_2(k+1)$, dos funciones de pertenencia gaussianas uniformemente distribuidas para cada una de las variables de estado, no incorporándose la entrada al antecedente de estas reglas, generando 4 reglas para la segunda de las salidas del modelo borroso.

Para los para los valores con los que obtiene un mejor rendimiento, $[\alpha, \beta] = [10^{-2}, 10^{-3}]$, los errores en línea se muestran en las Figuras 2 y 3, los errores de los modelos finales obtenidos a partir de los datos de entrenamiento son mostrados en las Figuras 4 y 5, y los errores de los modelos finales obtenidos a partir de los datos de validación se muestran en las Figuras 6 y 7, respectivamente. La Figura 8 muestra la salida del sistema y del modelo obtenido durante el modelado en línea, y finalmente, la Figura 9 muestra la respuesta del modelo final para la entrada de validación.

Como se puede observar, el algoritmo propuesto permite modelar sistemas dinámicos no lineales de múltiple entrada y múltiple salida en línea y en presencia de ruido, obteniendo muy buenos resultados.

Tabla 1: Error cuadrático medio (RMSE) de cada prueba.

$[\alpha, \beta] =$	$[1e-3, 1e-1]$	$[1e1, 1e-2]$	$[1e3, 1e1]$	$[1e-2, 1e-3]$	$[1e2, 1e-4]$
RMSE para x_1	$6.495995e-03$	$2.291368e-02$	$1.330786e-02$	$6.685541e-03$	$1.776174e+04$
RMSE para x_2	$1.131838e-02$	$1.744776e-02$	$1.890073e-02$	$9.344264e-03$	$2.313962e+04$

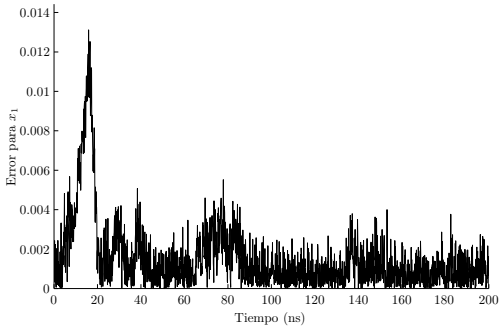


Figura 2: Evolución del error absoluto en línea de $x_1(k)$.

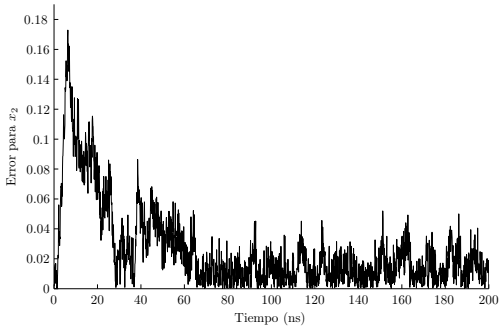


Figura 3: Evolución del error absoluto en línea de $x_2(k)$.

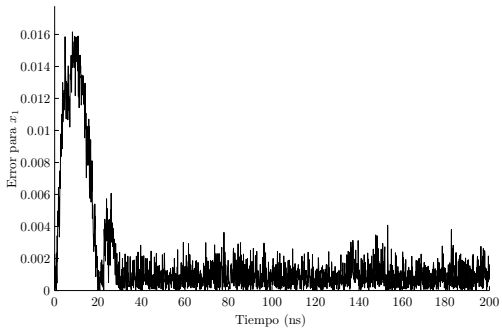


Figura 4: Error del modelo final para los datos de entrenamiento de $x_1(k)$.

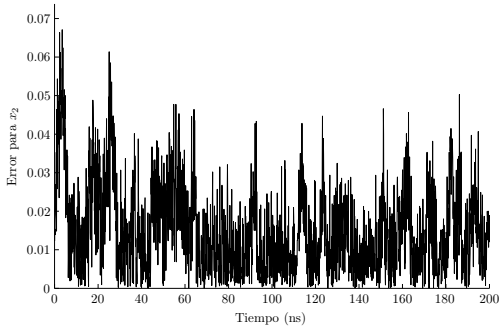


Figura 5: Error del modelo final para los datos de entrenamiento de $x_2(k)$.

5 CONCLUSIONES

En este trabajo se ha presentado un algoritmo para el uso del filtro de Kalman extendido en la estimación de los parámetros de un modelo borroso general, esto es, sin restricciones ni en el tamaño de los vectores de entrada o salida, ni en el tipo o distribución de las funciones de pertenencia utilizadas en la definición de los conjuntos borrosos del modelo. También se ha presentado un ejemplo práctico de aplicación del algoritmo propuesto, donde se muestra la buena respuesta del mismo y su rápida convergencia incluso en presencia de ruido, y la importancia de una inicialización adecuada del algoritmo. Actuamente los autores están trabajando en una metodología que permita realizar la inicialización en función de los primeros datos capturados del sistema con la intención de que ésta pueda adecuarse a las características del mismo y redunde en un comportamiento más robusto del modelado.

Agradecimientos

Este trabajo es una contribución del proyecto DPI2010-17123 financiado por el Ministerio de Educación y Ciencia y por la Unión de Desarrollo Regional Europea.

Referencias

[1] José Manuel Andújar and Antonio Javier Barragán. A methodology to design stable nonlinear fuzzy control systems. *Fuzzy Sets and Systems*, 154(2):157–181, September 2005.

[2] José Manuel Andújar, Antonio Javier Barragán, and Manuel Emilio Gegúndez. A general and formal methodology for designing stable nonlinear fuzzy control systems. *IEEE Trans. Fuzzy Syst.*, 17(5):1081–1091, October 2009.

[3] José Manuel Andújar, Antonio Javier Barragán, and Jesús Graiño. Una metodología de análisis de

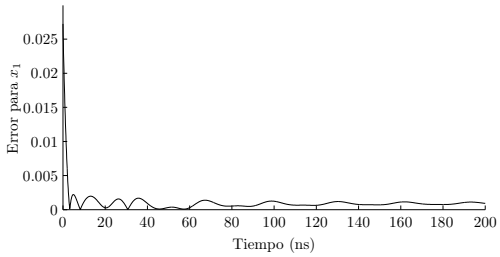


Figura 6: Error del modelo final para los datos de validación de $x_1(k)$.

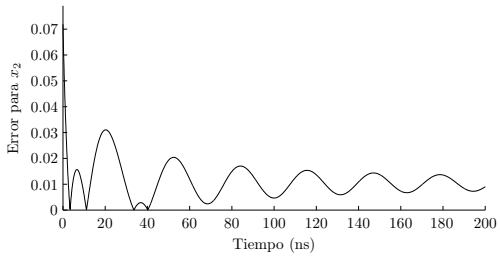


Figura 7: Error del modelo final para los datos de validación de $x_2(k)$.

sistemas dinámicos mediante lógica borrosa. In CEA-IFAC, editor, *XXX Jornadas de Automática*, Valladolid, España, September 2009.

[4] Robert Babuška. Fuzzy modeling – a control engineering perspective. In *Proceedings of FUZZ-IEEE/IFES'95*, volume 4, pages 1897–1902, Yokohama, Japan, March 1995.

[5] Leon O. Chua, Charles A. Desoer, and Ernest S. Kuh. *Linear and nonlinear circuits*. McGraw-Hill series in electrical and computer engineering: Circuits and systems. McGraw-Hill Book Company, New York, 1987.

[6] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. John Wiley & Sons, Inc., 2nd edition, 2001.

[7] Tao Jiang and Yao Tang Li. Generalized defuzzification strategies and their parameter learning procedures. *IEEE Trans. Fuzzy Syst.*, 4(1):64–71, August 1996.

[8] Agustín Jiménez, Basil M. Al-Hadithi, and Fernando Matía. An optimal T-S model for the estimation and identification of nonlinear functions. *WSEAS Trans. Sys. Ctrl.*, 3(10):897–906, January 2008.

[9] Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Transactions on ASME-Journal of Basic Engineering*, 82(series D):35–45, 1960.

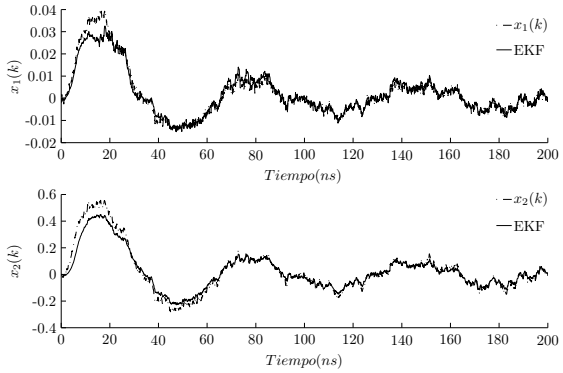


Figura 8: Respuesta del modelado en línea.

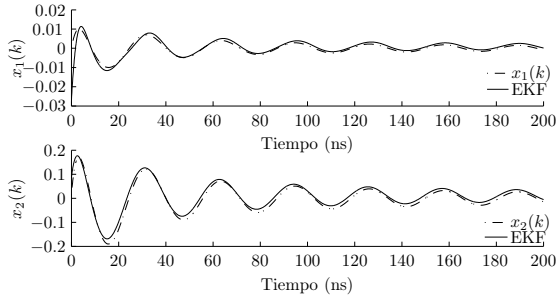


Figura 9: Respuesta del modelo final frente a la entrada de validación.

[10] Wang Liang and Yen John. Extracting fuzzy rules for system modeling using a hybrid of genetic algorithms and Kalman filter. *Fuzzy Sets and Systems*, 101(3):353–362, February 1999.

[11] Peter S. Maybeck. *Stochastic models, estimation, and control*, volume 141 of *Mathematics in Science and Engineering*. Academyc Press, 1979.

[12] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst., Man, Cybern.*, 15(1):116–132, 1985.

[13] L. X. Wang. *Adaptive fuzzy systems and control*. Prentice Hall, New Jersey, 1994.

[14] Li-Xin Wang. *A course in fuzzy systems and control*. Prentice Hall, New Yersey, USA, 1997.

[15] L.K. Wong, F.H.F. Leung, and P.K.S. Tam. Stability design of TS model based fuzzy systems. In *IEEE International Conference on Fuzzy Systems*, volume 1, pages 83–86, Barcelona, Spain, July 1997.

Modelos Neuro-Difusos basados en FasArt. Aplicaciones

Jose-Manuel Cano-Izquierdo, Julio Ibarrola, Miguel Almonacid

Universidad Politécnica de Cartagena, JoseM.Cano@upct.es

Resumen

El modelo FasArt se ha propuesto como una alternativa para la integración de los principios de la teoría de conjuntos difusos en arquitecturas de tipo ART. A FasArt le han seguido otra serie de modelos: FasBack, dFasArt y S-dFasArt que se han aplicado con éxito en diferentes problemas tanto asociados con el reconocimiento de patrones como la identificación y el control de sistemas dinámicos.

Palabras Clave: Arquitecturas Neuro-difusas, Modelos FasArt, Aplicaciones.

1. Introducción

Los modelos Fuzzy ART y Fuzzy ARTMAP [15, 14] se han presentado como el resultado de la integración de la Teoría de Conjuntos Difusos dentro del marco de la Teoría de la Resonancia Adaptativa (ART). Esta integración resulta cuestionable ya que se basa en la mera sustitución de los operadores "crisp" ($\bigcup, \bigcap$) por operadores fuzzy ($\bigvee, \bigwedge$). El modelo FasArt [6] se propuso como una alternativa que, manteniendo la arquitectura ART, establecía una inclusión más rigurosa de los conceptos difusos. El modelo FasArt utiliza una ley de aprendizaje basada en el principio de emparejamiento utilizado en ART, la inclusión de aprendizaje guiado por el error condujo a la propuesta del modelo FasBack [7].

El modelo dFasArt [10] incorpora una formulación de ecuaciones diferenciales a las principales expresiones que describen el funcionamiento de FasArt. Con ello se pretende que el factor tiempo cobre importancia a la hora de realizar la clasificación. Aspectos como la duración de un determinado estímulo (entrada) o el ordenamiento temporal en el que se presenta pueden influir en la clasificación. dFasArt y su versión supervisada S-dFasArt se plantean como alternativa a la hora de clasificar señales de carácter temporal como puede ser el error en un lazo de control o la lectura de un sensor.

En este trabajo se presentarán brevemente los diferentes modelos mencionados, aportándose en cada caso una descripción y referencias a las distintas aplicaciones para las que se han propuesto. De forma más de detallada se describe una aplicación de S-dFasArt para el aprendizaje online de señales EEG.

2. FasArt y FasBack

FasArt mantiene la estructura básica de la arquitectura Fuzzy ARTMAP, Figura 1, que consta de dos módulos de clasificación de tipo Fuzzy ART y una memoria asociativa denominada Mapa Inter-ART. Cada uno de los módulos Fuzzy ART consta de tres niveles:

- F_0: Registra la entrada y la normaliza utilizando código complementario.
- F_1: Realiza la comparación entre la entrada y los prototipos asociados a las categorías ya almacenadas.
- F_2: Constituida por unidades asociadas a cada una de las categorías que se han generado. La actividad de cada unidad se asocia al emparejamiento entre el prototipo (almacenado en forma de vector de pesos) y la entrada.

El mecanismo de Reset, que se dispara cuando el emparejamiento entre una entrada y el prototipo seleccionado no es suficiente, permite mantener el denominado principio de *estabilidad-plasticidad*, aprender nueva información sin que ello destruya la ya almacenada. El Mapa Inter-ART permite asociar las categorías que se generan en los clasificadores y, gracias al mecanismo de Reset Inter-ART, afinar la clasificación en Fuzzy ART-a cuando la asociación no resulta valida. Este mecanismo se traduce en el principio de *máxima generalización con mínimo error de predicción*

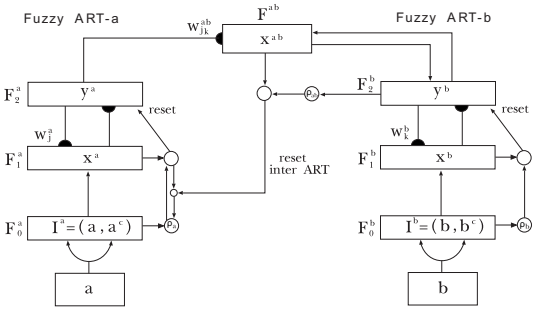


Figura 1: Arquitectura Fuzzy ARTMAP.

A esta estructura básica, FasArt incorpora el concepto de dualidad neuro-difusa, para ello se define una nueva función de activación para las unidades del nivel F_2 que también puede interpretarse como una función de pertenencia a una clase difusa asociada a la unidad y que se define como:

$$\eta_j(\mathbf{I}) = \prod_{i=1}^{M} \eta_{ji}(I_i) \tag{1}$$

donde j hace referencia a la *j-esima* unidad, $\mathbf{I} = (I_1 \dots I_M)$ representa la entrada, y la función $\eta_{ji}(I_i)$ (Figura 2) queda determinada a partir de los vectores de pesos: $\mathbf{W}, \mathbf{V}, \mathbf{C}$, asociados a la unidad.

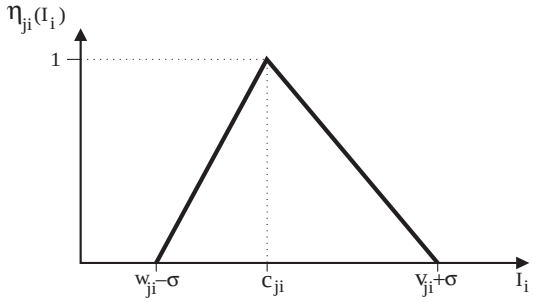


Figura 2: Función de activación/pertenencia en FasArt.

Siguiendo con la dualidad neuro-difusa que propone FasArt, el mapa Inter-ART define un conjunto de relaciones entre unidades (categorías difusas) que pueden describirse en forma de reglas difusas del tipo:

IF I_1^a is A_1^j AND $\dots I_{M_A}^a$ is $A_{M_A}^j$ THEN I_1^b is Y_1^j AND $\dots I_{M_B}^b$ is $Y_{M_B}^j$

donde los conjuntos difusos A_i y Y_j vienen definidos por los pesos de las unidades de nivel F_2 de acuerdo con la Figura 2 y las relaciones entre ellos por los pesos del mapa Inter-ART.

Utilizando esta interpretación es posible considerar que FasArt define un Sistema Lógico Difuso con inferencia mediante producto y fuzzificación por punto aislado y, utilizando defuzzificación mediante la media de los centros, calcular la salida del mismo como:

$$y_m(\mathbf{I}) = \frac{\sum_{l=1}^{N^b} \sum_{k=1}^{N^a} c_{lm}^b w_{kl}^{ab} \eta_k(\mathbf{I})}{\sum_{l=1}^{N^b} \sum_{k=1}^{N^a} w_{kl}^{ab} \eta_k(\mathbf{I})} \tag{2}$$

donde c_{lm} es el valor para el que $\eta_{R_l}^b$ es máximo.

Este resultado, unido a las propiedades de los Sistemas Lógico Difusos [16], permite que FasArt pueda aproximar cualquier función f, continua a tozos, definida sobre un compacto U, $f : U \subseteq \Re^M \rightarrow V \subseteq \Re$.

El modelo FasBack incluye aprendizaje guiado por el error, en concreto la utilización del método de Gradiente Descendente para la minimización del error de predicción. Utilizando éste método se modifican los pesos asociados a las unidades y los pesos de las conexiones que determinan las reglas.

2.1. Aplicaciones con FasArt/FasBack

La propiedad de representación de funciones mediante FasArt permite aventurar la idoneidad del modelo para tareas de identificación. En este sentido se ha propuesto y comparado con otras arquitecturas en problemas de representación de funciones a partir de datos con ruido [8, 13]. Como aplicación de esta propiedad se propuso la identificación y control de sistemas dinámicos [5, 4] y se han estudiado diferentes casos en los que se utilizaron datos provenientes de instalaciones industriales, como los de una torre de desisopentanización de la empresa REPSOL [2], biorreactores para producción de penicilina en Antibióticos SAU [1] o una máquina de refrigeración por energía solar [20].

Se puede destacar la utilización de FasArt/FasBack para varios cometidos:

1. Como sensor software para la estimación de variables con dificultad de ser medidas on-line.
2. Como modelo de la planta para estimar la salida de la misma en un esquema de control IMC.
3. Como controlador, ya sea de forma directa o en diferentes esquemas de control: feedforward, control adaptativo, control IMC.

Por otra parte FasArt también se ha aplicado a problemas de reconocimiento de patrones como el reconocimiento de caracteres manuscritos [18, 17]. En estos trabajos se utilizaron datos generados en el proyecto UNIPEN mediante el que se generó un benckmark internacional que permite hacer comparaciones de diferentes soluciones. En otra línea

de actuación se ha utilizado para el tratamiento de documentos estructurados [23, 24, 25] con el objetivo de generar de forma automática librerías digitales. También se ha propuesto su uso para detección de fallos en motores AC [26], donde se explora la posibilidad de interpretar el modelo como un conjunto de reglas de tipo difuso que describen, de una forma cercana al lenguaje natural, el mecanismo que conduce a la clasificación.

3. dFasArt y S-dFasArt

En principio, el orden de presentación de las entradas es un factor que no influye en los modelos de tipo Fuzzy ARTMAP y, por extensión, en el modelo FasArt. Hay ocasiones en las que el orden de presentación puede ser importante a la hora de clasificar una señal. Con el objetivo de tratar este tipo de problemas se formula el modelo dFasArt [10], que incorpora una formulación dinámica a las ecuaciones de FasArt. Para ello se proponen como ecuaciones de activación y Reset las expresiones:

$$\begin{aligned}
&\frac{dT_j}{dt} = -A_T T_j + B_T \prod_{i=1}^{M} \eta_{ji}(I_i^a(t)) \\
&T_j(0) = 1 \\
&\frac{dR_j}{dt} = (-A_R R_j + B_R \sum_{i=1}^{M} \frac{l_{ji}}{|2c_{ji}|+\varepsilon})(R_{max} - R_j) \\
&R_j(0) = 0 \\
&l_{ji} = max(v_{ji}, I_i^a) - min(w_{ji}, I_i^a)
\end{aligned} \tag{3}$$

Las ecuaciones de aprendizaje para los pesos asociados a las categorías también se consideran como ecuaciones diferenciales de la forma:

$$\begin{aligned}
\frac{d\mathbf{W}}{dt} &= -A_W \mathbf{W} + B_W min(\mathbf{I}^a(t), \mathbf{W}) \\
\frac{d\mathbf{C}}{dt} &= A_C(\mathbf{I}^a(t) - \mathbf{C}) \\
\frac{d\mathbf{V}}{dt} &= -A_V \mathbf{V} + B_V max(\mathbf{I}^a(t), \mathbf{V})
\end{aligned} \tag{4}$$

dFasArt se propuso como un modelo para aprendizaje no supervisado, a diferencia de FasArt que se utilizaba en aprendizaje supervisado, por lo que la versión supervisada se denominó S-dFasArt (Figura: 3) [28]. Como se aprecia en la figura se ha simplificado la estructura ARTMAP, sustituyendo el segundo clasificador por una capa se supervisión. En principio, S-dFasArt se aplica en problemas de clasificación supervisada donde la clase de los patrones de aprendizaje está ya establecida.

3.1. Aplicaciones con dFasArt/S-dFasArt

dFasArt se ha aplicado a problemas relacionados con la monitorización de lazos de control. En [11] se muestra como, de forma no supervisada, dFasArt puede detectar cambios en los parámetros de un controlador PID, y por lo tanto clasificar el funcionamiento del controlador, monitorizando el error de seguimiento. De forma parecida en [9] se

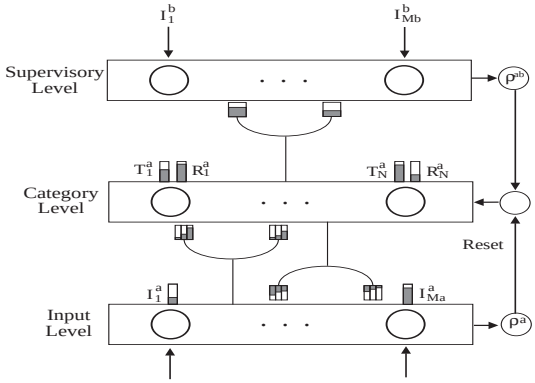


Figura 3: Arquitectura S-dFasArt.

propone la utilización de dFasArt para detectar cambios en las prestaciones de un controlador. dFasArt "aprende" cuál es el estado "aceptable" de la señal de error y detecta cuando se está produciendo una anomalía para, en su caso, proceder a una adaptación del controlador. Cuando el error vuelve al estado "aceptable" dFasArt lo detecta y permite que se detenga la adaptación del controlador, evitando problemas de "parameter drift". Un ejemplo de aplicación de este esquema al control de posición de un motor DC puede consultarse en [19].

S-dFasArt se ha propuesto como elemento clasificador para sistemas de navegación [22, 27]. Utilizando como entrada la señal proveniente de un localizador GPS, un módulo S-dFasArt permite identificar el tipo de maniobra que está realizando el vehículo (parada, velocidad constante, aceleración). En los trabajos citados se ha trabajado con datos de un vehículo experimental y se han comparado los resultados con otros métodos de clasificación. Los resultados muestran que S-dFasArt es adecuado para trabajar con datos que presentan incertidumbre.

Una última aplicación de S-dFasArt se puede encontrar en [12] y se inscribe en el diseño de interfaces cerebro-computador (BCI). Concretamente en el desarrollo de interfaces BCI espontaneas y asíncronas. Para ello, mediante una serie de sensores dispuestos en la cabeza se recoge la actividad eléctrica del cerebro (EEG). Estas señales son tratadas con el objeto de determinar procesos cognitivos que, en el caso de los sistemas espontáneos, no son activados por ninguna señal externa y, en los sistemas asíncronos no mantienen ninguna disposición temporal determinada sino que es el individuo de prueba quien determina el suceso. Se puede decir que el desarrollo de este tipo de BCI es, probablemente, el que presenta una mayor dificultad, sobre todo a la hora de detectar los cambios y clasificar correctamente los procesos cognitivos. En [12] se han utilizado datos de un benckmark internacional [3] (BCI Competition III Data Set V) para probar el funcionamiento de S-dFasArt para la clasificación de tres procesos cognitivos:

mover mano derecha, izquierda, pensar una palabra. Se ha hecho un estudio comparativo con los resultados que se obtuvieron en la BCI Competition y resultados publicados posteriormente (clasificadores estadísticos, RRNN, Algoritmos Genéticos, RDA, SVM ...).

4. Aprendizaje on-line de señales EEG

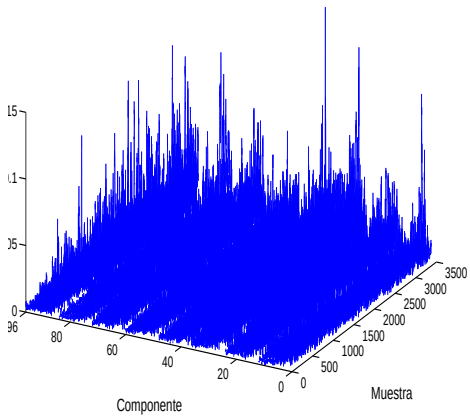
Uno de los aspectos que se destacan en el estudio de las señales EEG es la alta variabilidad de las mismas si tenemos en cuenta los diferentes usuarios e incluso, para un mismo usuario, cuando se consideran diferentes sesiones. Esta variabilidad ha inducido a pensar en la necesidad de un aprendizaje on-line, al principio de la sesión, que permita reajustar los clasificadores [21].

Para el problema del diseño de interfaces BCI, en las soluciones que se han venido aportando se observa la diferencia de prestaciones entre diferentes usuarios, lo que lleva a plantear la necesidad de realizar un reaprendizaje on-line con el objetivo de mejorar el funcionamiento para el caso de ciertos usuarios. En la Figura 4 se muestran los datos pre-tratados mediante el cálculo de power spectrum density para una de las sesiones de uno de los usuarios participantes en el benckmark BCI Competition III Data Set V. Se han utilizado datos de 8 canales EEG y para cada uno de ellos se ha estimado la PSD 16 veces por segundo para 12 frecuencias en el rango: 8-30 Hz.. Así mismo se pueden ver las etiquetas correspondientes a cada una de las tareas que se han considerado (3=imaginar mover la mano izquierda, 2=imaginar mover la mano derecha, 7=pensar en una palabra). Se trata, en resumen, de un problema de clasificación de vectores de dimensión 96.

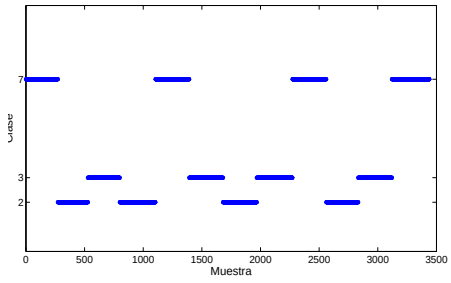
Para los experimentos que se van a describir, los parámetros utilizados en S-dFasArt van a ser siempre los mismos, con el objeto de poder comparar resultados. $h = 1., \delta = 0{,}17, A_w = 0{,}8, A_v = 0{,}8, A_c = 0{,}1, A_r = 0{,}05, A_t = 0{,}1, RESET = 0{,}1$. Se dispone de tres sesiones etiquetadas (S1,S2,S3) de datos similares a la mostrada en la Figura 4. En un primer experimento se propone utilizar una sesión para el aprendizaje y otra para el test. La tabla 1 recoge los resultados (en porcentaje de acierto) que se obtienen cuando se consideran la sesión de test dividida en dos tramos iguales (T1,T2) o la sesión de test completa (T1+T2). El valor de N se corresponde con el número de clases/reglas que se han generado.

Una posible alternativa para mejorar los porcentajes de acierto sería utilizar parte de la sesión de test como sesión de aprendizaje. En la tabla 2 se muestran los resultados de esta opción. Se ha utilizado la mitad de los datos de la sesión como conjunto de aprendizaje y la otra mitad como conjunto de test.

SdFasArt permite que los ciclos de aprendizaje y test se puedan ir alternando, es decir hacer un aprendizaje on-line.



(a)



(b)

Figura 4: (a) Vectores (96-dimensiones) PSD: 8 canales, 12 frecuencias. (b) Clases 2=iquierda,3=derecha,7=palabra

Se ha planteado un experimento en el que para la primera mitad de la sesión de test (T1), después de la predicción, se realimenta al sistema cual sería la opción correcta, y se realiza un ciclo de aprendizaje. Los resultados se recogen en la tabla 3. N1 es el número de reglas que se han generado despues del aprendizaje de la primera sesión y N2 el numero de reglas final.

Analizando los valores de la tabla 3 destaca que en el tramo T1, donde se realiza el aprendizaje on-line, los valores de acierto son muy altos. Esto se explica por la forma de las clases (Figura 4(b)). Una vez detectado un cambio en la clase, ya sea de forma autónoma o gracias a la supervisión, el carácter dinámico de S-dFasArt hace que la regla ganadora se refuerce y tienda a mantener la predicción. Así mismo se percibe una mejora en los resultados de la predicción en el tramo T2, con respecto a los obtenidos en los dos casos anteriores (Tablas 1,2).

Para poder hacer una comparación, en la tabla 4 se han promediado los resultados obtenidos para los tramos T2 de las sesiones. Se observa como el aprendizaje on-line permite añadir nueva información al modelo que mejora las prestaciones del mismo, manteniendo la información ya aprendida.

Cuadro 1: Aprendizaje con una sesión y test con otra.

Aprendizaje/test	N	T1	T2	T1+T2
S1/S2	123	40.65	50.47	45.56
S1/S3	123	36.92	35.87	36.40
S2/S1	122	40.65	41,24	40.95
S2/S3	122	34.71	33.84	34.27
S3/S1	123	40.77	38.61	39.69
S3/S2	123	40.77	42.52	41.65

Cuadro 2: Aprendizaje con media sesión y test con el resto.

Aprendizaje/test	N	T2
S1(T1)/S1(T2)	62	35.98
S2(T1)/S2(T2)	62	39.60
S2(T1)/S2(T2)	62	35.52

Cuadro 3: Aprendizaje con una sesión, aprendizaje on line en media sesión y test con la otra mitad.

Aprendizaje/test	N1	N2	T1	T2
S1+S2(T1)/S2(T2)	123	185	99.59	51.52
S1+S3(T1)/S3(T2)	123	185	99.59	37.03
S2+S1(T1)/S1(T2)	122	184	99.59	45.27
S2+S3(T1)/S3(T2)	122	184	99.59	36.92
S3+S1(T1)/S1(T2)	123	185	99.65	41.00
S3+S2(T1)/S2(T2)	123	185	99.47	43.69

Cuadro 4: Resumen de resultados.

Test SX	Aprendizaje SY	Aprendizaje SX(T1)	Aprendizaje SY+SX(T1)
S1(T2)	40.32	35.98	**43.14**
S2(T2)	43.61	39.60	**47.61**
S2(T3)	35.34	35.52	**36.98**

5. Conclusiones

En el presente trabajo se han descrito los diferentes modelos que se han propuesto a partir del modelo FasArt. Estos modelos parten del objetivo de aunar los sistemas lógico difusos con las arquitecturas neuronales de tipo ART. Para ello se establece una equivalencia: funciones de pertenencia-activación, conexiones Inter-ART-reglas. Este hecho permite asumir que el modelo FasArt se puede interpretar como un sistema lógico difuso y, por tanto, aplicar resultados en tareas de representación de funciones, recogidos en la literatura.

La incorporación de una formulación dinámica en las ecuaciones de FasArt conducen al modelo dFasArt que permite tratar entradas teniendo en cuenta el carácter temporal de las mismas.

Los modelos propuestos se han aplicado en diferentes problemas asociados al reconocimiento de patrones, ya sea de forma supervisada o no supervisada, y a la identificación y control de sistemas dinámicos. Las referencias aportadas muestran un amplio abanico de aplicaciones exitosas, incluyendo varias en las que se ha trabajado con datos provenientes de la industria o se han comparado los resultados con los que se obtienen con otros modelos en benchkmarks internacionales.

Agradecimientos

En recuerdo de nuestro amigo y compañero Dr. Miguel Pinzolas y de lo que su presencia ha aportado a nuestras vidas.

Referencias

[1] M. Araúzo, J. Cano-Izquierdo, E. Gómez, López M., Y. Dimitriadis, and J. Coronado. Automatization of a penicillin production process with soft sensors and a adaptive controller based on neuro fuzzy systems. *Control Engineering Practice*, 12:1073–1090, 2004.

[2] J. Balanza, J. Cano-Izquierdo, J. Coronado, A. Alvarez, and M. Luna. Estudio y desarrollo de sensores software basados en sistemas neuro-difusos. aplicación en procesos petroquímicos. In *Actas de las XXI Jornadas de Automática*, Sevilla, España, 2000.

[3] B. Blankertz, K. Muller, D. Krusienski, G. Schalk, Wolpaw J., A. Schlögl, G. Pfussrtscheller, J. Millan, M. Schröder, and N. Birbaumer. The BCI competition III: Validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):153–159, 2006.

[4] J. Cano, M. Arauzo, Y. Dimitriadis, and J. Coronado. Non linear systems identification using fuzzy neural networks. In *Proceedings of EUFIT'97*, Aachen, Deuchland, 1997.

[5] J. Cano, Y. Dimitriadis, M. Arauzo, F. Abajo, and J. Coronado. A neuro-fuzzy architecture for automatic development of fuzzy controllers. In *Proceedings of CESA'96*, Lille, France, 1996.

[6] J. Cano, Y. Dimitriadis, M. Arauzo, and J. Coronado. FasArt: A new neuro-fuzzy architecture for incremental learning in systems identification. In *Proceedings of IFAC'96*, San Francisco, USA, 1996.

[7] J. Cano, Y. Dimitriadis, and J. Coronado. FasBack : Matching-error based learning for automatic generation of fuzzy logic systems. In *Proceedings of FUZZ-IEEE'97*, Barcelona, Spain, 1997.

[8] J. Cano, Y. Dimitriadis, E. Gómez, and J. Coronado. Learning from noisy information in FasArt

and FasBack neuro-fuzzy systems. *Neural Networks*, 14:407–425, 2001.

[9] J. Cano-Izquierdo, M. Almonacid, J. Ibarrola, and M. Pinzolas. Use of dynamic neuro fuzzy model dFasArt for identification of stationary states in closed-loop controlled systems. In *Proceedings of EUROFUSE 2007*, Jaen, Spain, 2007.

[10] J. Cano-Izquierdo, M. Almonacid, M. Pinzolas, and J. Ibarrola. dFasArt: Dynamic neural processing in FasArt model. *Neural Networks*, 22:479–487, 2009.

[11] J. Cano-Izquierdo, J. Ibarrola, and M. Almonacid. Performance monitoring of closed-loop controlled systems using dFasArt. In *Proceedings of IWINAC'07*, La Manga del Mar Menor, Spain, 2007.

[12] J. Cano-Izquierdo, J. Ibarrola, and M. Almonacid. Improving motor imagery classification with a new BCI design using neuro-fuzzy S-dFasArt. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, doi:10.1109/TNSRE.2011.2169991, 2011.

[13] J. Cano-Izquierdo, M. Pinzolas, J. Ibarrola, and J. Coronado. Identificación de funciones utilizando sistemas lógicos difusos. In *Actas de las XXI Jornadas de Automática*, Sevilla, España, 2000.

[14] G. Carpenter, S. Grossberg, and J. Reynolds. A Fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems. *IEEE Transactions on Neural Networks*, 6:1330–1336, 1995.

[15] G. Carpenter, S. Grossberg, and D. Rosen. Fuzzy ART: Fast stable learning and categorization of analog patterns by a adaptive resonance system. *Neural Networks*, 4:759–771, 1991.

[16] J. Castro and M. Delgado. Fuzzy systems with defuzzification are universal approximators. *IEEE Transaction on Systems, Man, and Cybernetics-Part B: Cybernetics*, 26(1):149–152, February 1996. 4.120.

[17] E. Gómez, Y. Dimitriadis, M. Sánchez, P. Sánchez, J. Cano, and J. Coronado. On-line UNIPEN character recognition with fuzzy neural networks. *Intelligent Automation and Soft Computing*, 7(3):163–175, 2001.

[18] E. Gómez, J. Gago, Y. Dimitriadis, J. Cano, and J. López. Experimental study of a novel neuro-fuzzy system for on-line handwritten UNIPEN digit recognition. *Pattern Recognition Letters*, 19:357–364, 1998.

[19] Cano-Izquierdo J, Almonacid M., and Ibarrola J. Applying neuro-fuzzy model dFasArt in control systems. *Engineering Applications of Artificial Intelligence*, 23:1053–1063, 2010.

[20] García-Cascales J., Vera-García F., J. Cano-Izquierdo, J. Delgado-Marín, and Martínez-Sánchez R. Modelling an absorption system assisted by solar energy. *Applied Thermal Engineering*, 31:112–118, 2011.

[21] J. Millan. On the need for on-line learning in brain-computer interfaces. In *Proceedings of International Joint Conference on Neural Networks*, Budapest, Hungary, 2004.

[22] Toledo-Moreo R., Pinzolas M., and Cano-Izquierdo J. Maneuver prediction for road vehicles based on a neuro-fuzzy architecture with a low-cost navigation unit. *IEEE Transactions on Intelligent Transportation Systems*, 11(2):498–504, 2010.

[23] G. Sainz, J. Cano, Y. Dimitriadis, and J. López. A new neuro-fuzzy system for logical labelling of documents. In *13th International Conference on Pattern Recognition, IAPRO*, Viena, Austria, 1996.

[24] G. Sainz, Y. Dimitriadis, J. Cano, E. Gómez, and E. Parrado. *Neuro-Fuzzy Pattern Recognition*, chapter Neural ART-Based Model Set for Pattern Recognition: FasArt Family, pages 145–175. World Scientific, 2000.

[25] G. Sainz, Y. Dimitriadis, R. Sanz, and J. Cano. Neuro-fuzzy ART based document management system: Application to mail distribution and digital libraries. *Engineering Applications of Artificial Intelligence*, 15:17–29, 2002.

[26] G. Sainz-Palmero, J. Juez-Santamaria, E. Moya, and J. Perán-González. Fault detection and fuzzy rule extraction in AC motors by a neuro-fuzzy ART-based system. *Engineering Applications of Artificial Intelligence*, 18:867–874, 2005.

[27] A. Toledo, R. Toledo-Moreo, J. Cano-Izquierdo, and M. Pinzolas-Prado. Maneuver prediction for road vehicles based on a novel neuro-fuzzy dynamic architecture. *Robotics and Autonomous Systems*, 58:1316–1320, 2010.

[28] R. Toledo, M. Pinzolas, and J. Cano-Izquierdo. Supervised dFasArt: a neuro-fuzzy dynamic architecture for maneuver detection in road v vehicle collision avoidance support systems. In *Proceedings of IWINAC'07*, La Manga del Mar Menor, Spain, 2007.

ANÁLISIS DE EFICIENCIA CON DATOS INCIERTOS. UNA APLICACIÓN A LA INDUSTRIA TEXTIL

Leonor Plá[1], Trinidad Casasús[2], Juan Carlos Pérez[3], Vicente Liern[2]

[1]*Dep. Matemática Aplicada, Universidad Politécnica de Valencia*, mlpla@mat.upv.es

[2]*Dep. Matemáticas para la Economía y la Empresa, Universidad de Valencia*, trinidad.casasus@uv.es, vicente.liern@uv.es

[3]*IES Andreu Sempere, Alcoy*, jcpc4655@telefonica.net

Resumen

En este trabajo presentamos modelos para analizar la eficiencia de un conjunto de Unidades de Toma de Decisión, DMUs, y la ordenación de éstas teniendo en cuenta su nivel de eficiencia. Normalmente, los datos que se manejan para el análisis suelen estar afectados de incertidumbre que difícilmente admiten un tratamiento aleatorio. En este contexto el uso de modelos de análisis de eficiencia fuzzy, en concreto Data Envelopment Analysis Fuzzy Models, resultan muy útiles. Para ilustrar las propuestas aplicamos los modelos a la evaluación y clasificación a un grupo de empresas textiles de las zonas de Alcoi, Cocentaina y Muro en Alicante.

Palabras Clave: Fuzzy Mathematical Programming, Fuzzy DEA Models, Ranking Fuzzy Numbers.

1 INTRODUCCIÓN

La actividad textil es una de las más antiguas de la Comunidad Valenciana. En las comarcas de L'Alcoià y el Comtat se concentra una actividad elevada dedicada al sector textil-hogar. En particular, en este estudio, nos centramos en los municipios de Alcoi, Cocentaina y Muro siendo el tipo pequeñas y medianas empresas (PYMEs).

En primer lugar, la evaluación de la eficiencia ha adquirido en los últimos años un gran interés en el ámbito empresarial, entendiendo por eficiencia el uso más adecuado de los recursos disponibles para obtener la mayor cantidad de productos o beneficios posibles. En este sentido, la eficiencia está ligada al aprovechamiento óptimo de los recursos.

Entre los métodos más utilizados para la medición de la eficiencia se encuentra el Análisis Envolvente de Datos DEA, que surge como una extensión del trabajo de Farrell [7] que proporciona una medida satisfactoria de la eficiencia productiva [3].

En segundo lugar, la competitividad de las empresas, entendida como que el coste final del producto coincida con la voluntad de pago del potencial cliente [3]; si bien no es sinónimo de eficiencia al menos ésta se erige en un determinante directo de su nivel de competitividad [6 13], en un elemento de especial importancia para el análisis del posicionamiento competitivo de las empresas así como un método que fija de modo esencial sus estrategias [5].

La utilización eficiente de los recursos productivos representa una estrategia que permite a la empresa mejorar su rentabilidad. Así pues, no cabe separar completamente la eficiencia y el conjunto de acciones posibles para ser competitivos [6]; al contrario, su presencia facilita la competitividad y, por ello, la existencia de competidores induce a la búsqueda de la eficiencia [12]; aunque también es cierto que la eficiencia no agota el conjunto de estrategias para el logro de la competitividad.

Por otro lado, las técnicas tradicionales de análisis de eficiencia asumen que los planes de producción se conocen con precisión y se desarrollan bajo la hipótesis de que los valores de los inputs y los outputs son determinados con certeza. Sin embargo en el mundo real muchos procesos de producción involucran a menudo inputs y outputs que son muy difíciles de medir de una forma precisa, lo que sugiere que la disponibilidad de una metodología que permita el tratamiento de datos imprecisos podría resultar de gran utilidad, y no sólo en este contexto textil.

El tratamiento de la incertidumbre en los modelos DEA puede hacerse desde planteamientos estocásticos o bien desde planteamientos fuzzy **(**[8], [9]**)**. El planteamiento estocástico [11] requiere conocer o suponer y justificar alguna distribución de probabilidad sobre el error cometido en los datos o las ineficiencias. El planteamiento fuzzy es el que abordamos en nuestro trabajo porque no requiere

ninguno de estos condicionamientos y la metodología es fácilmente ampliable a diferentes y variadas situaciones reales. Pensamos que puede ser de gran utilidad práctica.

Sin embargo rara vez es tenido en cuenta que, para tomar decisiones, puede no ser suficiente conocer simplemente si la empresa usa los recursos de manera más eficiente, por ello nos proponemos además establecer una clasificación de las empresas a partir de la ordenación fuzzy de sus eficiencias y que puede utilizarse como una primera aproximación al estudio de la competitividad relativa entre ellas.

2 METODOLOGIA DEA

2.1. MODELOS DEA

El modelo matemático básico formulado por Charnes, Cooper y Rhodes (CCR) [2] adopta la versión fraccional,

$$(MF)\ h_{j_0} = \text{Max} \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}}$$
$$s.a. \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad j = 1,2,...,n, \tag{1}$$
$$u_r \ge 0,\ r = 1,2,\cdots,s,$$
$$v_i \ge 0,\ i = 1,2,\cdots,m.$$

La unidad productiva cuya eficiencia se calcula se denota por el subíndice j_0, las variables y_{rj} y x_{ij} representan, respectivamente, las cantidades de output r y de input i de la entidad j. Finalmente u_r y v_i representan las ponderaciones o pesos atribuidas al output r y al input i respectivamente. La resolución de este programa permite determinar los valores de las variables u_r y v_i y, por tanto, el índice de eficiencia h_{j_0} asignado a la unidad productiva evaluada.

En la práctica, se suele utilizar el modelo dual de (MF) porque generalmente el número de restricciones es menor que en el caso primal y mucho menor que el número de DMUs. Este modelo llamado envolvente con orientación input es:

$$E_{j_0} = Min\ \theta_{j_0}$$
$$s.a. \quad \sum_{j=1}^{n} \lambda_j x_{ij} \le \theta_{j_0} x_{ij_0},\ i = 1,2,\cdots,m, \tag{2}$$
$$\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{rj_0},\ r = 1,2,\cdots,s,$$
$$\lambda_j \ge 0,\ \ j = 1,2,\cdots,n.$$

2.2. MODELOS DEA CON DATOS IMPRECISOS

A pesar de que un proceso de producción, o un mecanismo de servicio, por lo general consiste en manipular inputs a fin de obtener outputs, no significa que la medición sea sencilla. En los últimos años han aparecido algunos modelos que permiten calcular la eficiencia teniendo en cuenta la incertidumbre.

El punto de vista que vamos a desarrollar es el de Kao y Liu [9] pero en la versión dual del problema, porque es la que más se utiliza en Economía. En un conjunto de n DMUS, supongamos que los m inputs y s outputs vienen determinados por los números fuzzy $\widetilde{X}_{ij}$ e $\widetilde{Y}_{rj}$ cuyas funciones de pertenencia son respectivamente $\mu_{\widetilde{X}_{ij}}$ y $\mu_{\widetilde{Y}_{rj}}$.

Los α-cortes de estos números fuzzy pueden expresarse mediante intervalos del siguiente modo

$$\widetilde{x}_{ij} = \left[\min_{x_{ij}} \left\{ x_{ij} \in \widetilde{X}_{ij} \middle| \mu_{\widetilde{X}_{ij}}(x_{ij}) \ge \alpha \right\}, \max_{x_{ij}} \left\{ x_{ij} \in \widetilde{X}_{ij} \middle| \mu_{\widetilde{X}_{ij}}(x_{ij}) \ge \alpha \right\} \right] \tag{3}$$
$$= \left[x_{ij}(\alpha)^L, x_{ij}(\alpha)^U \right],\ 1 \le i \le m, 1 \le j \le n,$$
$$\widetilde{y}_{rj} = \left[\min_{y_{rj}} \left\{ y_{rj} \in \widetilde{Y}_{rj} \middle| \mu_{\widetilde{Y}_{rj}}(y_{rj}) \ge \alpha \right\}, \max_{y_{rj}} \left\{ y_{rj} \in \widetilde{Y}_{rj} \middle| \mu_{\widetilde{Y}_{rj}}(x_{ij}) \ge \alpha \right\} \right] \tag{4}$$
$$= \left[y_{rj}(\alpha)^L, y_{rj}(\alpha)^U \right],\ 1 \le r \le s,\ 1 \le j \le n.$$

Para incorporar los valores expresados en (3) y (4) al modelo (2), Kao y Liu proponen el modelo lineal fuzzy siguiente:

$$\widetilde{E}_{j_0} = Min\ \theta_{j_0}$$
$$s.a. \quad \sum_{j=1}^{n} \lambda_j \widetilde{x}_{ij} \le \theta_{j_0} \widetilde{x}_{ij_0},\ i = 1,2,\cdots,m, \tag{5}$$
$$\sum_{j=1}^{n} \lambda_j \widetilde{y}_{rj} \gtrsim y_{rj_0},\ r = 1,2,\cdots,s,$$
$$\lambda_j \ge 0,\ \ j = 1,2,\cdots,n.$$

NOTA: En todo nuestro trabajo podemos suponer, sin pérdida de generalidad, que todas las observaciones son fuzzy, ya que los valores crisp se representan por funciones de pertenencia degeneradas.

Basado en el principio de extensión de Zadeh ([17]) la función de pertenencia de la k-ésima DMU puede obtenerse como

$$\mu_{\widetilde{E}_k}(z) = \sup_{x_{ij}, y_{rj}} \min \left\{ \mu_{\widetilde{X}_{ij}}(x_{ij}), \mu_{\widetilde{Y}_{rj}}(y_{rj}), \forall i,j,r \middle| z = E_k(x_{ij}, y_{rj}) \right\}, \tag{6}$$

donde E_k se calcula como en (2).

Desde el punto de vista práctico, Kao y Liu [9] proponen describir la función de pertenencia dada en (6) mediante intervalos para cada nivel de pertenencia $\alpha \in [0,1]$,

$$E_{j_0}(\alpha) = \left[E_{j_0}^L(\alpha), E_{j_0}^U(\alpha) \right], \tag{7}$$

que se obtienen mediante los dos programas lineales siguientes:

$$
\begin{aligned}
& E^L_{j_0}(\alpha) = Min\ \theta_{j_0} \\
& s.a. \quad \sum_{j=1,\ j\neq j_0}^{n} \lambda_j x^L_{ij}(\alpha) + \lambda_{j_0} x^U_{ij_0}(\alpha) \le \theta_{j_0} x^U_{ij_0},\ i = 1,2,\cdots,m, \\
& \quad \sum_{j=1,\ j\neq j_0}^{n} \lambda_j y^U_{rj}(\alpha) + \lambda_{j_0} y^L_{rj_0}(\alpha) \ge y^L_{rj_0},\ r = 1,2,\cdots,s, \\
& \quad \sum_{j=1}^{n} \lambda_j = 1, \\
& \quad \lambda_j \ge 0, \quad j = 1,2,...,n,
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
& E^U_{j_0}(\alpha) = Min\ \theta_{j_0} \\
& s.a. \quad \sum_{j=1,\ j\neq j_0}^{n} \lambda_j x^U_{ij}(\alpha) + \lambda_{j_0} x^L_{ij_0}(\alpha) \le \theta_{j_0} x^L_{ij_0},\ i = 1,2,\cdots,m, \\
& \quad \sum_{j=1,\ j\neq j_0}^{n} \lambda_j y^L_{rj}(\alpha) + \lambda_{j_0} y^U_{rj_0}(\alpha) \ge y^U_{rj_0},\ r = 1,2,\cdots,s, \\
& \quad \sum_{j=1}^{n} \lambda_j = 1, \\
& \quad \lambda_j \ge 0, \quad j = 1,2,...,n.
\end{aligned}
\tag{9}
$$

Los orígenes de los intervalos expresados en (7) son las soluciones del modelo (8), llamado caso desfavorable (mayores inputs y menores outputs para la DMU j_0 y lo contrario para las otras) y los extremos de los intervalos son las soluciones del modelo (9), caso favorable (menores inputs y mayores outputs para la DMU j_0 y lo contrario para las otras).

En este contexto ordenar las DMUs basándonos en la eficiencia consiste en ordenar los números fuzzy $\widetilde{E}_j$ que poseen como soporte los intervalos $E_{j_0}(\alpha)$.

Para obtener una ordenación de las DMU en función de su eficiencia fuzzy existen varias propuestas, de hecho Kerre ([14], [15]) recoge, clasifica y relaciona algunos de estos métodos. En la subsección siguiente mostraremos algunos métodos basados en α-cortes.

2.2.1. Algunos procedimientos de clasificación

(1) *Aproximación de Yager*. R.R. Yager [16] propuso cuatro índices para clasificar cantidades borrosas en [0, 1], del que probablemente el más utilizado e intuitivo sea

$$
Y\left(\widetilde{E}_j\right) = \int_0^{\alpha_{max}} \tfrac{1}{2}\left(E^L_j(\alpha) + E^U_j(\alpha)\right) d\alpha. \tag{10}
$$

(2) *Aproximación de Chen y Klein* [4]. Estos autores consideran el índice

$$
CK\left(\widetilde{E}_j\right) = \frac{\sum_{l=0}^{N}\left(E^U_j(\alpha_l) - c\right)}{\sum_{l=0}^{N}\left(E^U_j(\alpha_l) - c\right) - \sum_{l=0}^{N}\left(E^L_j(\alpha_l) - d\right)} \tag{11}
$$

donde $c = \min_{ij}\left\{E^L_{ij}(\alpha_l)\right\}$ y $d = \max_{ij}\left\{E^U_{ij}(\alpha_l)\right\}$.

(3) *Distancia de Hamming* **[**10**]**

Si dos borrosos coinciden la distancia de Hamming d_H vale 0 por lo que tomamos como medida el valor $1 - d_H$:

$$
H\left(\widetilde{E}_j\right) = 1 - \left(\sum_{l=0}^{N}\left(\left|E^U_j(\alpha)\right| - 1\right) + \sum_{l=0}^{N}\left(\left|E^L_j(\alpha)\right| - 1\right)\right) \tag{12}
$$

Los tres índices anteriores y otros muchos, que representaremos genéricamente por Ψ, permiten establecer una ordenación de la forma siguiente [1, 14, 15]:

Definición 1: *El conjunto de números fuzzy* $\left\{\widetilde{E}_j\right\}_{j=1}^{n}$ *se puede ordenar como:*

$$
\begin{aligned}
& \widetilde{E}_j \succ \widetilde{E}_k \text{ para } \Psi \ \Leftrightarrow\ \Psi_j > \Psi_k \\
& \widetilde{E}_j \approx \widetilde{E}_k \text{ para } \Psi \ \Leftrightarrow\ \Psi_j = \Psi_k \\
& \widetilde{E}_j \succeq \widetilde{E}_k \text{ para } \Psi \ \Leftrightarrow\ \Psi_j \ge \Psi_k
\end{aligned}
\tag{13}
$$

(4) *Método estadístico de Boscá et al.* [1].

Tomamos *n* DMUs cuyos inputs y outputs se expresan por los números LR-fuzzy

$$
\widetilde{X}_{ij} = \left\langle x^L_{ij}, x^U_{ij}, \gamma^L_{ij}, \gamma^U_{ij} \right\rangle, 1 \le i \le m,\ 1 \le j \le n \tag{14}
$$

$$
\widetilde{Y}_{rj} = \left\langle y^L_{rj}, y^U_{rij}, \delta^L_{rj}, \delta^U_{rj} \right\rangle, 1 \le r \le s,\ 1 \le j \le n \tag{15}
$$

El procedimiento para alcanzar la clasificación consiste en seleccionar números crisp de los inputs y outputs LR-fuzzy, el algoritmo es el siguiente.

PASO 1: Obtención números crisp:

$$
\begin{aligned}
& x^t_{ij} := x^L_{ij} - \gamma^L_{ij}, x^A_{ij} := \frac{x^L_{ij} + x^U_{ij}}{2}, x^T_{ij} := x^U_{ij} - \gamma^U_{ij}, \\
& y^t_{rj} := y^L_{rj} - \delta^L_{rj}, y^A_{rj} := \frac{y^L_{rj} + y^U_{rj}}{2}, y^T_{rj} := y^U_{rj} - \delta^U_{rj}.
\end{aligned}
\tag{16}
$$

PASO 2: Para analizar la eficiencia de la DMU j_0 se resuelve el siguiente modelo DEA

$$
\begin{aligned}
& E\left(x^a_{ij}, x^b_{ij_0}, y^c_{rj}, y^d_{rj_0}\right) = Max\ \eta_0 \\
& s.a. \quad \sum_{j=1}^{n} \lambda_j x^a_{ij} \le x^b_{i0},\ i = 1,2,\cdots,m, \\
& \quad \sum_{j=1}^{n} \lambda_j y^c_{rj} \ge \eta_0 y^d_{rj_0},\ r = 1,2,\cdots,s, \\
& \quad \lambda_j \ge 0, \quad j = 1,2,\cdots,n,
\end{aligned}
\tag{17}
$$

donde $a,b,c,d \in \{t, A, T\}$.

PASO 3: Construcción del conjunto

$$
\Gamma_{j_0} = \left\{E\left(x^a_{ij}, x^b_{ij_0}, y^c_{rj}, y^d_{rj_0}\right) \middle| a,b,c,d \in \{t,A,T\}\right\}. \tag{18}
$$

PASO 4: Cálculo de los siguientes ratios para la j_0-unidad:

$$
R^1_{ij_0} = \frac{e_{j_0}}{card\left(\Gamma_{j_0}\right)},\quad R^2_{ij_0} = \begin{cases} \dfrac{S_{j_0} - e_{j_0}}{card\left(\Gamma_{j_0}\right) - e_{j_0}}, & card\left(\Gamma_{j_0}\right) \neq e_{j_0} \\ 0, & card\left(\Gamma_{j_0}\right) = e_{j_0} \end{cases} \tag{19}
$$

donde

$$S_{j_0} = \sum_{a,b,c,d} E\left(x_{ij}^a, x_{ij_0}^b, y_{rj}^c, y_{rj_0}^d\right),$$

e_{j_0} = número de veces que la j_0-ésima DMU es eficiente en (17).

$card\left(\Gamma_{j_0}\right)$ = cardinal de Γ_{j_0} .

Con este método, Boscá et al. [1] establecen la siguiente ordenación:

Definición 2: *El conjunto de números fuzzy* $\left\{\widetilde{E}_j\right\}_{j=1}^n$ *se puede ordenar como sigue:*

$$\begin{aligned}
&\widetilde{E}_j \succ \widetilde{E}_k \;\Leftrightarrow\; R_j^1 > R_k^1 \text{ o bien } [R_j^1 = R_k^1 \text{ y } R_j^2 > R_k^2],\\
&\widetilde{E}_j \approx \widetilde{E}_k \;\Leftrightarrow\; R_j^1 = R_k^1 \text{ y } R_j^2 = R_k^2, \qquad (20)\\
&\widetilde{E}_j \succeq \widetilde{E}_k \;\Leftrightarrow\; R_j^1 \geq R_k^1 \text{ o bien } R_j^2 \geq R_k^2.
\end{aligned}$$

3 APLICACIÓN NUMÉRICA

El análisis de la eficiencia supone centrar la atención en la tecnología existente, los recursos y los precios de ellos. La clave consiste en aprovechar al máximo los recursos y hacerlo adaptándose a los precios quien lo consiga será eficiente; quien no, incurre en ineficiencias que le suponen un deterioro para competir y, en cualquier caso, soportará un coste innecesario por tal desaprovechamiento [3].

En este estudio nos proponemos aplicar la metodología DEA a empresas del sector textil-hogar para posteriormente incorporar, como un elemento más del problema, la incertidumbre que aparece de forma natural. Existe el hecho de que a menudo los inputs y outputs involucrados en las aplicaciones reales no pueden ser medidos de forma precisa. En este sentido cabe preguntarse si nuestros datos son imprecisos, ¿se verán afectados los resultados de la medición de la eficiencia? Vamos a realizar un tratamiento de la incertidumbre desde una perspectiva fuzzy.

Los datos utilizados se refieren al año 2009 y han sido extraídos de la base de datos Sabi, elaborada por Bureau Van Dijk, de los municipios Alcoi, Cocentaina y Muro cuya actividad se encuadra en alguno de los grupos en que la clasificación CNAE-93 divide el sector textil. Del total de empresas se han escogido aquellas con un número de empleados superior a 9 y que estén activas por lo menos hasta acabar el año 2009. En total se eligieron 54 empresas.

Antes de proceder a la evaluación de eficiencia del conjunto de empresas es necesario seleccionar las variables input y output a partir de las cuales se definirá la eficiencia y que permitirán construir el modelo DEA.

Se han considerado cuatro inputs y dos outputs que reflejan las características de las empresas.

INPUTS:

I1. Existencias: Son los bienes de propiedad de la empresa destinados a la venta o a su transformación e incorporación al proceso productivo.

I2. Número de empleados: Conjunto de personas, fijas o eventuales, que en el año de referencia se encuentran trabajando en la empresa.

I3. Gastos de personal: Retribuciones al personal, cualquiera que sea la forma o el concepto por el que se satisfacen; cuotas a la Seguridad Social a cargo de la empresa y los demás gastos de carácter social.

I4. Dotaciones para amortización de inmovilizado: Expresión de la depreciación sistemática anual efectiva sufrida por el inmovilizado intangible y material, por su aplicación al proceso productivo.

OUTPUTS:

O1. Importe neto de cifras de ventas: La renta que una empresa recibe de sus actividades comerciales normales, por lo general de la venta de bienes y servicios a los clientes.

O2. Ingresos de explotación: Ingresos obtenidos por la empresa al realizar su actividad típica.

Supondremos que todas las observaciones son fuzzy. Para cada observación, tanto inputs como outputs, hemos construido un número fuzzy cuyo soporte es un intervalo con desviaciones variables de hasta un 10% del valor observado, y elegidas estas desviaciones de un modo arbitrario para cada observación.

A excepción de la variable referida al número de empleados, las restantes se encuentran expresadas en miles de euros. Para el cálculo de índices de eficiencia se ha utilizado el programa GAMS.

Los datos del año 2009 correspondientes a los inputs y outputs de 12 de las 54 DMUs se detallan en la Tabla 1, con notación desde K1 hasta K12.

Tabla1: Relación datos de inputs y outputs de las empresas

DMU	I1	I2	I3	I4	O1	O2
K1	4 250	149	3 929	1 549	30 665	31 372
K2	620	23	674	156	1 571	1 601
K3	39	14	336	100	880	946
K4	234	34	866	487	4 863	4 894
K5	8 284	143	2 778	1 811	15 047	15 489
K6	562	44	1 140	205	2046	2 257
K7	3 903	56	1 897	651	11 307	11 580
K8	1 770	34	1 137	388	5 902	5 930
K9	2	27	1 023	8	1 869	1 875
K10	208	31	853	210	1 911	2 027
K11	222	49	1 430	177	2 839	2 839
K12	1 736	131	5 074	1 945	11 994	12 092

Aplicando la metodología DEA con datos imprecisos del epígrafe 2.2 obtenemos los intervalos de eficiencia $E_k(\alpha)$ de Kao y Liu [9] expuestos en la Tabla 2 para 12 de las 54 empresas.

Tabla 2: Intervalos de eficiencia para ciertos valores de α

α	K1	K2	K3	K4	K5	K6
0	[0.696, 1]	[0.231, 0.415]	[0.603, 1]	[0.821, 1]	[0.253, 0.454]	[0.231, 0.422]
0.2	[0.724, 0.998]	[0.245, 0.392]	[0.640, 1]	[0.855, 1]	[0.268, 0.429]	[0.246, 0.397]
0.4	[0.754, 0.958]	[0.260, 0.370]	[0.680, 0.975]	[0.890, 1]	[0.284, 0.405]	[0.261, 0.374]
0.6	[0.785, 0.921]	[0.275, 0.349]	[0.722, 0.918]	[0.927, 1]	[0.301, 0.382]	[0.277, 0.352]
0.8	[0.817, 0.885]	[0.292, 0.329]	[0.767, 0.865]	[0.964, 1]	[0.319, 0.360]	[0.294, 0.332]
1	[0.850, 0.850]	[0.310, 0.310]	[0.814, 0.814]	[1, 1]	[0.339, 0.339]	[0.312, 0.312]
α	K7	K8	K9	K10	K11	K12
0	[0.460, 0.687]	[0.392, 0.705]	[0.740, 1]	[0.412, 0.615]	[0.417, 0.750]	[0.330, 0.601]
0.2	[0.479, 0.660]	[0.416, 0.666]	[0.789, 1]	[0.429, 0.591]	[0.442, 0.708]	[0.350, 0.566]
0.4	[0.499, 0.634]	[0.440, 0.628]	[0.835, 1]	[0.447, 0.568]	[0.468, 0.668]	[0.372, 0.533]
0.6	[0.519, 0.609]	[0.467, 0.592]	[0.887, 1]	[0.465, 0.545]	[0.497, 0.630]	[0.395, 0.502]
0.8	[0.540, 0.585]	[0.495, 0.558]	[0.942, 1]	[0.484, 0.524]	[0.527, 0.594]	[0.419, 0.473]
1	[0.562, 0.562]	[0.526, 0.526]	[1, 1]	[0.504, 0.504]	[0.559, 0.559]	[0.445, 0.445]

Los resultados obtenidos al aplicar a los intervalos de eficiencia los 4 procedimientos de clasificación expuestos en el epígrafe 2.2.1 vienen expuestos en la Tabla 3.

Tabla 3: Resultados de los procedimientos

	Yager	Chen&Klein	Boscá et. al	1-Hamming
K1	0.8531	0.7903	0.8531	0.0833//0.8397
K2	0.3148	0.2831	0.3148	0//0.3148
K3	0.8166	0.7435	0.8166	0//0.8166
K4	0.9549	0.9113	0.9549	0.5833//0.8917
K5	0.3442	0.3136	0.3442	0//0.3442
K6	0.3175	0.2863	0.3175	0//0.3175
K7	0.5664	0.5292	0.5664	0//0.5664
K8	0.5342	0.4984	0.5342	0//0.5342
K9	0.9325	0.8731	0.9325	0.5833//0.8381
K10	0.5072	0.4721	0.5072	0//0.5072
K11	0.5681	0.5293	0.5681	0//0.5681
K12	0.4526	0.4215	0.4526	0//0.4526
K13	1	1	1	1//0

4 RANKING DE LAS DMUS

A partir de los resultados obtenidos en la Tabla 3, de acuerdo con (13) y (20), ordenamos las DMUs teniendo en cuenta la ordenación de sus escores de eficiencia $\{\widetilde{E}_j\}_{j=1}^{54}$.

Método de Yager

K13≈K20≈K41 ≻ K33 ≻ K17 ≻ K26 ≻ K49 ≻ K4 ≻ K9≈K42 ≻ K30≈K45 ≻ K52 ≻ K18 ≻ K14 ≻ K1 ≻ K36 ≻ K27 ≻ K53 ≻ K3 ≻ K28 ≻ K31 ≻ K15 ≻ K50 ≻ K29 ≻ K38 ≻ K35 ≻ K39 ≻ K11 ≻ K19 ≻ K7 ≻ K25 ≻ K51 ≻ K8 ≻ K44 ≻ K10≈K43 ≻ K34 ≻ K16 ≻ K12 ≻ K54 ≻ K40 ≻ K5 ≻ K37 ≻ K23 ≻ K47 ≻ K6 ≻ K2 ≻ K48 ≻ K22 ≻ K32 ≻ K46 ≻ K24 ≻ K21.

Método de Chen y Klein

K13≈K20≈K41 ≻ K33 ≻ K17 ≻ K26 ≻ K49 ≻ K4 ≻ K9≈K42 ≻ K30≈K45 ≻ K52 ≻ K18 ≻ K1 ≻ K14 ≻ K36 ≻ K27 ≻ K53 ≻ K3 ≻ K28 ≻ K31 ≻ K15 ≻ K50 ≻ K29 ≻ K38 ≻ K35 ≻ K39 ≻ K11 ≈K19 ≻ K7 ≻ K25 ≻ K51 ≻ K8 ≻ K44 ≻ K10≈K43 ≻ K34 ≻ K16 ≻ K12 ≻ K54 ≻ K40 ≻ K5 ≻ K37 ≻ K23 ≻ K47 ≻ K6 ≻ K2 ≻ K48 ≻ K22 ≻ K32 ≻ K46 ≻ K24 ≻ K21.

Método 1-Hamming

K13≈K20≈K41 ≻ K33 ≻ K17 ≻ K26 ≻ K49 ≻ K4 ≻ K9≈K42 ≻ K30≈K45 ≻ K52 ≻ K18 ≻ K14 ≻ K1 ≻ K36 ≻ K27 ≻ K53 ≻ K3 ≻ K28 ≻ K31 ≻ K15 ≻ K50 ≻ K29 ≻ K38 ≻ K35 ≻ K39 ≻ K11 ≻ K19 ≻ K7 ≻ K25 ≻ K51 ≻ K8 ≻ K44 ≻ K10≈K43 ≻ K34 ≻ K16 ≻ K12 ≻ K54 ≻ K40 ≻ K5 ≻ K37 ≻ K23 ≻ K47 ≻ K6 ≻ K2 ≻ K48 ≻ K22 ≻ K32 ≻ K46 ≻ K24 ≻ K21.

Método de Boscá et al.

K13≈K17≈K20≈K33≈K41 ≻ K26 ≻ K49 ≻ K4 ≻ K9≈K42 ≻ K30≈K45 ≻ K52 ≻ K18 ≻ K14 ≻ K1 ≻ K36 ≻ K27 ≻ K53 ≻ K3 ≻ K28 ≻ K31 ≻ K15 ≻ K50 ≻ K29 ≻ K38 ≻ K35 ≻ K39 ≻ K11 ≻ K19 ≻ K7 ≻ K25 ≻ K51 ≻ K8 ≻ K44 ≻ K10 ≻ K43 ≻ K34 ≻ K16 ≻ K12 ≻ K54 ≻ K40 ≻ K5 ≻ K37 ≻ K23 ≻ K47 ≻ K6 ≻ K2 ≻ K48 ≻ K22 ≻ K32 ≻ K46 ≻ K24 ≻ K21.

5 CONCLUSIONES

El análisis de la eficiencia propuesto incorpora la incertidumbre en la formulación del modelo y plantea situaciones en las que una DMU es particularmente sensible a la medición de la variable.

Los resultados de eficiencia para cada DMU, al aplicar el DEA fuzzy, pertenecen a un intervalo proporcionando más información que el modelo el DEA crisp en el que obtenemos un número.

Observamos que al aplicar los 4 procedimientos de clasificación el orden es muy similar y afianza más, si cabe, los resultados obtenidos al aplicar el modelo DEA fuzzy.

Este ranking ofrece un buen dato para el estudio de la competitividad entre las empresas pues la diferencia entre las valoraciones obtenidas en la Tabla 3, que permite establecer la ordenación, da idea sobre los esfuerzos que debe realizar una empresa para obtener una eficiencia similar a su inmediata anterior en la clasificación. Actualmente el método está siendo implementado en un programa que se pasará a alguna de las empresas del estudio para comprobar la utilidad del mismo.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el Ministerio de Educación y Ciencia a través del proyecto de investigación TIN2008-06872-C04-02.

Referencias

[1] J.E. Boscá, V. Liern, R. Sala, A. Martínez: Ranking decision making units by means of soft computing DEA models, *International Journal of Uncertainty, Fuzziness and Knowledge-based systems,* 19, 115-134, 2011.

[2] A. Charnes, W.W. Cooper, y E. Rhodes: Measuring the Efficiency of Decision Making Units, European Journal 1, Operational Research, 2, 429-444, 1978.

[3] V. Coll, O. Blasco: Evaluación de la eficiencia de la industria textil española a partir de información económico-financiera: una aplicación del análisis envolvente, *Revista investigación operacional* 28, 61-91, 2007.

[4] C. B. Chen, C. M. Klein: An efficient approach to solving fuzzy MADM problems, *Fuzzy Sets and Systems* 88, 51-67, 1997.

[5] N. Duch: Posición competitiva y estrategias de las empresas catalanas. Análisis del Programa Créixer (2003-2005*). Documents de treball*. CIDEM, Generalitat de Catalunya. 2006.

[6] J. Esteban, V. Coll: Competitividad y eficiencia, *Estudios de Economía Aplicada,* 21, 423-450. 2003.

[7] Farrell, M.J. (1957): The Meausurement of Productive Efficiency. Yournal of the Royal Statistical Society. Series A, 120, Part III, 253-290.

[8] P. Guo, H. Tanaka: Fuzzy DEA: A Perceptual Evaluation Method, *Fuzzy Sets and Systems* 119, 149-160, 2001.

[9] C. Kao, S.T. Liu: Fuzzy Efficiency Measures in Data Envelopment Analysis, *Fuzzy Sets and Systems* 119, 149-160, 2000.

[10] J.M. León Rojas: Juicios por comparación, inferencias lingüísticas y actos de decisión en sistemas de representación de conocimiento efectivamente computables basados en unidades vagamente perfiladas. Tesis doctoral Universidad de Extremadura, 2003.

[11] O.B. Olesen, y N.C. Petersen (1995): Chance Constrained Efficiency Evaluation, Management Science, 41, 442-457.

[12] M. Porter: The Competitive Advantage of Nations. The Free Press. USA. 1990.

[13] O. Roca, H. Sala: Producción, empleo y eficiencia productiva de la empresa española: Una radiografía a partir de SABE, *Boletín Económico del* ICE, 2857, 21-38, 2005.

[14] X. Wang, E. Kerre: Reasonable properties for the ordering of fuzzy quantities (I), *Fuzzy Sets and Systems* 118, 375 -385, 2001.

[15] X. Wang, E. Kerre: Reasonable properties for the ordering of fuzzy quantities (II), *Fuzzy Sets and Systems* 118, 387 - 405, 2001.

[16] R.R. Yager: A Procedure for Ordering Fuzzy Subsets of the Unit Interval, *Information Sciences* 24, 143-161, 1981.

[17] L.A. Zadeh: Fuzzy Sets, Information and Control 8, 338-353, 1965.

Controladores borrosos para la dirección de vehículos autónomos en maniobras dentro de entornos urbanos

Joshué Pérez Rastelli[1], **Teresa de Pedro Lucio**[1], **Matilde Santos Peñas**[2]

[1]*Centro de Automática y Robótica, UPM-CSIC, Arganda del Rey, Madrid, España*, {joshue.perez,teresa.pedro}@csic.es
[2]*Departamento de arquitectura de computadores y automática, Facultad de Informatica, Universidad Complutense de Madrid, España*, msantos@dacya.ucm.es

Resumen

Hasta la fecha, los sistemas de ayuda a la conducción desarrollados en el sector de la automoción se centran especialmente en el control de velocidad del vehículo. Sin embargo, sistemas que involucren el control (ya sea parcial o total) sobre la dirección del vehículo se encuentran todavía en fase experimental. Este trabajo está centrado en el diseño, desarrollo e implementación de un sistema de control lateral en cascada para vehículos autónomos reales, basado en controladores borrosos de alto nivel para maniobras en circuitos urbanos. Diferentes experimentos se han llevado a cabo en curvas de distinto radio y a diferentes velocidades (dentro de entornos urbanos), además, se han implementado dos nuevas maniobras: la marcha atrás y conducción en rotondas, mostrando un buen funcionamiento.

Palabras Clave: Control lateral, logica difusa, entornos urbanos, vehículos autónomos, control inteligente.

1 Introducción

Las investigaciones realizadas en los Sistemas de Trasporte Inteligente - ITS - (del ingles, *Intelligent Transportation Systems*) están encaminadas a la mejora de la seguridad y el confort en la conducción de vehículos. Diferentes maniobras han sido desarrolladas e implementadas en vehículos comerciales, entre ellas: el control de crucero (CC) [1] y el control de crucero adaptativo, sistemas de frenado ABS [2], sistemas que actúan de forma autónoma sobre el vehículo [3], sistemas de aviso de ángulo muerto [4], etc. La mayoría de estas aplicaciones están enfocadas al control longitudinal (acción sobre el freno y el acelerador), aunque algunas, a muy baja velocidad, toman el control del volante como: el aparcamiento asistido [5], y los sistemas de dirección asistida (EPS o EPAS) [6], los cuales solo toman el control parcial del volante para avisar al conductor sobre situaciones de peligro.

La escasa implantación comercial del control lateral se debe, en parte, a las fatales consecuencias que implican cambios bruscos en el volante (a veces giros de pocos grados pueden hacer que el vehículo vuelque), dejando al conductor sin tiempo para reaccionar y recuperar el control del volante. Por ello, la implementación de un sistema capaz de controlar de forma precisa el volante se antoja fundamental para el desarrollo de los vehículos autónomos. Este sistema no puede limitarse solo al control de giro del volante [7, 8], también es necesario un giro suave y seguro, y para ello hay que considerar la velocidad angular del volante como parámetro de control. En este sentido, en el presente trabajo se presenta una arquitectura de control que utiliza dos variables, la posición y la velocidad angular del volante en el lazo de control que gobierna la dirección del vehículo.

Por otra parte, sabemos que los controladores borrosos tienen la ventaja de utilizar el conocimiento de expertos y esto, permite una intuitiva y rápida sintonía para diferentes maniobras. En este sentido, se ha incluido de la velocidad giro del volante como parámetro de control, así como variables que, hasta ahora solo, estaban relacionadas con el control longitudinal, mejorando la conducción tanto en tramos rectos como en tramos curvos.

Muchos grupos de investigación sólo han considerado la posición del volante como único parámetro de control, sin tomar en cuenta la velocidad del vehículo. Aunque los resultados de este tipo de control son aceptables, en muchos casos, el volante oscila, sobre todo en tramos curvos. El objetivo de la arquitectura presentada en este trabajo es incluir la velocidad del vehículo y la distancia a la próxima curva como nuevas variables de entrada al sistema, manteniendo además la simplicidad de la estructura general del sistema. Los experimentos que se han llevado a cabo permiten demostrar la arquitectura de control planteada se puede aplicar a nuevas maniobras que, hasta ahora, no habían sido consideradas: la marcha atrás en calles ciegas y la conducción en rotondas.

La estructura de este artículo se divide como sigue: la sección 2 describe los trabajos previos, así como las principales aportaciones. La sección 3 describe el vehículo y las pistas de pruebas del CAR. Los controladores borrosos utilizados en los diferentes experimentos, se explican en la sección 4. Con el objetivo de validar los diferentes controladores, enmarcados dentro de la misma arquitectura de control, se han realizados experimentos con un vehículo real (sección 5). Finalmente, se remarcan las conclusiones más importantes, así como los trabajos futuros.

2 Trabajos previos

En el mundo del control existe una brecha entre el control de sistemas dinámicos simulados y la implementación de este control en sistemas reales. El control inteligente es una buena alternativa para la automatización de sistemas complejos, en este caso, vehículos de serie.

Sugeno y Nishida [9] han utilizado lógica borrosa para el control de robots móviles. En [10] se presentan los fundamentos para sintonizar estos controladores. Por lo tanto, el problema queda reducido a establecer un conjunto de reglas de control y sintonizarlas en función de la experiencia humana.

Los experimentos descritos en este trabajo han sido realizados en una furgoneta eléctrica del programa AUTOPÍA, en la que se ha implementado un nuevo sistema de control para la dirección en vehículos autónomos. Aunque el trabajo es la continuación de otras anteriores, [11, 8], las mejoras conseguidas son considerables. Estas se mencionan a continuación:

- El sistema de control propuesto introduce la velocidad angular como parámetro de control que no había sido considerado en trabajos previos [8], con el objetivo de mejorar la suavidad y el confort del control.
- De acuerdo con la experiencia de los conductores humanos, dos nuevos parámetros de control han sido tenidos en cuenta, v.g., la velocidad actual y distancia a la próxima curva. De esta forma, el sistema es *multivariable* - MIMO - (del ingles, *Multiple Input Multiple Output*) e introduce, variables asociadas tradicionalmente al control longitudinal en el control lateral.
- La velocidad para curvas próximas a los 90º estaba fijada a 8 km/h [11, 8]. El sistema propuesto no limita la velocidad para ajustar la trayectoria al carril de circulación. De hecho, las curvas de más 90º han podido ser tomadas hasta los 24 km/h.
- Finalmente, otros escenarios como caminos bloqueados y rotondas ha sido considerando dentro del esquema de control. Éstos solo requieren un ajuste de las reglas o de las funciones de pertenencia asociadas a cada maniobra, manteniendo la intuición y simplicidad en el ajuste de los controladores.

3 Descripción del sistema

El vehículo utilizado para validar el sistema de control propuesto en este trabajo es una furgoneta eléctrica del programa AUTOPIA, del Centro de Automática y Robótica (UPM-CSIC). La figura 1 representa el esquema de control, también se especifican los periféricos utilizados en la percepción del entorno. Como puede apreciarse el control longitudinal está separado del lateral. El control longitudinal se presupone solucionado [12].

Teniendo en cuenta que el objetivo de este trabajo es mejorar el control del volante en todo tipo de trazados, se ha utilizado el control de velocidad previamente desarrollado [8]. Para controlar el acelerador se ha utilizado una tarjeta analógica [8], y para el control del freno se ha instalado una bomba hidráulica [13]. La arquitectura de control ha sido adaptada, permitiendo la inclusión de nuevos sensores [14], como RFIDs, IMU, sensores de efecto Hall (para el cálculo de la velocidad), entre otros, de forma fiable y sencilla [12]. Sin embargo, y para las aplicaciones presentadas en este trabajo, solo se hará uso del sistema de posicionamiento (GPS + IMU), y del sensor de velocidad.

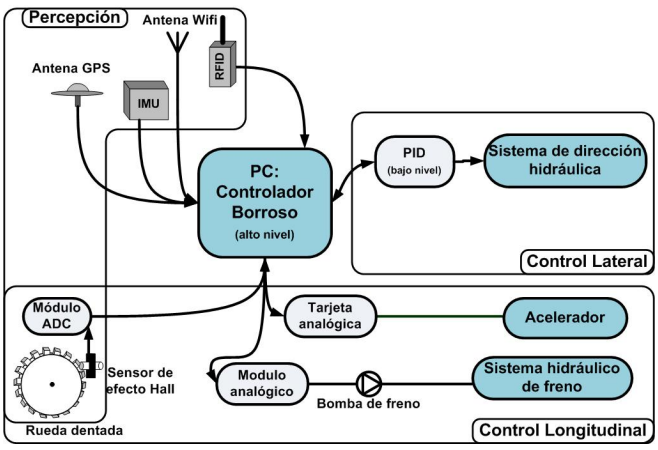


Figura 1: Esquema de control del vehículo

La figura 2 se muestra las pistas de pruebas utilizadas en los experimentos de este trabajo, en las instalaciones del CAR en Arganda del Rey, Madrid. La imagen superior muestra el trazado original de pistas, y la inferior la rotonda construida recientemente, con un radio de 13 metros y dos carriles internos.

Figura 2: Pistas de pruebas: trazado original y rotonda

4 Controladores borrosos para entornos urbanos

Para el control lateral del vehículo se ha utilizado una arquitectura en cascada. El lazo externo de control es un controlador borroso que emula al conductor humano y cierra el bucle de control con la señal proveniente del GPS y la del sensor de velocidad dispuesto [12]. El lazo interno de control es un PID discreto, el cual recibe la consigna desde el controlador borroso y lleva al volante hasta ella lo más rápida y suavemente posible.

Para el control lateral del vehículo se ha utilizado una arquitectura en cascada. El lazo interno de control está gobernado por un PID discreto[15]. El lazo externo de control es un controlador borroso que emula al conductor humano y cierra el bucle de control con la señal proveniente del GPS y del sensor de velocidad dispuesto para ello (figura 1).

Es sabido que la ventaja de los controladores borrosos radica en la facilidad para introducir la información y la experiencia humana del control de algunos procesos [10]. En este trabajo realizamos el control de tres maniobras en escenarios urbanos. Este control permiten demostrar lo intuitivo que son los controladores borrosos para una de las tareas más comunes y cotidianas de nuestros días: conducir un vehículo.

4.1 Controlador para curvas a diferentes velocidades

El control de alto nivel (figura 1) es controlador borroso. El ciclo de control lo cierran los datos provenientes del GPS y la velocidad leída directamente desde un sensor situado en una rueda delantera. Este controlador envía el ángulo del volante, junto con la velocidad de giro a la que el control de bajo nivel debe mover el motor solidario con el volante. Para la implementación del controlador borroso se ha utilizado la librería ORBEX, desarrollada por el grupo AUTOPÍA en sus comienzos [11].

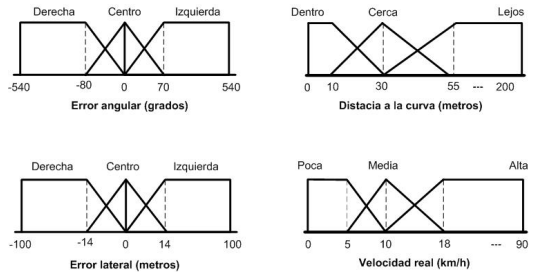


Figura 3: Funciones de pertenencia para el control del volante en rectas y curvas

El bucle de control tiene una frecuencia de 5 Hz. Como se verá en los resultados, esta frecuencia es suficiente para el buen desempeño del sistema. Las cuatro variables de entrada del controlador borroso se definen a continuación:

- ***error angular***: Ángulo entre la dirección del vehículo y la trayectoria predefinida, medido en grados.
- ***error lateral***: Distancia en metros desde la proyección en el suelo del morro del vehículo y la trayectoria predefinida.
- ***distancia a la curva***: Distancia del vehículo al centro de la próxima curva, esto es, el punto de intersección entre dos rectos consecutivos. Se estima utilizando la posición real del vehículo y el mapa de la trayectoria.
- ***velocidad real***: Velocidad del vehículo, calculada con el sensor de efecto hall, medida en kilómetros por hora (km/h).

Las reglas utilizadas para el control de posición del volante son:

SI *Error lateral* Izq. ENTONCES *Posición volante* Der.

SI *Error lateral* Ctr. ENTONCES *Posición volante* Mant.

SI *Error lateral* Der. ENTONCES *Posición volante* Izq.

SI *Error angular* Izq. ENTONCES *Posición volante* Der.

SI *Error angular* Ctr. ENTONCES *Posición volante* Mant.

SI *Error angular* Der. ENTONCES *Posición volante* Izq.

A continuación se presenta la base de reglas para controlar la velocidad angular. En un principio se habían utilizado 5 etiquetas para las funciones de pertenencia asociadas a la distancia a la próxima curva. Pero, debido a la simétrica con respecto al centro de la curva, se ha presentado solo los valores positivos de la de ésta, tal como se mencionó en la figura 3:

SI *Dist. a la curva* Dentro Y *Vel. real* Poca ENTONCES *Vel. angular* Media

SI *Dist. a la curva* Dentro Y *Vel. real* Media ENTONCES *Vel. angular* Media-Alta

SI *Dist. a la curva* Dentro Y *Vel. real* Alta ENTONCES *Vel. angular* Alta

SI *Dist. a la curva* Cerca Y *Vel. real* Poca ENTONCES *Vel. angular* Media

SI *Dist. a la curva* Cerca Y *Vel. real* Media ENTONCES *Vel. angular* Media-Alta

SI *Dist. a la curva* Cerca Y *Vel. real* Alta ENTONCES *Vel. angular* Media-Alta

SI *Dist. a la curva* Lejos Y *Vel. real* Poca ENTONCES *Vel. angular* Media

SI *Dist. a la curva* Lejos Y *Vel. real* Media ENTONCES *Vel. angular* Media

SI *Dist. a la curva* Lejos Y *Vel. real* Alta ENTONCES *Vel. angular* Poca

La fusión de ambos conjuntos de reglas en un mismo controlador permite que el volante gire suavemente, en función de las condiciones del entorno.

Los valores de salida del controlador son: Izquierda (-1) Mantener (0) y Derecha (+1) para el controlador de posición, y los valores Poca (0.4), Media (0.6), Media-Alta (0.8) y Alta (1) para el controlador de velocidad angular.

4.2 Controlador para la marcha atrás

Para el movimiento en marcha atrás se ha implementado el mismo esquema de control presentado en el figura 1. Incluso han sido utilizadas las mismas funciones de pertenencia, adaptadas a la cinemática del vehículo en marcha atrás. Solo se han cambiado, cuatro de las seis reglas de control de la posición del volante, sustituyéndolas por reglas duales como sigue:

SI *Error lateral* Izq. ENTONCES *Posición volante* Izq.

SI *Error lateral* Der. ENTONCES *Posición volante* Der.

SI *Error angular* Izq. ENTONCES *Posición volante* Izq.

SI *Error angular* Der. ENTONCES *Posición volante* Der.

4.3 Controlador para conducción en rotondas

Tomando como base la generación de trayectoria con curvas paramétricas propuesto en [16], se ha ajustado un nuevo

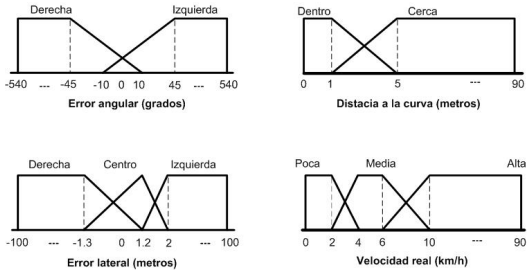


Figura 4: Funciones de pertenencia para el control del volante en rotondas

conjunto de reglas y funciones de pertenencia para la conducción autónomas en rotondas. En la figura 4 se muestran las funciones de pertenencia de este nuevo controlador.

En el conjunto de reglas es análogo al presentado en el apartado de conducción en curvas y rectas. Si bien la *distancia a la curva* ha sido reducida a dos funciones de entrada, esto se debe a que el efecto sobre el controlador se reduce debido la generación de puntos en la trayectoria de la rotonda [16]. Sin embargo, su efecto se aprecia cuando se realizan los cambios de carril dentro de la misma, esto es cuando se utiliza la etiqueta *Cerca*.

5 Experimentos

Los experimentos se han realizado en las instalaciones del CAR (UPM-CSIC). En todos se asume una velocidad crucero constante, aunque su valor puede ser cambiado en cada experimento.

5.1 Curvas a diferentes velocidades

En la figura 5 se muestran los experimentos con diferentes curvas y velocidades, desde los 8 km/h hasta los 24 km/h. Debido a la inclusión de la velocidad angular en el esquema de control se ha mejorado considerablemente el control en este tipo de escenarios.

En el experimento a 12 km/h se observa una pérdida en la corrección diferencial de la señal del GPS, por lo tanto un fallo en sistema de posicionamiento. No es parte de este trabajo dar una solución a este problema, que ya ha sido abordado en trabajos previos [17], y que actualmente se encuentra comercializado en sistemas de posicionamiento modernos. Sin embargo, la robustez del controlador borroso en cascada es capaz de filtrar ésta señal, debido a la acción sobre la velocidad angular, evitando saltos repentinos en la dirección, y manteniendo al vehículo en el camino.

5.2 Marcha atrás

En la figura 6 se muestra el comportamiento del vehículo circulando marcha atrás. Para realizar los cambios de mar-

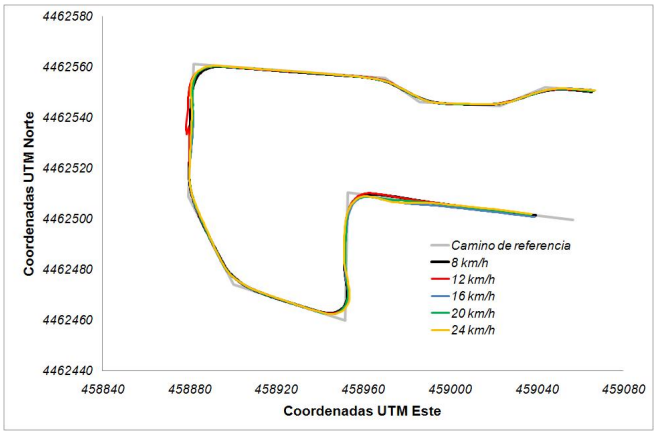


Figura 5: Curvas a diferentes velocidades

cha de forma autónoma, se ha habilitado una de las salidas analógicas del modulo ADC (figura 1), ya que la palanca de cambios en los vehículos eléctricos funciona como una resistencia (o *trimer*), lo que permite que solo con cambios del voltaje, se realicen los cambios de marcha desde el programa de control.

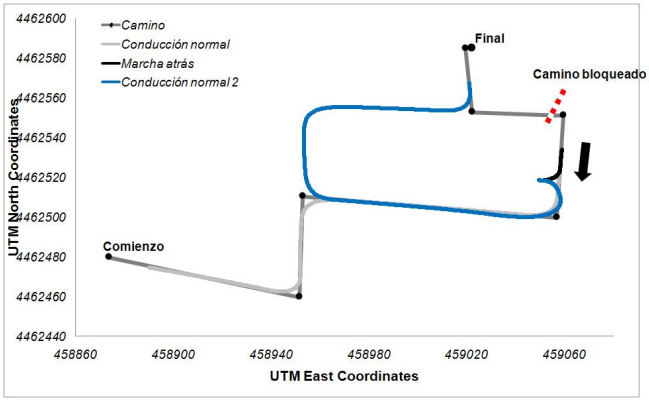


Figura 6: Secuencia de maniobra para la marcha atrás

Cuando el vehículo e n circulación recibe desde la estación central una señal indicando que el camino está bloqueado ([14]), el vehículo se detiene, define una nueva trayectoria y va marcha atrás hasta que encuentra un lugar para girar, vuelve a activar el controlador de marcha hacia a adelante.

5.3 Conducción en rotondas

En el último experimento el vehículo entra y sale de la rotonda, se muestra en la figura 7. En estos experimentos se asume que el vehículo viene por un camino recto, hasta que detecta una rotonda. Las rotondas pueden definirse como intersecciones que tienen un radio y centro definidos [16]. Con estos datos pueden generarse, en cada ciclo de control, la trayectoria a la que debe circular el vehículo, utilizando las ecuaciones paramétricas de una circunferencia [16].

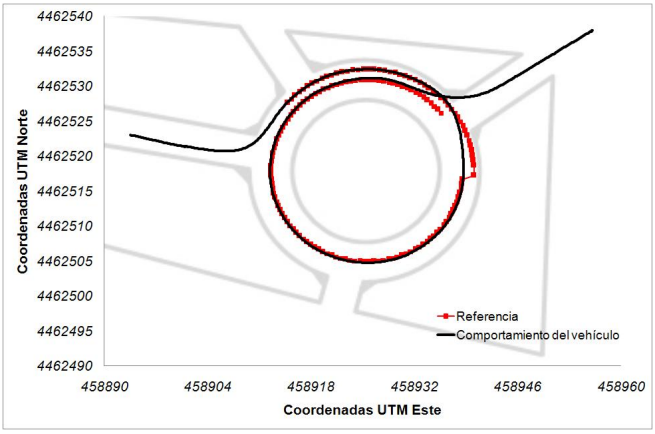


Figura 7: Maniobras en rotondas: entrada, cambio de carril y salida

Los puntos en rojo dibujados en la figura 7 muestran el mapa generado, mientras la gráfica negra representa la trayectoria del vehículo. Al entrar en la rotonda el vehículo circula en sentido anti horario y va por el carril central. En la primera vuelta intenta salir por la segunda salida, pero está congestionada [18]). Por eso el sistema de control decide dar otra vuelta, y cuando la salida queda libre, realiza un cambio de carril y sale de la rotonda.

6 Conclusiones y trabajos futuros

El sistema de control lateral basado en lógica borrosa propuesto en este trabajo, presenta buenos resultados en cada de uno de los tres escenarios urbanos que se han considerado: curvas, caminos bloqueados (marcha atrás) y rotondas.

El esquema de control que se presenta es muy intuitivo para el diseño de controladores borrosos. La sintonización y ajustes de los mismos se basan en las condiciones del vehículo y la experiencia de conductores expertos. La inclusión de una nueva variable de control, -la velocidad angular - asociada al comportamiento longitudinal del vehículo, permite que el movimiento del volante se asemeje más al de que haría un conductor humano.

Las pruebas en curvas de diferentes radios y a diferentes velocidades mejoran notablemente el estado de la técnica. Además, las maniobras de la marcha atrás y en la rotonda abren la posibilidad de considerar nuevos escenarios que resultan difíciles para algunos conductores.

Como parte de los trabajos futuros, se pretende desarrollar otras técnicas de control, considerando redes neuronales u otros algoritmos inteligentes que permitan ajustar mejor los controladores laterales, así como se ha hecho en trabajos previos [19]. Además, se presente incluir el control longitudinal dentro del esquema general de control, así como

considerar otros escenarios con otros vehículos, en las rotondas.

Agradecimientos

Este trabajo fue desarrollado gracias al apoyo del Plan Nacional, España, bajo el proyecto Tránsito (TRA2008-06602- C03-01), por la Comisión Interministerial de Ciencia y tecnología, España, bajo el Proyecto GUIADE (Ministerio de Fomento T9/08), y la financiación del proyecto CICYT DPI2009-14552-C02-01.

Referencias

[1] B. van Arem, C. J. G. van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," vol. 7, no. 4, pp. 429–436, 2006.

[2] J.-S. Lin and W.-E. Ting, "Nonlinear control design of anti-lock braking systems with assistance of active suspension," *IET Control Theory & Applications*, vol. 1, no. 1, pp. 343–348, 2007.

[3] T. hua Hsu, J.-F. Liu, P.-N. Yu, W.-S. Lee, and J.-S. Hsu, "Development of an automatic parking system for vehicle," in *Proc. IEEE Vehicle Power and Propulsion Conf. VPPC '08*, 2008, pp. 1–6.

[4] R. OMalley, M. Glavin, and E. Jones, "Vision-based detection and tracking of vehicles to the rear with perspective correction in low-light conditions," *IET Intelligent Transport Systems*, vol. 5, no. 1, pp. 1–10, 2011.

[5] I. Totzke, S. Jessberger, D. Muhlbacher, and H.-P. Kruger, "Semi-autonomous advanced parking assistants: Do they really have to be learned if steering is automated?" *IET Intelligent Transport Systems*, vol. 5, no. 2, pp. 141–147, 2011.

[6] J. Yoon, W. Cho, B. Koo, and K. Yi, "Unified chassis control for rollover prevention and lateral stability," vol. 58, no. 2, pp. 596–609, 2009.

[7] Y. Hayakawa, R. White, T. Kimura, and G. Naito, "Driver-compatible steering system for wide speed-range path following," vol. 9, no. 3, pp. 544–552, 2004.

[8] J. E. Naranjo, C. Gonzalez, R. Garcia, T. de Pedro, and R. E. Haber, "Power-steering control architecture for automatic driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 4, pp. 406–415, 2005.

[9] M. Sugeno and M. Nishida, "Fuzzy control of model car," *Fuzzy Sets and Systems*, vol. 16, pp. 103–113, 1985.

[10] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[11] R. García, T. de Pedro, and A. Roselli, "Estrategias borrosas para pilotos automáticas," in *3 Jornadas Iberoamericanas de Automáticas, Cartagena de Indias, Colombia.*, 1997.

[12] J. Pérez, F. Seco, V. Milanés, A. Jiménez, J. Díaz, and T. de Pedro, "An rfid-based intelligent vehicle speed controller using active traffic signals," *Sensors*, vol. 10, no. 6, pp. 5872–5887, 2010.

[13] V. Milanés, C. González, J. Naranjo, E. Onieva, and T. de Pedro, "Electro-hydraulic braking system for autonomous vehicles," *International Journal of Automotive Technology*, vol. 11, no. 1, pp. 89–95, Feb. 2010.

[14] J. Pérez, C. Gonzalez, V. Milanes, E. Onieva, J. Godoy, and T. de Pedro, "Modularity, adaptability and evolution in the autopia architecture for control of autonomous vehicles," in *Proc. IEEE International Conference on Mechatronics ICM 2009*, 14–17 April 2009, pp. 1–5.

[15] V. Milanés, J. Pérez, E. Onieva, C. González, and T. de Pedro, "Lateral power controller for unmanned vehicles," *Electrical Review*, vol. 86, no. 1, pp. 207–211, Jan. 2010.

[16] J. Pérez, V. Milanes, T. de Pedro, and V. L., "Autonomous driving manoeuvres in urban road traffic environment: a study on roundabouts," in *18th World Congress of the International Federation of Automatic Control (IFAC), Milán-Italia*, 14–17 April 2011, pp. 1–5.

[17] V. Milanés, J. Naranjo, C. Gonzalez, J. Alonso, and T. de Pedro, "Autonomous vehicle based in cooperative gps and inertial systems," *Robotica*, vol. 26, no. 5, pp. 627–633, Sept. 2008.

[18] J. Pérez, J. Villagra, E. Onieva, V. Milanes, T. de Pedro, and L. Vlacic., "Driving by driverless vehicles in urban environment," in *Computer Aided Systems Theory - EUROCAST 2011*, ser. Lecture Notes in Computer Science, S. B. . Heidelberg, Ed. Springer LNCS, 2011, vol. 4739/2007, pp. 1151–1158.

[19] J. Pérez, A. Gajate, V. Milanes, E. Onieva, and M. Santos, "Design and implementation of a neuro-fuzzy system for longitudinal control of autonomous vehicles," in *Proc. IEEE World Congress on Computational Intelligence WCCI 2010*, 14–17 April 2010, pp. 1–5.

DETECCIÓN DE INCOHERENCIAS DOCUMENTALES MEDIANTE EL USO DE SISTEMAS NEURO-DIFUSOS Y CONOCIMIENTO EXPERTO

Susana Martín Toral[1], **Susana San José Alonso**[1], **Gregorio I. Sáinz Palmero**[2]

[1]*CARTIF Centro Tecnológico. Parque Tecnológico de Boecillo, parcela 205, 47151 Boecillo (Valladolid), España*, {susmar,sussan}@cartif.es

[2]*Departamento de Ingeniería de Sistemas y Automática. Escuela de Ingeniería Industrial, Universidad de Valladolid. Paseo del Cauce s/n, 47011 Valladolid, España*, gresai@eii.uva.es

Resumen

La forma en que se generan, modifican y actualizan las colecciones de documentos puede ocasionar la aparición de errores y contradicciones en la forma de incoherencias documentales. Para hacer frente a este problema real, en este trabajo se propone una solución basada en sistemas neuro-difusos capaz de detectar incoherencias entre los conceptos, valores o referencias contenidas en la documentación de una compañía del sector eléctrico. Para ello se emplean representaciones de la documentación en la forma de N-tuplas. Posteriormente estas representaciones son enfrentadas entre sí empleando algoritmos de comparación de cadenas de caracteres. La decisión final sobre la existencia real de una incoherencia se obtiene mediante el entrenamiento supervisado de diversos sistemas neuro-difusos (FasArt y Nefprox) utilizando el conocimiento de los expertos del dominio, y teniendo en cuenta su subjetividad así como su diversidad de opinión. Estos sistemas son capaces de recopilar y unificar este conocimiento en la forma de reglas difusas y lingüísticas. El sistema final ofrece de este modo una solución semi-automática para la detección de incoherencias y el soporte a la decisión.

Palabras Clave: Incoherencia, N-tupla, sistema neuro-difuso, conocimiento experto, sistema de toma de decisión.

1 INTRODUCCIÓN

Los documentos, tanto en formato papel como en formato electrónico, constituyen el elemento básico de almacenamiento e intercambio de información en diversos contextos de la actividad humana. Por ello, la información y el conocimiento que contienen deben ser correctos y claros, sin posibilidades de confusión o contradicción. Lamentablemente, este objetivo no es fácil de alcanzar, y resulta complicado encontrar organizaciones que manejen de forma apropiada y formalizada conjuntos de documentación relacionada y heterogénea, que dispongan de una única formulación para su generación, gestión y control. Todo ello deriva en la aparición de incoherencias en documentación relacionada: errores en las referencias cruzadas, existencia de información redundante, contradictoria, o incorrecta, que da lugar a interpretaciones ambiguas [6, 11].

El impacto que la aparición de incoherencias puede ocasionar en una organización, tanto en sus relaciones internas como externas, se puede traducir en problemas económicos, legales, técnicos e incluso sociales. A modo de ejemplo, existen algunos sectores que se ven especialmente afectados por el problema de las incoherencias, como son el sector sanitario [12], las compañías de software [1], el sector legal [14] o el sector eléctrico [11]. El presente trabajo se centra en este último sector.

Una documentación libre de incoherencias redundará en una mejora en la gestión coherente de la misma, así como en la generación de productos y servicios de mayor calidad.

Por otro lado, el problema de la existencia o no de incoherencias tiene asociada una componente subjetiva. Esto se debe a que es un experto del dominio el que finalmente decide si un error o contradicción en la documentación puede ser considerada como una incoherencia documental, y cuál puede ser su relevancia. Además, diferentes expertos pueden opinar de forma distinta sobre una incoherencia, complicando la resolución del problema planteado.

Teniendo en cuenta estos problemas, el presente artículo trata sobre la existencia de incoherencias en la documentación de una empresa del sector eléctrico, y plantea una posible solución para su resolución. Dicha propuesta se basa en obtener diversas representaciones de la documentación en función de la incoherencia que se desee detectar, la aplicación de técnicas de enfrentamiento en-

tre representaciones, y finalmente la detección de las incoherencias mediante el entrenamiento y test supervisado de diversos clasificadores neuro-difusos. El empleo de estos clasificadores nos permitirán además recopilar el conocimiento subjetivo del experto del dominio a través de las reglas difusas obtenidas. Estas reglas se pueden incorporar más tarde a un sistema de toma de decisión. De ahí el interés en utilizar este tipo de sistemas.

El resto del artículo se organiza de la siguiente manera: en la siguiente sección se presentará una definición formal de qué se considera una incoherencia, así como una clasificación de las mismas. A continuación se expone la propuesta de solución para la detección de incoherencias. Más tarde se mostrará la experimentación realizada que prueba la validez de la solución propuesta. Finalmente se discutirán los resultados más interesantes obtenidos junto con las principales conclusiones extraídas en la elaboración de este trabajo.

2 DEFINICIÓN DE INCOHERENCIA DOCUMENTAL

Llegados a este punto, parece necesario proporcionar algún tipo de descripción sobre lo que es considerado una incoherencia en este trabajo: *se denomina incoherencia a la falta de coherencia entre documentos relacionados, o dentro de un mismo documento, o la falta o exceso de información en él, detectadas por la comparación de documentos o piezas de un documento entre sí* [11].

Partiendo de la muestra documental usada en este trabajo, procedente de una empresa del sector eléctrico, se ha podido establecer una clasificación de los diversos tipos de incoherencias que aparecen con más frecuencia en su documentación, según los expertos de este dominio [6, 10]:

- *Numéricas y de Atributos*. Estas incoherencias tienen que ver con los valores numéricos o atributos técnicos (como colores, formas, estados, etc.) contenidos en un documento y que deben estar de acuerdo con los valores indicados en las normas, estándares o documentos de referencia. Una contradicción entre documentos para el mismo concepto no está permitido en el dominio eléctrico.

- *Conceptuales*. Estas incoherencias aparecen cuando un concepto importante se denomina de forma diferente en el mismo o distintos documentos. Es muy importante utilizar los conceptos de una forma adecuada al contexto en el que trabajamos.

- *De Referencia*. Este tipo de incoherencia ocurre cuando un documento hace uso de una referencia a otro documento que no es el apropiado o que no existe, o bien si debe hacerle referencia y no lo hace.

Generalmente, para cada tipo de incoherencia a ser detectada, será necesario aplicar diferentes técnicas de procesamiento de información. Pero en la literatura técnica y científica, no es habitual encontrar la formulación del problema tal y como se presenta en este trabajo, o al menos con el mismo significado. Aún así, existen técnicas bien conocidas que pueden aplicarse en la detección de incoherencias como la minería de textos, reconocimiento de patrones, o el análisis semántico. En general, muchas de ellas se basan en el uso de soluciones heurísticas, con criterios similares a los encontrados en [8].

3 PROPUESTA PARA LA DETECCIÓN DE INCOHERENCIAS

El principal objetivo que se quiere alcanzar con esta propuesta es poder detectar cuándo una incoherencia está contenida en un documento. Con esta finalidad, hemos llevado a cabo el procedimiento mostrado en la figura 1. Su aplicación está orientada a la detección de incoherencias numéricas y de atributos, por ser los tipos más habituales e importantes de incoherencias que se dan en el dominio eléctrico.

Según esta figura, en primer lugar los documentos son resumidos o representados mediante un conjunto de términos y conceptos clave muy relevantes en el dominio, en la forma de N-tuplas (ver sección 3.1). Esta representación se obtiene a través de un procedimiento semi-automático basado en técnicas de extracción de información.

El siguiente paso se centra en el uso de técnicas de enfrentamiento de representaciones, con el fin de establecer el nivel de similitud entre los elementos de dos N-tuplas (ver sección 3.2). Tras este paso, se decide de una forma supervisada si existen incoherencias en el contenido de un documento o bien entre documentos conforme a sus medidas de similitud, al conocimiento del dominio y al conocimiento subjetivo del experto.

Finalmente se emplean dos tipos de clasificadores neurodifusos para recopilar este conocimiento experto y subjetivo, y del mismo modo ser capaces de detectar de forma automática cuándo se produce una incoherencia en la documentación y en qué grado.

Además de la detección y clasificación de incoherencias, esta propuesta es capaz de proporcionar un resultado que va más allá: la generación de una base de conocimiento que utiliza reglas difusas sobre el modo en el que las incoherencias son detectadas.

3.1 REPRESENTACIÓN DE DOCUMENTOS MEDIANTE N-TUPLAS

En este trabajo se emplean técnicas de extracción de información [2] con el fin de obtener representaciones que

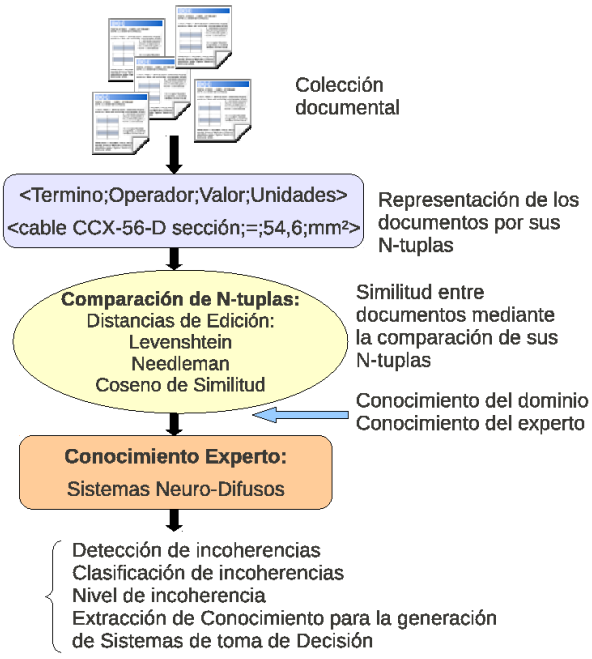


Figura 1: Propuesta para la detección de incoherencias documentales basada en sistemas neuro-difusos.

resuman cada documento del corpus. Aquí la extracción de información está basada en heurística [8], conforme a los patrones de información detectados dentro del cuerpo documental y que son relevantes para los expertos en el dominio eléctrico. Un ejemplo de este resumen es la representación de un documento mediante sus datos técnicos en la forma de "N-tuplas":

$$< \textit{Término ; Operador ; Valor ; Unidades} >$$
$$< \textit{Término ; Operador ; Atributo ;} >$$

Donde *Término* es el conjunto de palabras que representan un concepto relevante, *Operador* puede indicar que un término es mayor que, menor que, o igual a un valor/atributo específico, *Valor/Atributo* representa el valor numérico, o el atributo del término, y finalmente, las *Unidades* se emplean únicamente cuando el valor es numérico y con unidades de medida. Por lo tanto el documento es representado por un conjunto de N-tuplas, que han sido generadas mediante un procedimiento similar al proporcionado por las técnicas de la Minería de Reglas de Episodio (ERM - Episode Rule Mining) [9].

Un ejemplo real de N-tupla puede ser:

$$< \textit{cable CCX-56-D sección ; = ; 54,6 ; } mm^2 >$$
$$< \textit{cubierta del cable CCX-56-D color ; = ; verde ;} >$$

Esta forma de representación facilita la detección de incoherencias numéricas, de atributo y de unidades de medida. Este trabajo se centra en la detección de este tipo de incoherencias debido a su relevancia y alto grado de aparición en la documentación técnica, tal y como nos lo ha notificado el experto del dominio eléctrico.

3.2 MEDIDAS DE SIMILITUD APLICADAS A LOS ELEMENTOS DE LAS N-TUPLAS

Para determinar el grado de similitud que existe entre dos N-tuplas se han tenido en cuenta en este trabajo dos aproximaciones, una basada en la distancia de edición, con las distancias de Levenshtein [5, 4] y Needleman [4], y la otra en el espacio del vector, con el coseno de similitud [4] entre otras medidas.

En este trabajo, los documentos se representan mediante tuplas que contienen 4 elementos. Teniendo ésto en cuenta, el coseno de similitud se ha propuesto para establecer la similitud entre términos. Esta propuesta también se ha utilizado en trabajos como [7] y [5], donde se evalúan diferentes métodos de comparación de cadenas de caracteres en otros contextos. Por otro lado, los operadores, valores numéricos y unidades de medida son cadenas de caracteres cortas formadas por una palabra. Para comparar dos de estos elementos en las N-tuplas se propone el uso de distancias de edición.

El objetivo de comparar dos tuplas es el de detectar la existencia de incoherencias entre sus contenidos.

3.3 APRENDIZAJE DEL CONOCIMIENTO EXPERTO: SISTEMAS NEURO-DIFUSOS

Con la finalidad de incorporar tanto el resultado de la comparación entre N-tuplas, así como el conocimiento del dominio y el conocimiento subjetivo del experto del sistema, se emplean en este trabajo clasificadores neuro-difusos. Estos sistemas son capaces de recopilar y unificar este conocimiento en la forma de reglas difusas y lingüísticas.

Ésta parece ser la solución más apropiada por diferentes razones. Los sistemas neuro-difusos son aproximaciones muy populares para la generación de sistemas basados en reglas difusas que se benefician de la capacidad de aprendizaje de las redes neuronales, así como de la capacidad explicativa de la lógica difusa. Además, las tecnologías difusas son capaces de recopilar la vaguedad e incertidumbre que proporciona la subjetividad del experto en su decisión sobre la existencia real o no de una incoherencia.

En este trabajo se emplean dos sistemas neuro-difusos distintos: FasArt [3, 15] y NefProx[1] [13]. Ambos emplean reglas tipo Mamdani con el fin de obtener modelos de alta precisión. Pero la forma en la que cada algoritmo genera estos modelos es distinta, con lo que su rendimiento será diferente también, como se verá en la parte de experimentación y resultados.

[1] http://fuzzy.cs.uni-magdeburg.de/nefprox

4 METODOLOGÍA EXPERIMENTAL

Los documentos utilizados en este trabajo han sido representados a través de N-tuplas empleando para ello un procedimiento semi-automático. Para la experimentación se han empleado en concreto dos conjuntos de documentos que contienen N-tuplas:

1. Un conjunto de documentos proporcionado por una compañía del sector eléctrico consistente en 1 manual técnico y las 10 normas más referenciadas en el mismo. Estos documentos han sido representados en la forma de N-tuplas recogiendo los términos y conceptos más relevantes. Esta muestra está compuesta por un total de 3265 tuplas.

2. Un conjunto de documentos sintéticos que contiene 5 documentos generados a través de un procedimiento manual, con 29 tuplas ideales y diversas versiones de éstas con diferentes niveles de incoherencias. Este grupo está compuesto por un total de 1185 tuplas.

Una vez dispuestos los datos experimentales, se han aplicado técnicas de enfrentamiento entre N-tuplas para calcular las medidas de similitud. El experto de dominio ha evaluado cada caso de comparación, decidiendo si existe o no incoherencia, y etiquetando cada patrón de un modo supervisado. Como resultado de la etapa de comparación se han generado 242 ficheros con información real agrupados en 11 grupos, y 25 ficheros con información sintética agrupados en 5 grupos, con el total de medidas mostradas en la tabla 1.

Tabla 1: Resultado de la etapa de enfrentamiento

Tipo	# Ficheros	# Grupos	Medidas por grupo
Sintético	25	5	Más de 100.000
Real	242	11	Más de 1 millón

Los sistemas neuro-difusos son utilizados en este punto para aprender el conocimiento sobre las incoherencias contenidas en los documentos y representados mediante N-tuplas. La salida del sistema será la presencia de incoherencias, habiendo utilizado como entradas al sistema las medidas de similitud entre dos tuplas.

Por otra parte, los sistemas se han entrenado y testeado utilizando validación cruzada, empleando un grupo real o sintético para el entrenamiento y el resto de grupos reales o sintéticos para el test. Como criterio de comparación de resultados se ha calculado el error cuadrático medio del total de las pruebas (ver ecuación 1) y la complejidad del sistema final a través del número de reglas difusas (ver ecuación 2).

$$RMSE(f) = \sqrt{\frac{\sum_{i=1}^{n}(\hat{X}_i - X_i)^2}{n}} \quad (1)$$

$$Complejidad = NumeroReglas \quad (2)$$

Se han realizado los siguientes experimentos: a) Entrenamiento y test utilizando documentos sintéticos; b) Entrenamiento y test utilizando documentos reales; c) Entrenamiento con documentos sintéticos y test con documentos reales; d) Entrenamiento con documentos reales y test con documentos sintéticos.

5 RESULTADOS

5.1 RESULTADOS UTILIZANDO EL CLASIFICADOR FASART

Haciendo uso de los dos conjuntos de documentación presentados, en la tabla 2 se muestran las características del clasificador FasArt, así como la tasa de error conforme al RMSE.

Tabla 2: Mejores resultados en la detección de incoherencias utilizando el clasificador FasArt.

Similitud	Entrenamto	Test	N# Reglas	Error
Lev.	Syn.	Syn.	979	3.54 %
Lev.	Syn.	Real	957	**0.91 %**
Lev.	Real	Syn.	850	12.37 %
Lev.	Real	Real	893	1.55 %
Need.	Syn.	Syn.	1081	2.90 %
Need.	Syn.	Real	971	**0.94 %**
Need.	Real	Syn.	814	78.79 %
Need.	Real	Real	813	0.97 %

Atendiendo a la tasa de error, el mínimo error se obtiene cuando el sistema FasArt se entrena usando tuplas sintéticas y se testea usando tuplas reales, con un error de 0.91 % usando la distancia de Levenshtein y de 0.94 % empleando la distancia de Needleman. La complejidad del sistema ronda las 957-971 reglas. Este resultado parece coherente, ya que las tuplas sintéticas cubren un mayor número de casos que las tuplas reales. Con ello el conjunto de tuplas sintéticas representan un modelo ideal y teórico en la experimentación.

Un hecho destacado ocurre cuando el sistema es entrenado con tuplas reales y testeado con tuplas sintéticas. En esta situación se obtiene un error de 12.37 % usando la distancia de Levenshtein, y de 78.79 % usando la distancia de Needleman. Este resultado parece posible, ya que el uso de un único grupo de tuplas reales para cada entrenamiento no parece cubrir los mismos casos que con las tuplas sintéticas.

5.2 RESULTADOS UTILIZANDO EL CLASIFICADOR NEFPROX

Haciendo uso de los mismos conjuntos de datos que en el caso anterior, en la tabla 3 se presentan las características del clasificador Nefprox, así como la tasa de error obtenida según RMSE. En esta ocasión, debido a la inicialización aleatoria con la que trabaja el clasificador Nefprox, se han ejecutado tres pruebas. El error final se calcula como el valor medio de las tres ejecuciones.

Tabla 3: Mejores resultados en la detección de incoherencias utilizando el clasificador Nefprox.

Similitud	Entrenamto	Test	N# Reglas	Error
Lev.	Syn.	Syn.	121	**7.64 %**
Lev.	Syn.	Real	121	7.83 %
Lev.	Real	Syn.	58	11.48 %
Lev.	Real	Real	58	7.85 %
Need.	Syn.	Syn.	104	**8.52 %**
Need.	Syn.	Real	114	8.15 %
Need.	Real	Syn.	32	78.44 %
Need.	Real	Real	33	8.16 %

Conforme a los resultados presentados en la tabla 3, empleando la distancia de Levenshtein se obtiene una reducción fundamental en el número de reglas en comparación con el clasificador FasArt, aunque redundando en un incremento de la tasa de error en todos los casos. Esta reducción en la complejidad se obtiene definiendo el número y la forma de los conjuntos difusos del sistema, tanto en su entrada como en su salida. Por lo tanto, Nefprox ofrece un mayor control en la definición del sistema neuro-difuso pero a costa de incrementar la tasa de error del clasificador. En este caso el mínimo error se obtiene cuando el sistema es entrenado y testeado usando tuplas sintéticas, con un error del 7.64 % y una complejidad de 121 reglas para un sistema con 5 entradas y salidas difusas, y utilizando la distancia de Levenshtein. Para un sistema con 15 entradas se puede obtener un error de 1.64 % y un complejidad de 500 reglas.

El peor caso de clasificación se obtiene de nuevo cuando el sistema es entrenado usando tuplas reales y testeado usando tuplas sintéticas.

5.3 COMPARACIÓN DE LOS CLASIFICADORES

La complejidad en ambos sistemas indica el número de reglas difusas que necesitamos para recopilar el conocimiento experto, por lo que se hace necesario equilibrar la complejidad del sistema con el error para obtener una solución apropiada. El algoritmo FasArt ofrece unos resultados de clasificación más precisos a cambio de incrementar la complejidad del sistema. Por contra, el algoritmo Nefprox permite configurar la complejidad del sistema pero a cambio de incrementar el error del clasificador.

Este conocimiento experto recopilado puede ser reutilizado con el fin de obtener un sistema de toma de decisiones basado en las reglas difusas extraídas. Con FasArt el número de reglas obtenidas es siempre mayor, con lo que deberían ser simplificadas para disponer de un sistema experto adecuado. Usando Nefprox, el usuario puede decidir la complejidad del sistema pero a cambio de delimitar la tasa de error que es aceptable para el clasificador.

Tabla 4: Matriz de confusión para el clasificador Nefprox

	Salida esperada		
	No Incoh.	I.valor	I.medida
Salida real	sintético-sintético		
No Incoh.	**99.61 %**	0.23 %	0.16 %
I.valor	1.68 %	**98.32 %**	0 %
I.unidad	1.75 %	1.04 %	**97.21 %**
Salida real	sintético-real		
No Incoh.	**99.74 %**	0.26 %	0 %
I.valor	0.29 %	**99.71 %**	0 %
I.unidad	0.8 %	0.89 %	**98.31 %**
Salida real	real-real		
No Incoh.	**98.48 %**	1.51 %	0.01 %
I.valor	0.23 %	**99.76 %**	0.01 %
I.unidad	0.3 %	0 %	**99.7 %**
Salida real	real-sintético		
No Incoh.	**97.94 %**	0.94 %	1.12 %
I.valor	3.25 %	**90.69 %**	6.06 %
I.unidad	0.73 %	0 %	**99.27 %**

Por último, obtenemos la matriz de confusión del clasificador Nefprox representando el resultado real y el deseado en la clasificación de incoherencias (ver tabla 4). La situación ideal buscada es la de obtener una matriz diagonal para cada caso. Los resultados muestran que ciertas incoherencias se han clasificado de forma incorrecta. La utilización de un clasificador difuso puede explicar esta situación al obtener como salida grados de incoherencia de uno u otro tipo.

6 CONCLUSIONES

Este trabajo introduce el problema real de la existencia de incoherencias en la documentación de una empresa del sector eléctrico en el que los documentos conectados entre sí pueden contener errores, información incorrecta o contradictoria. Los efectos de la existencia de estas incoherencias pueden derivar en problemas económicos, legales, técnicos o sociales.

La detección de este tipo de problemas supone una serie de dificultades extra con respecto al uso habitual del problema de reconocimento de patrones: cuando una incoherencia ocurre en un documento, su detección depende de la

documentación del dominio y de sus expertos. Se trata por tanto de una cuestión no objetiva en la que es necesario aplicar el conocimiento del experto del dominio para alcanzar el éxito en la detección de incoherencias. En este trabajo, este conocimiento experto se incorpora al sistema a través de un procedimiento de aprendizaje supervisado, soportado por varios sistemas neuro-difusos.

En este trabajo se presenta una propuesta para el procesamiento de los documentos involucrados, con el fin de detectar incoherencias.

Los experimentos han mostrado que el sistema es capaz de cubrir la mayoría de los casos de coherencias e incoherencias que pueden darse lugar en un conjunto de documentación, con una tasa de éxito del 98 % usando la distancia de Levenshtein, y del 92 % usando la distancia de Needleman, en la mayoría de los casos. Las pruebas realizadas utilizando tanto muestras reales como sintéticas han demostrado que el sistema es capaz de aprender y detectar incoherencias por medio de las medidas de similitud obtenidas de comparar dos N-tuplas que contienen información numérica o de atributo.

En estos momentos el trabajo sigue en desarrollo basándose en la especialización de los sistemas neuro-difusos para que sean capaces, no sólo de detectar la existencia o no de incoherencias, sino también de determinar la categoría de la incoherencia detectada, empleando la representación mediante N-tuplas. Por otro lado, empleando esta aproximación difusa, es posible extraer de los sistemas neuro-difusos el conocimiento subjetivo del experto del dominio, a través de un conjunto de reglas difusas que puedan dar soporte a un sistema de toma de decisiones sobre este problema complejo y no objetivo.

Referencias

[1] F. Arango: Gestión de inconsistencias en la evolución e interoperación de los esquemas conceptuales OO, en el marco formal de OASIS. PhD tesis, Universidad Politécnica de Valencia, Valencia, España, 2003.

[2] M. W. Berry: *Survey of Text Mining: Clustering, Classification, and Retrieval*. Springer, New York, 2004.

[3] J. M. Cano Izquierdo, Y. A. Dimitriadis, E. Gómez Sánchez, J. Coronado López: Learning from noisy information in FasArt and fasback neuro-fuzzy systems. *Neural Networks* 14(4-5), pp. 407–425, 2001.

[4] S. Chapman: Sam's String Metrics page. Available at http://www.dcs.shef.ac.uk/šam/stringmetrics.html (Accedido dic.09).

[5] W. W. Cohen, P. Ravikumar, S. E. Fienberg: A comparison of string metrics for matching names and records. En: *Proceedings of the KDD-2003 Workshop on Data Cleaning, Record Linkage, and Object Consolidation*, Washington DC (USA), pp. 13–18, 2003.

[6] GIDIN: Gestor inteligente de documentación incoherente (2008-2009). *Proyecto del Plan Avanza I+D*. Referencia del proyecto: TSI-020302-2008-73. Financiado por el Ministerio de Industria, Turismo y Comercio, España, 2008.

[7] N. Koudas, A. Marathe, D. Srivastava: SPIDER: flexible matching in databases. En: *SIGMOD'05: Proceedings of the 2005 ACM SIGMOD international conference on Management of Data*, New York (USA), pp. 876–878, 2005.

[8] B. Krulwich, C. Burkey: The infofinder agent: Learning user interests through heuristic phrase extraction. En: *IEEE Expert: Intelligent Systems and Their Applications* 12(5), pp. 22–27, 1997.

[9] H. Mannila, H. Toivonen, A. I. Verkamo: Discovery of frequent episodes in event sequences. *Data Mining Knowledge Discovery* 1(3), pp. 259–289, 1997.

[10] S. Martín, V. Arribas, G. I. Sáinz: Detection of incoherences in a document corpus based on the application of neuro-fuzzy system. En: *Tenth International Conference on Document Analysis and Recognition*, Barcelona (España), pp. 118–123, 2009.

[11] S. Martín, V. G. I. Sáinz, Y. Dimitriadis: Detection of incoherences in a technical and normative document corpus. En: *Tenth International Conference on Enterprise Information Systems, volume Artificial Intelligence and Decision Support Systems*, Barcelona (España), pp. 282–287, 2008.

[12] L. Mingshan, A. M. Ching-to: Consistency in performance evaluation reports and medical records. *The Journal of Mental Health Policy and Economics* 5(4), pp. 191–192, 2002.

[13] D. Nauck, R. Kruse: Neuro-fuzzy systems for function approximation. *Fuzzy Sets and Systems.* 101(2), pp. 261–271, 1999.

[14] M. Ruiz: *Sistemas jurídicos y conflictos normativos*. Dykinson, Universidad Carlos III de Madrid, Instituto de Derechos Humanos Bartolomé de las Casas, Madrid, 2002.

[15] G. I. Sainz Palmero, Y. Dimitriadis, J. Cano Izquierdo, E. Gómez Sánchez, E. Parrado Hernández : ART based model set for pattern recognition: FasArt family. En: H. Bunke y A. Kandel editores, *Neuro-fuzzy pattern recognition*, World Scientific Pub. Co., pp. 147–177, 2000.

Selección de Características para la conservación de vía ferroviaria mediante PCA, FA y operadores OWA, MA-OWA

Rogelio Contreras[2], **Marta Galende**[1], **María Jesús Fuente**[2], **Gregorio I. Sainz**[1,2]

[1]*Centro Tecnológico CARTIF, Parque Tecnológico de Boecillo, 47151 Valladolid*, {margal,gresai}@cartif.es
[2]*Escuela de Ingenierías Industriales, Paseo del Cauce 59, 47011 Valladolid*, {maria,gresai}@eii.uva.es

Resumen

En la actualidad cualquier proceso de actividad, sea del tipo que sea, está sujeto a seguimiento y monitorización. En el caso de la conservación de vías ferroviarias esto se lleva a cabo mediante auscultaciones periódicas. Esta información será posteriormente utilizada para determinar las intervenciones a afectuar sobre las vías. En este punto aparece el problema de decidir cuál son las variables más adecuadas cara a un sistema inteligente de apoyo dichas tareas, teniendo en cuenta que el conocimiento experto no está siempre disponible, ni es evidente de obtener, y sobre todo, no es generalizable. En este marco el presente trabajo se centra en el *descubrimiento* de qué variables son de mayor interés cara a un proceso de extracción de conocimiento de la base de datos sobre el estado de conservación de la infraestructura. Para esto, se ha realizado una selección de características no supervisada basada en dos conocidas técnicas de extracción:Análisis de Componentes Principales (PCA) y Análisis de Factores (FA). Esta selección se realiza en base a distintos criterios, para obtener una selección final se aplica un criterio de toma de decisiones multicriterio mediante operadores OWA (Ordered Weighted Averaging) y MA-OWA (Majority additive-Ordered Weighted Averaging). Finalmente, se ha hecho una comparación de los resultados obtenidos con la experiencia de los técnicos de conservación.

Palabras Clave: Selección de Características, Análisis de Componentes Principales, Análisis de Factores, OWA, Conservación de Infraestructuras.

1. Introducción: Infraestructuras y Lógica Difusa

La aplicación de técnicas basadas en Soft Computing, y más concretamente en Lógica Difusa, en distintos campos y dominios industriales no es nueva: [1], [2]. La aplicación en el campo de las infraestructuras, y en especial en la conservación de las mismas, no ha sido excesivo debido a diferentes causas, entre otras y como en otros dominios, la dificultad de disponer de información suficiente, y/o conocimiento experto, para aplicar estas técnicas, así como restricciones temporales y económicas.

Dentro del área de las infraestructuras ferroviarias algunas aplicaciones, clasificadas de acuerdo a la problemática a abordar, son: Seguridad vial [3], Gestión de tráfico ferroviario [4], [5], Sistemas de ayuda a la conducción [6], [7] o Salud estructural [8].

Este trabajo se enmarca en una línea de I+D basada en la aplicación de técnicas de soft computing, y en especial de la lógica difusa, al dominio de la conservación de infraestructuras ferroviarias [9] con el objetivo de conseguir herramientas de ayuda a la toma de decisiones en el área de la conservación. En este marco, el trabajo aquí presentado se basa en la primera etapa consistente en determinar qué variables son las más relevantes entre las variables recogidas periódica y sistemáticamente en el proceso de auscultación de las vías férreas. Este proceso de auscultación se realiza mediante la sensorización correspondiente en la cabeza tractora de un convoy convencional de forma programada, o bajo demanda, una vez al mes en alguno de los horarios programados. Esta auscultación tiene como objetivo la confortabilidad del usuario y el estado de la vía.

A partir de esta selección, en fases posteriores se podrá abordar la detección, clasificación y predicción de eventos que permitan una mejor y más efectiva política de conservación.

Hoy en día, a partir de la información de auscultación recogida de forma automática el experto analiza los gráficos y determina de acuerdo a su conocimiento y experiencia

lo sucedido en la infraestructura, y las acciones correctivas a llevar a acabo. La sistematización de este proceso no es fácil ni evidente, e incluso la subjetividad de los expertos es muy relevante en ese análisis y, por otro lado, la infraestructura *"madura"* cambiando sus parámetros de funcionamiento. Para hacer frente a esta problemática se plantea un proceso de análisis básicamente no supervisado en todas las fases, y una posterior confrontación con el conocimiento de los expertos para conseguir una interpretación de los resultados de acuerdo al dominio.

La organización de este trabajo es como sigue: primero una muy breve descripción, dado que son técnicas muy conocidas, de los aspectos teóricos que se han aplicado para la selección de características no supervisada y la toma de decisiones. A continuación la metodología propuesta y llevada a cabo en este trabajo para la realización de la selección. Finalmente se presentan los resultados y un análisis de los mismos junto a unas conclusiones finales.

2. Soporte Teórico

2.1. Selección de Características: PCA y FA

La selección de características es un dominio ampliamente estudiado y aplicado en multitud de problemas y campos [10]. Dentro de las técnicas de extracción de características, el Análisis de Componentes Principales (PCA) [11] es una de las técnicas más conocidas y usadas en distintos propósitos. El Análisis de Factores (FA) [11] es otra técnica también muy popular en este ámbito y muy relacionada con PCA. Ambas técnicas pueden ser empleadas para la selección de características a pesar de ser técnicas de extracción. Con este propósito han sido empleadas en este trabajo dado el carácter *no supervisado* que era necesario en el proceso de selección.

2.2. Toma de decisiones: Operadores OWA, MA-OWA

A partir de los resultados obtenidos mediante PCA y FA, será necesario *unificar* la selección hecha por cada una de las aproximaciones utilizadas. Para tal tarea los conocidos operadores de agregación OWA (Ordered Weighted Averaging) [12], [13] y MA-OWA (Majority additive-Ordered Weighted Averaging) [14], [15] serán empleados para agregar las selecciones de características obtenidas.

3. Selección de Características de Auscultación: Propuesta y Metodología

El objetivo de la propuesta es explotar y poner en valor, de forma sistemática y metodológica, la información recogida en los procesos de monitorización/auscultación que se realizan sobre las líneas férreas de tal manera que sirva de ayuda a los técnicos encargados de la conservación. Esta ayuda se plantea desde un doble punto de vista:

- Dar soporte al conocimiento y experiencia existente.
- Aportar nuevo conocimiento.

A partir de aquí, una aproximación no supervisada para la selección de características es lo más apropiado en base al Análisis de Componentes Principales (PCA) y Análisis de Factores (FA): se busca hacer un *ranking* de variables/características que conserven una variabilidad alta pero reduciendo el número de variables a ser consideradas. De esta manera:

- Se reduce el número de variables a considerar pero se conserva aquellas que proporcionan mayor *información.*
- Se consigue una mayor interpretabilidad y facilidad de manejo de la información gracias a la reducción.

Al realizar esta selección en base a información histórica, se puede comprobar la evolución en el tiempo de la relevancia de cada característica, poniendo de manifiesto la *"maduración"* de la infraestructura y analizando los problemas que han ido apareciendo en el periodo considerado, y la efectividad de las soluciones desde dos puntos de vistas: la infraestructura y la confortabilidad del usuario.

La metodología seguida se puede ver en el Algoritmo 1

Algorithm 1 Metodología para la Selección de Características basada en PCA, FA, OWA y MA-OWA

```
for Trayecto= T1,T2 do
  for Via= Via1, Via2, Via12 do
    for Variabilidad=0.65, 0.75, 0.85, 0.95 do
      Fuzzification Resaltar Variables Relevantes μ_ResaltadoRelevancia(r)
      Selección basada en PCA (Trayecto, Via, Variabilidad)
      Selección basada en FA (Trayecto, Via, Variabilidad)
      Selección basada en OWA (PCA, FA, Trayecto, Via, Variabilidad)
      Selección basada en MA-OWA (PCA, FA, Trayecto, Via, Variabilidad)
    end for
  end for
end for
```

4. Trabajo Experimental

En este trabajo se han considerado dos trayectos, cada uno de los cuales tenía dos vías:

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
3	0,2391	2	0,5250	2	0,289	2	0,125
2	0,2364	3	0,3535	3	0,030	3	0,036
1	0,2148	1	0,0558	1	0,002	1	0,006
4	0,1778	6	0,0137	4	0,001	4	0,003
5	0,0648	5	0,0012	5	0,000	5	0,001
6	0,0170	4	0,0009	6	0,000	6	0,000

Tabla 1: Trayecto A, Vía 1. Relevancia de Variables (variabilidad 95 %) 2007-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
1	0,190	3	0,469	3	0,152	3	0,084
3	0,189	1	0,464	1	0,144	1	0,081
2	0,184	2	0,008	2	0,001	2	0,004
5	0,175	6	0,007	5	0,001	5	0,003
4	0,163	5	0,002	4	0,000	4	0,003
6	0,048	4	0,000	6	0,000	6	0,001

Tabla 2: Trayecto A, Vía 2. Relevancia de Variables (variabilidad 95 %) 2007-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

- **Trayecto A**, periodo 2007-2010, Vía 1: 1916 registros, Vía 2: 1487 registros
- **Trayecto B**, periodo 2005-2010, Vía 1: 7538 registros, Vía 2: 7473 registros

De cara al empleo de OWA y MA-OWA como operadores de agregación de los criterios PCA y FA, el valor (r) de cada variable obtenida por estos criterios es *fuzzificada* mediante $\mu_{ResaltadoRelevancia}(r)$ (Ec. 1). El objetivo es *resaltar* las variables más importantes, *mantener* la relevancia de las intermedias y *apenas considerar* las variables poco relevantes. Este paso puede ser *obviado* mediante una $\mu_{ResaltadoRelevancia}(r)$ lineal también considerada en el trabajo experimental.

$$\mu_{Resaltado}(r) = \begin{cases} 0 & \text{if } r \leq 0{,}07 \\ 2\left(\frac{r-0{,}07}{0{,}73}\right)^2 & \text{if } 0{,}07 \leq r \leq 0{,}5 \\ 1-2\left(\frac{r-0{,}07}{0{,}73}\right)^2 & \text{if } 0{,}5 \leq r \leq 0{,}8 \\ 1 & \text{if } r \geq 0{,}8 \end{cases} \quad (1)$$

Esta función ha sido parametrizada mediante el estudio estadístico de los valores obtenidos por las variables consideradas en el PCA y FA.

El cuantificador lingüístico considerado en este trabajo es tipo RIM con $andness = 0{,}4$ ($orness = 0{,}6$) (Ec. 2):

$$Q(r) = r^{\alpha} \qquad \alpha = \frac{1-orness}{orness} \quad (2)$$

4.1. Resultados

Por razones de espacio sólo se describen los resultados para variabilidad del 95 % y $\mu_{ResaltadoRelevancia}(r)$ de la Ec. 1.

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
3	0,238	3	0,488	3	0,193	3	0,100
2	0,232	6	0,435	6	0,097	6	0,062
1	0,225	1	0,019	2	0,002	2	0,007
4	0,196	2	0,005	1	0,002	1	0,006
5	0,042	5	0,003	4	0,001	4	0,004
6	0,017	4	0,000	5	0,000	5	0,000

Tabla 3: Trayecto A, Vía 12 (conjunta). Relevancia de Variables (variabilidad 95 %) 2007-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
2	0,252	3	0,718	3	0,947	3	0,248
3	0,231	2	0,168	2	0,004	2	0,011
1	0,221	6	0,025	1	0,002	1	0,006
5	0,211	1	0,025	5	0,001	5	0,005
6	0,033	5	0,014	6	0,000	6	0,001
4	0,003	4	0,000	4	0,000	4	0,000

Tabla 4: Trayecto B, Vía 1. Relevancia de Variables (variabilidad 95 %) 2005-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

Trayecto A:

La información recopilada y disponible de este trayecto es menor que en el caso del Trayecto B, la razón es que es un tramo más moderno y, por tanto, el número de auscultaciones realizadas ha sido menor. Esto también puede permitir hacer una comparación entre dos infraestructuras de distinta *edad* y estado de *madurez*.

En las Tablas 1, 2 y 3 se pueden ver los resultados obtenidos para el trayecto A, Vía 1, Via 2, y Vía 12 (conjunta) para el periodo 2007-2010 mediante PCA y FA, conservando el 95 % de la variabilidad, y la toma de decisiones basada en operadores OWA y MA-OWA. Se puede observar que para la Vía 1, las variables más relevantes son Var.3, Var. 2, y Var. 1. En el análisis de PCA aparecen las 4 primeras variables con relevancia muy parecida, pero una vez aplicado OWA y MA-OWA: las variables 2 y 3 son las más relevantes, con gran prevalencia de la Var. 2.

En el caso de la Vía 2 de este trayecto, los resultados obtenidos por PCA para las variables muestran relevancias muy similares para todas las variables excepto la Var.6, que es claramente descartada. El orden de importancia es Var:1, 3, 2, 5 y 4. En FA las elegidas son claramente las Var. 3 y 1, selección que queda refrendada mediante OWA y MA-OWA. Esta diferencia de resultados con respecto a la Vía 1 es razonable al tratarse de dos trazados no necesariamente iguales, con posible diferencias de orografía, tráfico, etc. Cuando se hace un análisis conjunto de ambas vías, Tabla 3, se puede observar como novedad respecto a los análisis individuales que FA tomó como variables significativas las Var. 3 y 6, quedando finalmente la Var. 3 como la variable relevante seguida de lejos por la Var. 6. El resto tiene valores muy inferiores.

Trayecto B:

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
2	0,269	3	0,686	3	0,865	3	0,237
3	0,245	2	0,161	2	0,006	2	0,013
6	0,213	1	0,053	6	0,002	6	0,006
1	0,196	6	0,049	1	0,001	1	0,004
5	0,025	4	0,001	5	0,000	5	0,000
4	0,002	5	0,000	4	0,000	4	0,000

Tabla 5: Trayecto B, Vía 2. Relevancia de Variables (variabilidad 95 %) 2005-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

PCA		FA		OWA		MA-OWA	
Var.	Relevancia	Var.	Relevancia	Var.	Relevancia	Var.	Relevancia
2	0,242	4	0,571	4	0,439	4	0,153
3	0,229	2	0,339	2	0,024	2	0,032
5	0,203	6	0,023	3	0,002	3	0,007
1	0,196	1	0,010	5	0,001	5	0,005
6	0,079	5	0,006	1	0,001	1	0,004
4	0,001	3	0,000	6	0,000	6	0,001

Tabla 6: Trayecto B, Vía 12 (conjunta). Relevancia de Variables (variabilidad 95 %) 2005-2010: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

En las Tablas 4, 5 y 6 se muestran los resultados obtenidos para este trayecto: en la Vía 1 mediante PCA, que siempre es el método que muestra menos poder de discriminación, las variables Var.2, Var.3, Var.5 y Var.1 son las más destacadas, con FA claramente la Var.3 es la más relevante, seguida de la Var. 2. Aplicando los operadores OWA y MA-OWA: la Var.3 es la relevante seguida de la Var.2 de forma testimonial respecto a la primera. En la Vía 2, los resultados son similares con la aparición de la Var. 6 en la tercera posición mediante PCA. Para la Vía 12 (conjunta) 2, 3, 5 y 1 son las variables más relevantes, mediante FA son las Var.4 y Var 2., siendo estas dos las variables obtenidas mediante OWA y MA-OWA.

En las Figuras 1, 2 y 3 se puede observar una evolución temporal de la relevancia de cada variable en función de los criterios empleados: PCA muestra una evolución marcada por pocas variaciones, y salvo las Var. 4 y Var.6, las demás tiene una relevancia similar a lo largo del tiempo. Mediante FA, las Var.2 y Var. 3 marcan una evolución con grandes variaciones. Mediante OWA y MA-OWA la variable más significativa, de forma clara, es la Var. 2. En la vía 2 el comportamiento de PCA es similar al caso anterior, pero aquí las no consideradas son las Var, 4 y Var.5. FA muestra que las variables más importantes son Var. 2, Var.3 y Var.4 durante todo el periodo con fluctuaciones importantes. Finalmente en esta vía las variables más importantes mediante OWA y MA-OWA son la Var.2 y Var.3 al principio del periodo para luego decaer, a la mitad del periodo la Var4. tomó una relevancia media.

Si se considera la Vía 12 (conjunta): para PCA las Var. 4 y Var.6 son muy poco relevantes en comparación con el resto que presentan un nivel similar a lo largo del periodo considerado. Con FA la Var.2 es la más destacada en todo el periodo pero con fluctuaciones. La Var.3 también es relevante en el periodo, y la Var.4 se ha hecho relevante en la segunda mitad del periodo. OWA y MA-OWA muestran que la variable Var. 2 es la más relevante a lo largo del periodo, con relevancia de Var. 3 al principio del mismo y Var.4 al final.

Esta evolución refleja el proceso de maduración de la infraestructura, los distintos eventos o problemas que han ido apareciendo y las actuaciones efectuadas, así como el éxito alcanzado, para corregir las mismas. Si en vez de usar Ec. 1 se usará una función lineal, los resultados obtenidos son similares con la diferencia de valores de relevancia a través de OWA y MA-OWA con algo menos de diferencia entre las variables.

5. Conclusiones

El presente trabajo es una aplicación práctica en el dominio de la conservación de infraestructuras de conocidas técnicas: Análisis de Componentes Principales (PCA), Análisis de Factores (FA) y Operadores difusos como OWA y MA-OWA para la selección de características. PCA y FA permiten obtener una clasificación sobre la relevancia de cada variable, y OWA y MA-OWA permiten generar soluciones globales a partir de los mismos. El problema planteado consistía en conocer la relevancia de las variables consideradas en los procesos de auscultación de la vía ferroviaria, y así dar soporte a futuras herramientas de ayuda a la toma de decisiones de mantenimiento.

A modo de resumen, las variables más importantes obtenidas han sido las variables Var.2 y 3. Este resultado confrontado con los expertos del dominio ha sido satisfactorio por cuanto éstos consideran las Var. 1, 2 y 3 cómo las más relevantes según su experiencia y modo de hacer.

Por otro lado, se ha hecho un análisis temporal de la evolución de la relevancia de estas variables que también permite *diagnosticar* la evolución de la infraestructura y si las actuaciones han tenido efectividad en la línea deseada.

Agradecimientos

Los autores quieren expresar su agradecimiento a VIAS y Construcciones S.A. por la ayuda en el desarrollo de este trabajo.

Referencias

[1] P. Bonissoene, Y.-T. Chen, K. Goebel, and P. Khedkar, "Hybrid soft computing systems: industrial and commercial applications," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1641–1667, September 1999.

[2] F. Karray and C. d. De Silva, *Soft Computing and Intelligent Systems Design. Theory, Tools and Applications*. Addison Wesley, 2004.

[3] A. Mirabadi and S. Sharifian, "Application of association rules in iranian railways (rai) accident data analysis," *Safety Science*, vol. 48, pp. 1427 – 1435, 2010.

[4] Y.-H. Cheng and L.-A. Yang, "A fuzzy petri nets approach for railway traffic control in case of abnormality: Evidence from taiwan railway system," *Expert Systems with Applications*, vol. 36, pp. 8040 – 8048, 2009.

[5] A. Fay, "A fuzzy knowledge-based system for railway traffic control," *Engineering Applications of Artifficial Intelligence*, vol. 13, pp. 719 – 729, 2000.

[6] H.-r. Dong, S.-g. Gao, B. Ning, and L. Li, "Extended fuzzy logic controller for high speed train," *Neural Comput & Applic*, 2011.

[7] S. Sezer and A. E. Atalay, "Dynamic modeling and fuzzy logic control of vibrations of a railway vehicle for different track irregularities," *Simulation Modelling Practice and Theory*, vol. 19, pp. 1873 – 1894, 2011.

[8] J. Rabatel, S. Bringay, and P. Poncelet, "Anomaly detection in monitoring sensor data for preventive maintenance," *Expert Systems with Applications*, vol. 38, pp. 7003 – 7015, 2011.

[9] CIBIC, "Conservación de Infraestructuras Basada en Inteligencia Computacional." *MICINN*, vol. CIT-460000-2009-46, 2009-2010.

[10] H. Liu and H. Motoda, *Feature Extraction, Construction and Selection. A data mining perpective*. Kluwer Academic Publishers, 1998.

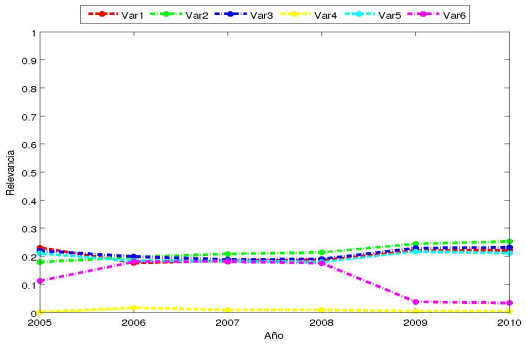
[11] I. Jolliffe, *Principal component analysis. (Second Edition)*. Springer-Verlag, 2002.

[12] R. R. Yager, "Families of owa operators," *Fuzzy Sets and Systems*, vol. 59, no. 2, pp. 125 – 148, 1993.

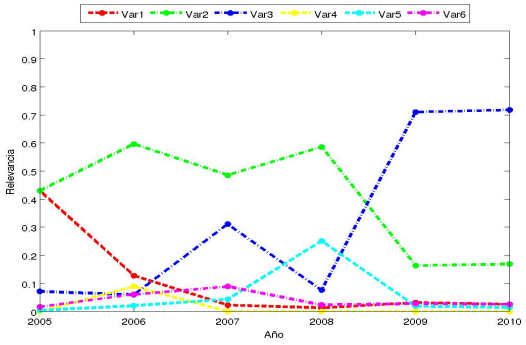
[13] D. Filev and R. R. Yager, "On the issue of obtaining owa operator weights," *Fuzzy Sets and Systems*, vol. 94, no. 2, pp. 157 – 169, 1998.

[14] J. Peláez, J. Doña, and J. Gómez-Ruiz, "Analysis of owa operators in decision making for modelling the majority concept," *Applied Mathematics and Computation*, vol. 186, no. 2, pp. 1263 – 1275, 2007.

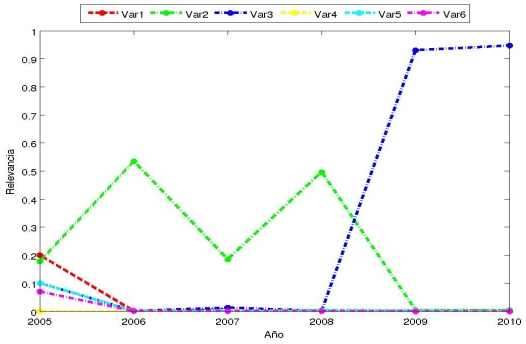
[15] J. I. Peláez and J. M. Doña, "Majority additive - ordered weighting averaging: A new neat ordered weighting averaging operator based on the majority process," *International Journal of Intelligent Systems*, vol. 18, no. 4, pp. 469–481, 2003.



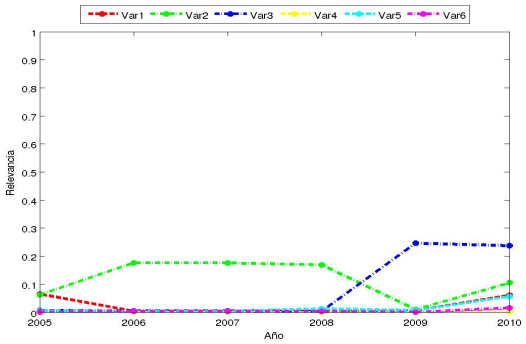
(a) PCA



(b) FA

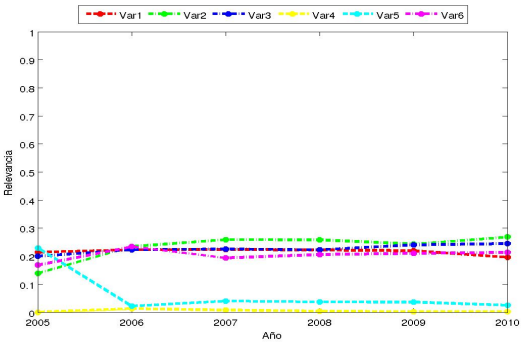


(c) OWA

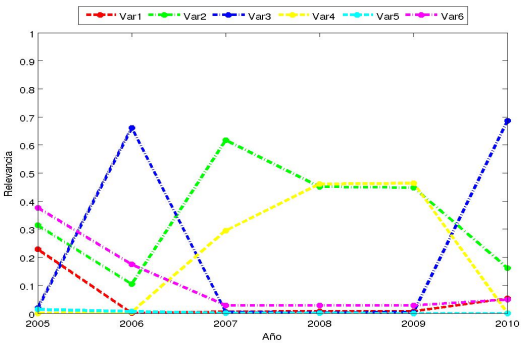


(d) MA-OWA

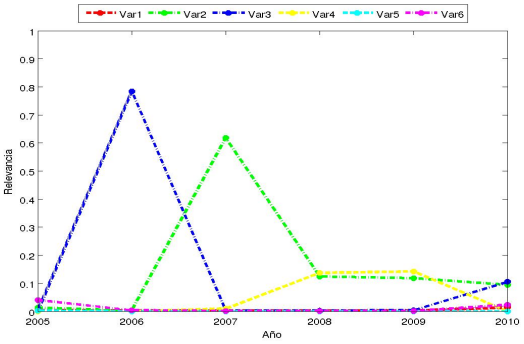
Figura 1: Trayecto B, Vía 1. Evolución de la Relevancia de Variables para variabilidad del 95 % mediante: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA



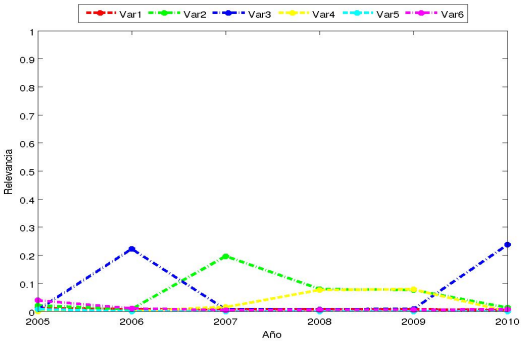
(a) PCA



(b) FA

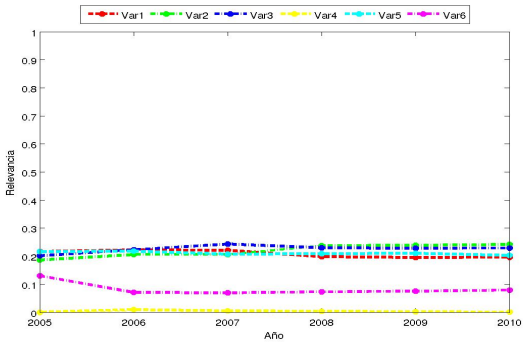


(c) OWA

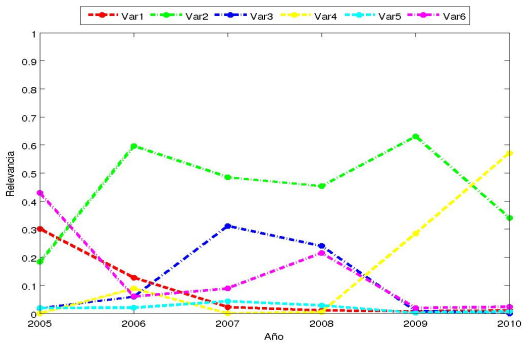


(d) MA-OWA

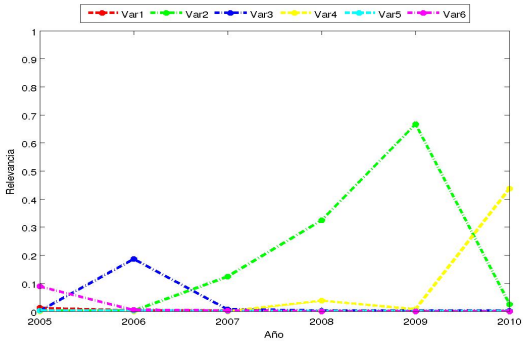
Figura 2: Trayecto B, Vía 2. Evolución de la Relevancia de Variables para variabilidad del 95 % mediante: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA



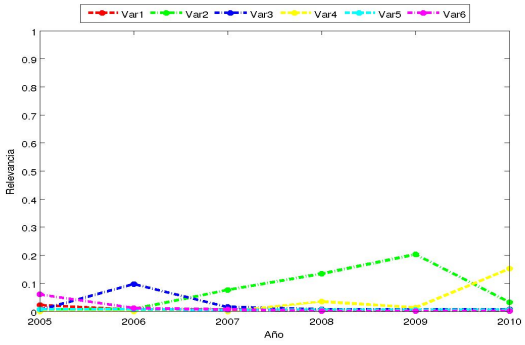
(a) PCA



(b) FA



(c) OWA



(d) MA-OWA

Figura 3: Trayecto B, Vía 12 (conjunta). Evolución de la Relevancia de Variables para variabilidad del 95 % mediante: (a) PCA, (b) FA, (c) OWA y (d) MA-OWA

ÁRBOLES DE DECISIÓN NO SUPERVISADOS DIFUSOS: APLICACIÓN A UN CENTRO DE ATENCIÓN DE LLAMADAS DE EMERGENCIA

Francisco Barrientos[1], **Gregorio Sáinz**[2,1], **Rubén García**[1]

[1]*CARTIF Centro Tecnológico, Parque Tecnológico de Boecillo 205*, {frabar,rubgar}@cartif.es
[2]*Universidad de Valladolid, E. I. Industriales, Dept. Ingeniería de Sistemas y Automática*, gresai@eii.uva.es

Resumen

Este artículo describe la aplicación de árboles de decisión no supervisados difusos (*FUDT*, de sus siglas en inglés) a un caso real: un centro de atención de llamadas de emergencia.

El principal problema en la gestión de un centro de llamadas de cualquier tipo es el de la planificación de los recursos, siendo el principal recurso los operadores que atienden las llamadas. Mediante la aplicación de *FUDT* se pretende obtener un sistema de apoyo a la toma de decisiones que ayude en la planificación de recursos, manteniendo un nivel de compromiso entre eficiencia y calidad de servicio.

La utilización de técnicas no supervisadas ha permitido caracterizar los distintos tipos de días en función de determinadas variables, solventando la disponibilidad de conocimiento experto. Los índices de calidad empleados para comparar los resultados obtenidos, proporcionan los criterios necesarios para la selección de la mejor solución.

Palabras Clave: Árboles de decisión no supervisados difusos, clustering difuso, emergencias.

1 INTRODUCCIÓN

En la actualidad, los centros de atención de llamadas son algo habitual en diversas áreas y actividades: atención al cliente, etc. Cada caso tiene sus particularidades, pero todos ellos deben gestionar sus recursos de manera que se alcance un equilibrio entre la eficiencia del centro y el acuerdo de nivel de servicio (*SLA*) con los usuarios.

El problema que subyace en un sistema de atención de llamadas es el del dimensionamiento. Deben establecerse los recursos, tanto técnicos como humanos, necesarios para dar un servicio de calidad y suficientes, de acuerdo al SLA establecido. El recurso principal en un sistema de estas características es el número de operadores disponibles para la atención de llamadas. Es necesario un cierto grado de sobredimensionamiento que permita atender picos de demanda puntuales, pero con mesura ya que en caso contrario supondría un coste añadido y restaría eficiencia al sistema.

Un centro de tamaño medio-grande puede gestionar miles de llamadas al día, y cada una debe ser atendida en unos pocos segundos [13]. La capacidad de atención de un centro depende fundamentalmente de la cantidad de personas que atienden las llamadas. La calidad del servicio suele medirse considerando la cantidad de llamadas perdidas. Debido a los costes de personal, la calidad del servicio se contrapone a la obtención de beneficios.

La planificación del personal es un medio para llegar a un compromiso entre costes y servicio. El método más usado es simular el centro mediante la teoría de colas [11]. En la mayoría de los casos, la predicción para determinar la cantidad de operadores necesarios se realiza de forma no automática, por lo que existe una dependencia del concurso de un experto humano. Las técnicas de predicción de series temporales pueden aplicarse en este contexto [5], que junto con los árboles de decisión difusos no supervisados, permiten al experto tomar decisiones más informadas.

En este trabajo se propone la utilización de FUDT aplicado a la caracterización de los distintos tipos de días, de tal manera que las reglas difusas obtenidas del árbol permitan establecer una relación entre los tipos de días y su carga de trabajo asociada. Los resultados obtenidos muestran que la propuesta genera una descripción lingüística de los días, útil para la asignación de recursos.

Este artículo está organizado de la siguiente manera. La Sección 2 describe las características de un centro de atención de llamadas de emergencia. La Sección 3 introduce la teoría asociada a los árboles de decisión y sus variantes. La Sección 4, la metodología empleada y la Sección 5 los resultados obtenidos. Finalmente, la Sección 6 presenta las conclusiones y futuras líneas de investigación.

2 EL CENTRO DE LLAMADAS DE EMERGENCIA

Este trabajo se centra en un tipo concreto de centro de llamadas: las llamadas de emergencia. Por tanto, se debe prestar especial atención a maximizar la disponibilidad del servicio para evitar la pérdida de llamadas, mientras que los aspectos económicos quedan en un segundo plano.

Un centro de atención de llamadas de emergencia está diseñado para responder eficientemente a llamadas individuales, asistiendo a las personas que realizan las llamadas y redirigiéndolos, si es necesario, al proveedor de servicios de emergencia adecuado [10] (e.g. sanidad, policía, bomberos, etc.), tal y como ilustra la Figura 1.

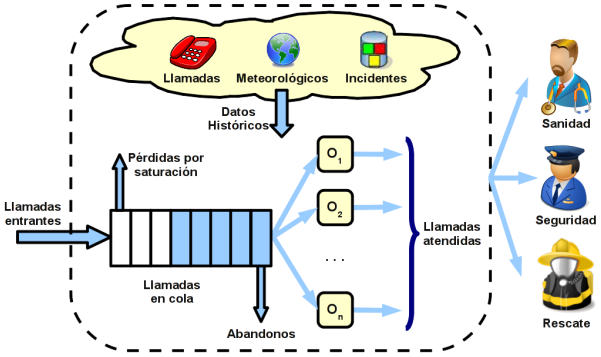


Figura 1: Esquema del sistema de atención de llamadas de emergencia.

El objetivo es sacar partido de la información y los datos recogidos por el centro de atención de llamadas. La extracción de dicho conocimiento puede ser útil en la toma de decisiones relacionadas con la asignación de recursos, gracias a la descripción lingüística de los diferentes tipos de días que se pueden presentar durante el normal funcionamiento del centro.

Los datos de entrada empleados para la caracterización de los tipos de días y la generación del conjunto de reglas difusas, son los datos recogidos en el propio centro de atención de llamadas (variables internas y parámetros recogidos por la centralita telefónica), así como otras variables externas, que a juicio de los gestores del sistema pueden tener influencia en la carga de trabajo, especialmente las condiciones atmosféricas (ver Tabla 1).

"No_Incidentes" es el porcentaje de llamadas diarias correspondiente con llamadas huecas, agradecimientos, bromas, etc. "Información" son las llamadas solicitando algún tipo de información. "Incidentes", es el porcentaje de llamadas que realmente se corresponden con incidentes que deben ser atendidos por algún tipo de servicio de emergencias.

Este último grupo a su vez se divide en: "Sanitarios" cuando se requiere atención de los servicios médicos, "Seguridad" si deben intervenir las Fuerzas y Cuerpos de Seguridad del Estado, "Salvamento y Rescate" (S.&R.) cuando deben intervenir los bomberos, Protección Civil, etc. y "Servicios Básicos" (también denominado "Otros") cuando se trata de otros asuntos, e.g. problemas relacionados con los suministros o fugas de agua, gas, luz, etc.

Tabla 1: Resumen de las variables de entrada.

NOMBRE	MÍN.	MÁX.	MED.	DES. E.
No_Incid. (%)	64.90	92.70	83.38	4.33
Inform. (%)	0.30	16.30	1.89	1.35
Incidentes (%)	6.00	31.20	14.72	4.34
Sanitarios (%)	33.70	75.20	57.03	5.94
Seguridad (%)	18.60	57.10	36.30	5.65
S. & R. (%)	0.30	24.50	5.04	3.69
Serv. Bás. (%)	0.00	9.90	1.64	1.35
Ll_Noche	8.00	495.00	163.33	91.83
Ll_Madrugada	0.00	214.00	33.44	21.89
Ll_Mañana	92.00	722.00	245.76	98.60
Ll_Mediodía	109.00	835.00	360.07	132.07
Ll_Tarde	52.00	770.00	237.51	118.68
Ll_Total	1658.00	9504.00	4788.09	1608.71
T_Med	-3.37	28.73	11.79	7.68
T_Max	1.30	38.90	19.42	9.01
T_Min	-11.90	17.90	3.92	6.36
Lluvia	0.00	192.20	6.10	15.07
Nieve	0.00	6.00	0.17	0.67
Niebla	0.00	6.00	0.54	1.12
Día	1.00	7.00	4.00	2.01
Mes	1.00	12.00	6.01	3.55

El siguiente grupo ("Llamadas Medianoche", "Llamadas Madrugada", "Llamadas Mañana", "Llamadas Mediodía", "Llamadas Tarde" y "Llamadas Total") se corresponde con la distribución de llamadas a lo largo del día, y el número total de llamadas del día.

Respecto a las variables externas, y considerando sólo las provincias de la comunidad autónoma de Castilla y León, "T_Med", "T_Max" y "T_Min" son las temperaturas media, máxima y mínima del día, respectivamente. "Lluvia" es la precipitación diaria acumulada, mientras "Nieve" y "Niebla" indican el número de provincias en las que se han producido dichos fenómenos. Por último, "Día" y "Mes" no necesitan mayor explicación.

3 ÁRBOLES DE DECISIÓN NO SUPERVISADOS DIFUSOS

El hecho de que dos de los "*Top Ten Algorithms in Data Mining*" [17] sean algoritmos basados en árboles demuestra la amplia popularidad de estos métodos en el campo de la minería de datos. Los árboles de decisión son, quizá, uno de los paradigmas más utilizados en el mundo del aprendizaje automático [16].

Un caso particular de los árboles de decisión son los árboles difusos, que trabajan con variables lingüísticas. Recientemente, se ha desarrollado una nueva variante de los árboles de decisión: los árboles de decisión no supervisados [2]. Son una solución híbrida entre los árboles y las técnicas de *clustering*.

De esta manera se combinan las ventajas de ambos métodos: la clasificación de los datos sin conocimiento previo del clustering y la facilidad de interpretación de los resultados de los árboles de decisión, ya que cada nodo hoja representa un cluster y la ruta desde la raíz a esa hoja, una regla de clasificación de la forma *if-then*.

La selección del atributo a considerar en cada nodo se hace de acuerdo a una medida de *heterogeneidad* [2], es decir, de manera que la heterogeneidad del conjunto de datos sea máxima con respecto a ese atributo.

$$H_a = -\left(\sum_{i,j} \mu_{ij}(1-\mu_{ij}^a) + \mu_{ij}^a(1-\mu_{ij})\right), \qquad (1)$$

donde μ_{ij} es la similitud entre los puntos X_i y X_j considerando todos los atributos, y μ_{ij}^a es la misma similitud considerando todos los atributos excepto a (f_a).

Se deben tener en cuenta diversos valores críticos para la construcción de UDT (*Unsupervised Decision Trees*):

- Umbral de Heterogeneidad (*IT: Inhomogeneity Threshold*) que está relacionado con la cantidad de información contenida en el conjunto de datos. Es mínimo cuando los datos son completamente homogéneos.
- Umbral de Tamaño de Segmento (*SST: Segment Size Threshold*) que determina cual puede ser el número mínimo de elementos en un nodo hoja o terminal.
- Umbral de separación de segmentos (*SGT: Segment Gap Threshold*) es la distancia entre dos segmentos de un mismo nodo, en función del rango de valores del atributo.

Es decir, si un nodo excede el umbral *IT*, debe ser dividido en segmentos de al menos *SST* elementos, siempre y cuando la distancia entre los segmentos sea mayor que *SGT*. Cada segmento será un nuevo nodo hoja. Por último, se debe establecer el *Tipo de árbol*, es decir, si se quiere que el árbol resultante sea binario o no.

Al considerar la versión difusa de UDT (FUDT: *Fuzzy UDT*), los parámetros *SGT* y *Tipo de árbol* carecen de sentido. Será la cardinalidad de las variables difusas de entrada la que determine el número de hijos de un nodo, en función de las etiquetas lingüísticas de la variable correspondiente. De esta manera, la complejidad del problema se reduce considerablemente (ver Tabla 2).

Tabla 2: Diferencia en el número de experimentos considerando las versiones difusas y no difusas de los árboles.

	UMBRAL	**LÍM. INF.**	**LÍM. SUP.**	**INC.**	**# EXP.**
UDT	IT	0.10	0.40	0.10	4
	SST	4	30	2	14
	SGT	0.3	0.9	0.1	7
	Árbol	Binario	n-ario	-	2
	Total				**784**
FUDT	IT	0.10	0.40	0.10	4
	SST	4	30	2	14
	Total				**56**

3.1 ÍNDICES DE CALIDAD

Al emplear un método no supervisado, no es posible calcular el error cometido en la clasificación. Por eso es necesario emplear algún índice de calidad, o combinación de ellos, para poder determinar la bondad de los resultados obtenidos.

Concretamente, se han empleado índices de calidad relacionados con las características y prestaciones de los árboles: *T-measure* [16], *IQN-T* [4], *IC* [1] y, debido a la relación de FUDT con el clustering difuso, medidas de calidad de los nodos o particiones: *SC* [3], *S* [3] y *XB* [18].

4 METODOLOGÍA

La metodología empleada es sencilla y está dividida en etapas claramente diferenciadas (ver Figura 2). El primer paso es recopilar los datos en bruto, sea cual sea su origen o formato. A continuación, se hace un pre-procesamiento, con el fin de homogeneizar formatos, normalizar los rangos de valores y recuperar datos ausentes.

Una vez que los datos están listos, se realiza una primera selección de características mediante análisis de componentes principales o PCA [12] [14]. Con ello se simplifica el problema, eliminando las variables que menos información aportan.

Lo siguiente es construir el FUDT con el conjunto de variables seleccionadas del paso anterior. El árbol se encarga de realizar una segunda selección de características de manera implícita, determinando qué atributos tienen mayor relevancia y en qué orden deben ser considerados.

Por último, se analizan los resultados, en base a los indicadores que hemos mencionado previamente, para determinar cuál es la solución elegida.

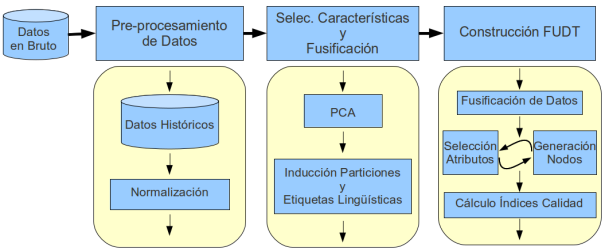


Figura 2: Representación de la metodología empleada.

5 RESULTADOS

El particionado difuso de la entrada se realizó mediante la herramienta KBCT[1] [8]. La Tabla 3 muestra los resultados obtenidos, utilizando los parámetros de configuración por defecto.

Tabla 3: Resultados del particionado difuso de la entrada.

VARIABLE	ÍND. CHEN	MÉTODO	#	CENTROS
No_Incid.	0.76	Kmeans	3	(77, 82.5, 87)
Inform.	0.99	HFP	3	(3.5, 11.5, 16.3)
Incidentes	0.78	Kmeans	3	(11.5, 17, 22.5)
Sanitarios	0.73	Kmeans	3	(49, 56, 64)
Seguridad	0.72	Kmeans	3	(28, 35.5, 42.5)
S. & R.	0.95	HFP	3	(7, 22.5, 24.5)
Serv. Bás.	0.92	HFP	3	(2.1, 7.2, 9.9)
Ll_Noche	0.77	Kmeans	3	(98, 205, 345)
Ll_Madrug.	0.96	HFP	3	(50, 165, 214)
Ll_Mañana	0.87	HFP	3	(240, 610, 722)
Ll_Mediod.	0.78	Kmeans	3	(230, 400, 560)
Ll_Tarde	0.87	HFP	3	(230, 650, 770)
Ll_Total	0.80	Kmeans	3	(3100, 5200, 7050)
T_Med	0.78	Kmeans	3	(4, 12, 22)
T_Max	0.78	Kmeans	3	(10, 18.5, 30.5)
T_Min	0.77	Kmeans	3	(-4, 3, 11)
Lluvia	0.98	HFP	3	(30, 110, 192.2)
Nieve	1	Regular	7	(1, 2, 3, 4, 5, 6, 7)
Niebla	1	Regular	7	(1, 2, 3, 4, 5, 6, 7)
Día	1	Regular	7	(1, 2, 3, 4, 5, 6, 7)
Mes	1	Regular	12[a]	(1, 2, 3, ..., 12)

[a] Valor estimado, ya que KBCT sólo trabaja en el rango $[2,9]$.

KBCT genera automáticamente todas las combinaciones posibles entre los tres métodos (HFP [6] [7], Regular y Kmeans [9]) y las particiones posibles en el intervalo $[2,9]$. Comparando los distintos casos y según los valores de determinados índices, se obtiene el número óptimo de funciones de pertenencia para cada variable y el método empleado para generarlas. Todas las funciones de pertenencia obtenidas son triangulares y las particiones completas, es decir, se cubre todo el rango de valores de las variables.

[1] Actualmente, la herramienta KBCT ha sido rebautizada como GUAJE.

En la construcción del FUDT, los valores de *IT* y *SST* se establecen experimentalmente para cada problema. Se ha encontrado que más allá de un cierto valor de los umbrales (ver Tabla 2), el árbol se colapsa, no creándose ningún nodo a parte del nodo raíz. Por este motivo, esos valores no se han considerado y el trabajo se centra en los casos realmente relevantes. Para *IT*, el valor máximo es 0.40 y para *SST* es 30. El valor inferior de *SST* se ha fijado en 4 elementos porque un segmento con menos de esos elementos puede ser considerado irrelevante o puede estar afectado por ruido o valores atípicos.

Por tanto, el árbol ha sido reconstruido fijando un valor de *IT* y variando el valor de *SST*. Así se evita la influencia de los valores extremos en la normalización de los índices. Las Figuras 3 y 4 muestran de manera gráfica los resultados tras la normalización.

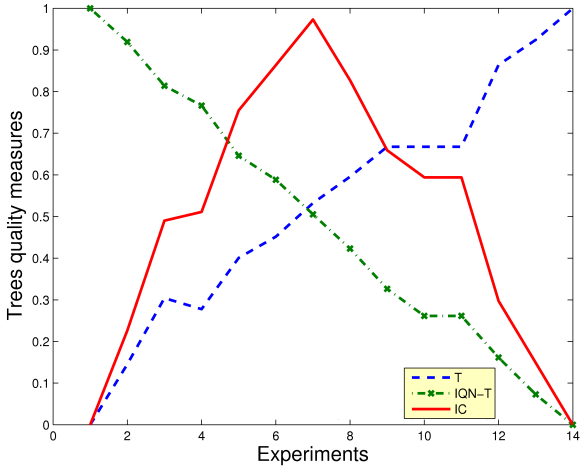


Figura 3: Medidas de calidad del árbol T, $IQN-T$ e IC.

La medida T, con $T \in [0,1)$, asigna un valor mayor a aquellos árboles que son más compactos, de manera que se alcanza el máximo ($T \simeq 1$) cuando todos los nodos hoja cuelgan directamente del nodo raíz.

Por otro lado, el valor de IQN_T es máximo cuando el árbol está completamente expandido. Dado que un árbol demasiado expandido puede ser menos eficiente debido al efecto del sobreaprendizaje [16] [15] y uno demasiado compacto puede tener un error de clasificación demasiado alto, *IC* proporciona un valor de equilibrio entre los dos anteriores, como muestra la Figura 3.

En el caso de los índices de calidad (Figura 4), un valor bajo de *SC* indica una mejor partición y el número óptimo de clusters debe minimizar el valor del índice *XB*. Se puede deducir, por tanto, que los mejores árboles son los más expandidos, puesto que los primeros experimentos tienen un umbral de tamaño de segmento menor. Como en el caso anterior, estos árboles pueden sufrir de falta de capacidad de generalización o sobreaprendizaje.

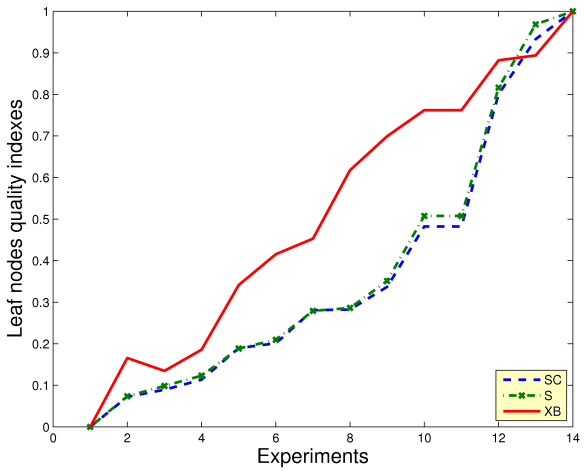


Figura 4: Índices de calidad de los clusters *SC*, *S* y *XB*.

El valor máximo de *IC* (ver Figura 3) se alcanza con un valor de umbral de heterogeneidad de 0.40 y un umbral de tamaño de segmento de 16. Para este experimento, el valor de *SC* y *XB* es razonablemente bajo, así que este parece un buen candidato para su estudio en profundidad.

El árbol tiene 37 nodos en total, de los que 22 son nodos hoja o terminales, y la profundidad es 5. Esto significa que como mucho habrá 22 clases diferentes, que se describen mediante reglas con 5 antecedentes como máximo. Por tanto, dichas reglas son compatibles con los criterios de legibilidad e interpretabilidad porque, según la literatura científica [19], se considera que 7 ± 2 antecendentes es el límite para que una persona pueda manejar adecuadamente las reglas.

La caracterización de los diferentes tipos de días encontrados se muestra en la Tabla 4. Por ejemplo, los nodos más poblados son el C9 y el C13. Su descripción, según los atributos lingüísticos, indica que el primero se corresponde con días laborables de temperatura suave, con un número total de llamadas bajo, de las cuales la mayoría estarían relacionadas con *Incidentes* y un número bajo de llamadas estarían relacionadas con *No_Incidentes*. El segundo nodo en cambio, caracteriza a los días no laborables con altas temperaturas, un número de llamadas bajo, pero igualmente distribuidas entre *Incidentes* y *No_Incidentes*.

6 CONCLUSIONES

En este trabajo se ha considerado una versión difusa de los árboles de decisión no supervisados, para hacer frente al problema de la asignación de recursos en un centro de atención de llamadas de emergencia. Con ello, se obtiene una caracterización en términos lingüísticos, dentro del dominio considerado, de los tipos de días en función de la carga de trabajo y los recursos necesarios para cada uno.

Tabla 4: Descripción de los Cluster-Hoja basada en características interpretables del dominio.

CLUSTER	MES												TIPO DE DÍA	LLAMADAS POR TURNO				TEMP. MED.	INCID.	NO INCID.	CARGA TRABAJO			
	E	F	M	A	M	J	J	A	S	O	N	D		M	T	N	Total				Sanit.	Seg.	S&R	Otros
C1	10 %	90 %											Laboral	bajo	bajo	bajo	med	bajo	bajo	alto	alto	bajo	bajo	bajo
C2			100 %										Laboral	bajo	bajo	bajo	med	bajo	bajo	alto	med	bajo	bajo	bajo
C3											37 %	63 %	Laboral	bajo	bajo	med	med	bajo	bajo	alto	alto	med	bajo	bajo
C4	10 %	18 %	33 %	26 %								13 %	Festivo	bajo	bajo	bajo	bajo	bajo	med	bajo	med	alto	bajo	bajo
C5	100 %												Laboral	bajo	bajo	bajo	bajo	bajo	alto	bajo	alto	med	bajo	bajo
C6		85 %	15 %										Festivo	bajo	bajo	bajo	bajo	bajo	med	bajo	med	alto	bajo	bajo
C7											16 %	84 %	Festivo	bajo	bajo	bajo	bajo	bajo	alto	bajo	med	alto	bajo	bajo
C8	31 %	13 %	21 %	15 %							10 %	10 %	Laboral	bajo	bajo	bajo	med	bajo	bajo	alto	med	alto	bajo	bajo
C9	4 %	4 %	13 %	1 %	4 %	3 %		4 %	4 %	29 %	24 %	10 %	Laboral	bajo	bajo	bajo	bajo	med	med	bajo	med	alto	bajo	bajo
C10			11 %	11 %	31 %	8 %				27 %	12 %		Laboral	bajo	bajo	med	med	med	bajo	alto	med	med	bajo	bajo
C11			59 %	29 %	12 %								Laboral	bajo	bajo	med	med	med	bajo	alto	med	med	bajo	bajo
C12				21 %	37 %	10 %			3 %	13 %	16 %		Laboral	bajo	bajo	bajo	med	med	bajo	alto	med	med	bajo	bajo
C13					1 %	15 %	37 %	22 %	21 %	3 %	1 %		Festivo	bajo	bajo	med	bajo	alto	med	med	med	alto	bajo	bajo
C14					6 %	60 %	22 %	6 %		6 %			Laboral	bajo	bajo	med	med	alto	bajo	alto	med	med	bajo	bajo
C15					29 %	21 %	25 %	17 %		8 %			Laboral	bajo	bajo	med	med	alto	bajo	alto	med	med	bajo	bajo
C16		36 %		29 %	12 %	10 %							Festivo	bajo	bajo	med	alto	bajo	bajo	med	med	med	bajo	bajo
C17			47 %	17 %	8 %						3 %	25 %	Festivo	bajo	bajo	med	alto	med	bajo	alto	med	med	bajo	bajo
C18				100 %									Festivo	bajo	bajo	med	alto	med	bajo	alto	med	med	bajo	bajo
C19					76 %	24 %							Festivo	bajo	bajo	med	med	alto	bajo	alto	med	med	bajo	bajo
C20									40 %	12 %	47 %	41 %	Festivo	bajo	bajo	med	alto	med	bajo	alto	med	med	bajo	bajo
C21					7 %	36 %	7 %			10 %			Festivo	bajo	bajo	med	med	alto	bajo	med	med	med	bajo	bajo
C22					11 %	56 %	33 %						Festivo	bajo	med	alto	alto	alto	bajo	alto	med	med	bajo	bajo

Para intentar conseguir la mejor solución posible, se han combinado medidas de calidad de los árboles de decisión con índices de calidad de los clusters.

El método desarrollado pretende ayudar a los gestores del centro de atención de llamadas estableciendo la línea base de funcionamiento normal del sistema. Posteriormente, utilizando algunas estimaciones y predicciones de ciertas variables, el sistema caracterizará el día deseado, ayudando en la planificación de los recursos necesarios.

Los resultados obtenidos pueden ser mejorados en futuros trabajos, por ejemplo refinando los criterios considerados para la construcción y selección de los árboles. También sería posible obtener una versión optimizada de la base de conocimiento, siempre que se mantenga el nivel de interpretabilidad de los resultados.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la Junta de Castilla y León, a través de la Agencia de Protección Civil y Consumo.

Referencias

[1] F. Barrientos, G. Sáinz: Interpretable knowledge extraction from emergency call data based on fuzzy unsupervised decision tree. *Knowledge-Based Systems* 25, pp. 77–87, 2012.

[2] J. Basak, R. Krishnapuram: Interpretable hierarchical clustering by constructing an unsupervised decision tree. *IEEE Transactions on Knowledge and Data Engineering* 17, pp. 121–132, 2005.

[3] A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, R. F. Murtagh: Validity-guided (Re)Clustering with applications to image segmentation. *IEEE Transactions on Fuzzy Systems* 4, pp. 112–123, 1996.

[4] D. Fournier, B. Cremilleux: A quality index for decision tree pruning. *Knowledge-Based Systems* 15, pp. 37–43, 2002.

[5] R. García, J. M. Benitez, G. Sainz: Feature selection for time series forecasting: A case study. *Hybrid Intelligent Systems, Eighth International Conference on*, pp. 555–560, 2008.

[6] S. Guillaume, B. Charnomordic: A new method for inducing a set of interpretable fuzzy partitions and fuzzy inference systems from data. En: Springer-Verlag (Eds.), *Studies in Fuzziness and Soft Computing*, pp. 148–175, 2003.

[7] S. Guillaume, B. Charnomordic: Generating an interpretable family of fuzzy partitions. *IEEE Transactions on Fuzzy Systems* 12, pp. 324–335, 2004.

[8] GUAJE: `http://www.softcomputing.es/guaje/` Última visita: septiembre de 2011.

[9] J. A. Hartigan, M. Wong: A k-means clustering algorithm. *Applied Statistics* 28, pp. 100–108, 1979.

[10] H. Jasso, W. Hodgkiss, C. Baru, T. Fountain, D. Reich, K. Warner: Using 9-1-1 call data and the space-time permutation scan statistic for emergency event detection. *Government Information Quarterly* 26, pp. 265–274, 2009.

[11] G. Koole, A. Mandelbaum: Queueing models of call centers. An introduction. *Annals of Operations Research* 113, pp. 41–59, 2003.

[12] A. Malhi, R. X. Gao: PCA-based feature selection scheme for machine defect classification. *IEEE Transactions on Instrumentation and Measurement* 53, pp. 1517–1525, 2004.

[13] A. Mandelbaum, O. Garnett, M. Reiman: Designing a call center with impatient customers. *Manufacturing and Service Operations Management* 4, pp. 208–227, 2002.

[14] K. Mao: Identifying critical variables of principal components for unsupervised feature selection. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 35, pp. 339–344, 2005.

[15] J. Mingers: An empirical comparison of prunning methods for decision tree induction. *Machine Learning* 4, pp. 227–243, 1989.

[16] S. Mitra, T. Acharya: *Data mining: multimedia, soft computing, and bioinformatics*. Wiley, Nueva York, 2003.

[17] W. Wu, V. Kumar: *The top ten algorithms in data mining*. CRC Press, 2009.

[18] X. L. Xie, G. A. Beni: Validity measure for fuzzy clustering. *IEEE Trans. PAMI* 3(8), pp. 841–846, 1991.

[19] S. M. Zhou, J. Q. Ganb: Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets and Systems* 159, pp. 3091–3131, 2008.

Aplicación de Técnicas Borrosas en Sistemas Altamente No Lineales

Juan Carlos Sanz Villanueva [1], Alberto Herreros López [2]

[1] *Dpt. Ingeniería de Sistemas y Automática (Universidad de Valladolid, Spain)*, jucasvill@orange.es

[2] *Instituto de las Tecnologías Avanzadas de la Producción (Universidad de Valladolid, Spain)*, albher@eis.uva.es

Resumen

Este artículo tratará de poner de manifiesto las ventajas e inconvenientes de las técnicas de control borrosas, o "*fuzzy*", en sistemas altamente no lineales, aplicándolas al control del sistema Pendubot, consistente en un péndulo invertido con movimiento rotacional.

Se analizará el sistema, tanto física como matemáticamente, y se propondrá un controlador "borroso", que mantenga en equilibrio el sistema, así como el algoritmo de control, analizando su desarrollo para optimizar el tiempo de respuesta del controlador. Por último, se comparará los datos obtenidos con los que proporciona un controlador creado mediante la teoría clásica.

El objetivo, en definitiva, será diseñar un controlador para un sistema complejo sin utilizar complicados algoritmos matemáticos sino sencillas reglas lógicas que cualquier persona sería capaz de obtener razonando un poco.

Palabras Clave: Control Borroso, Lógica Borrosa, Pendubot, Control no lineal.

1 INTRODUCCIÓN

El desarrollo de los instrumentos de control para la industria avanza a pasos de gigante. Hace pocos años la gran mayoría de las empresas funcionaban gracias al trabajo y a los conocimientos de las personas y el grado de automatización de la industria era más bien pequeño. Sin embargo, hoy en día, se han cambiado los papeles; el número de personas en las empresas comienza a ser menor que el de máquinas, que por su tecnología podríamos considerarlas casi autónomas. Pero este avance rápido y continuo en las teorías de control no solamente lo tienen que asimilar los investigadores o los ingenieros, sino también las personas que trabajan en las fábricas ya que, en definitiva, son los que al final manejan esa nueva tecnología que aparece, y sin la que, tarde o temprano, las empresas desaparecen. Es en este nuevo campo, el de facilitar la asimilación de estos avances, donde el control borroso está tomando más auge en la industria.

Pero las técnicas borrosas no nacieron con ese objetivo. Cuando Lofti A. Zadeh [9], profesor de la Universidad de California en Berkeley, en 1965, presenta al mundo su trabajo sobre técnicas borrosas, lo hace con un objetivo claro: enlazar las matemáticas con el razonamiento humano. Es Ebrahim Mandani quién, a principios de los años 70, usa la lógica difusa en un control práctico, para así aprovechar la experiencia de los trabajadores en el control de las máquinas.

A partir de entonces se han realizado grandes esfuerzos para conseguir aplicar estas técnicas a control de laboratorio e industrial, ver [3]. Las reglas de inferencia son encontradas muchas veces por la experiencia humana en el manejo de los sistemas y otras con ayuda de otras técnicas como la redes neuronales que aprenden a controlar la planta y tras ello se realiza estudios de simplificación de reglas para entender su significado, ver [2,4].

El objetivo de este articulo es desarrollar un controlador borroso para una planta de laboratorio altamente no lineal, ver [6], el Pendubot. El control adaptativo propuesto en la práctica fue diseñado por [1]. Este autor linealiza el sistema en diferentes puntos de equilibrio y halla controladores óptimos en dichos puntos. Otros autores han propuesto control fuzzy para conseguir controlar la planta, ver [8,9]. Proponen reglas de inferencias que consigan controlar el sistema en todos los puntos de equilibrio. Este artículo propone unas reglas muy sencillas, basadas en la experiencia, y una optimización del algoritmo para hacerlo más rápido, ver [5]. Con ello se consigue que las mismas reglas sean útiles en todos los puntos de equilibrio del sistema.

El esquema del artículo es el siguiente. En la sección segunda se describe la práctica de laboratorio Pendubot y las ecuaciones matemáticas altamente no lineales que rigen el sistema. En la sección tercera se describe el controlador borroso con las reglas que se han aplicado para conseguir sencillez y velocidad de computación. En la sección cuarta se analizan los resultados obtenidos en la planta real y las conclusiones del trabajo.

2 DESCRIPCIÓN DEL SISTEMA.

PENDUBOT, objeto de estudio en este estudio, es un péndulo invertido rotacional de tipo electro-mecánico desarrollado por Mechatronic Systems Inc. El conjunto, como se puede observar en la figura 1, se puede dividir en tres bloques funcionales bien diferenciados:

- Sistema a controlar, formado por un péndulo invertido de dos eslabones conexionados entre ellos mediante una unión libre y unidos solidariamente a un motor que permite moverlos.
- Interfaz de relación entre el sistema péndulo y el controlador, encargado de tomar datos y de adecuar las señales de control, compuesto por encoders y servo-amplificadores.
- Controlador digital, consistente en un PC que contiene los algoritmos de control así como las tarjetas de adquisición y envío de señales de control al interfaz.

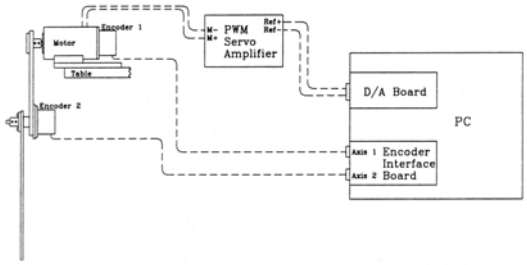


Figura 1: Esquema general del sistema.

2.1. ANÁLISIS MATEMATICO.

El sistema que sometemos a estudio es un péndulo invertido, formado por dos brazos unidos por articulaciones planas que permiten el giro relativo entre ellos, y que responde al siguiente esquema:

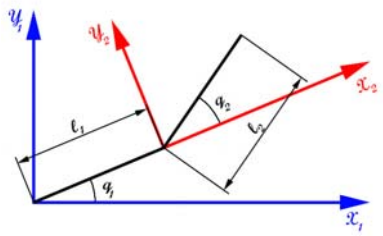


Figura 2: Sistemas de coordenadas utilizados y magnitudes principales.

Como se observa en la figura 2, vamos a utilizar dos sistemas de coordenadas para facilitar el cálculo de los vectores que nos definen tanto la posición como la velocidad de los puntos. El primero de los sistemas de coordenadas, S_1, va a estar solidario al punto de apoyo del péndulo, es decir, un sistema fijo al cual referiremos todos las magnitudes vectoriales del sistema. El segundo sistema de coordenadas, S_2, estará colocado en la unión U_2, solidario al eslabón E_1, lo que le convierte en un sistema relativo respecto al anterior pero que permite calcular la posición de los puntos del eslabón E_2 con facilidad.

Una vez definidos los sistemas de coordenadas y la relación entre ambos, estamos en disposición de abordar el cálculo de las ecuaciones que nos definen el comportamiento dinámico del sistema. Aplicando las ecuaciones de Euler-Lagrange que, basándose en la teoría de los trabajos virtuales, relacionan la energía de un sistema con el trabajo aplicado sobre él, y expresando los resultados en espacio de estados dado que, al tratarse de un sistema SIMO, es lo más conveniente, llegamos a las ecuaciones en variables de estado descritas en [1,5], cuyas variables de estado y sus derivadas corresponden a los valores de:

$$x_1 = q_1 \quad \text{(ángulo del eslabón } E_1\text{)} \qquad (1)$$

$$\dot{x}_1 = \dot{q}_1 \quad \text{(veloc. angular del eslabón } E_1\text{)} \qquad (2)$$

$$x_2 = q_2 \quad \text{(ángulo del eslabón } E_2\text{)} \qquad (3)$$

$$\dot{x}_2 = \dot{q}_2 \quad \text{(veloc. angular del eslabón } E_2\text{)} \qquad (4)$$

En dichas ecuaciones aparecen una serie de parámetros constantes $\theta_1 \ldots \theta_5$ dependientes de la longitud, la masa y los momentos de inercia de los eslabones que deben ser estimados. Mediante la realización de varios experimentos y utilizando Matlab para la realización de los cálculos se llega a la obtención de los siguientes valores para dichos parámetros:

$$\theta_1 = 0.0308 \text{ V*s}^2$$
$$\theta_2 = 0.0106 \text{ V*s}^2$$
$$\theta_3 = 0.0095 \text{ V*s}^2$$
$$\theta_4 = 0.2087 \text{ V*s}^2\text{/m}$$
$$\theta_5 = 0.0630 \text{ V*s}^2\text{/m}$$

Estos valores serán utilizados en las simulaciones y en el desarrollo de los controladores.

3 DISEÑO DEL CONTROLADOR BORROSO.

El control fuzzy o borroso nace en el año 1965 de la mano de Zadeh [9] y consiste en aplicar la experien-

cia, que posee el operador de una máquina, a la hora de controlarla automáticamente mediante un computador. En pocas palabras consiste en conseguir que el sistema de control se comporte como un *operador humano*, razonando de forma equivalente.

El esquema básico de un controlador borroso está formado fundamentalmente por tres pasos:

- Borrosificación.
- Aplicación de la Matriz de Inferencia.
- Desborrosificación.

3.1. BORROSIFICACIÓN

En este primer paso debemos definir las variables que vamos a tener, tanto las variables físicas que podemos cuantificar como las lingüísticas que vamos a utilizar en el proceso de control.

El objetivo del controlador que vamos a diseñar es mantener en una posición dada el segundo brazo del robot, es decir, conseguir que los ángulos q_1 y q_2 sean unos predefinidos y las velocidades angulares de los brazos, dq_1 y dq_2, nulas. Ahora bien, las posiciones de los brazos nos las proporcionan los encoders pero no las velocidades angulares, por lo que tendremos que calcularlas a partir de los datos de que disponemos. La forma de cálculo de la velocidad en el instante *k* va a ser la siguiente:

$$(dq_i)_k = \frac{(q_i)_k - (q_i)_{k-1}}{\Delta t} \qquad (6)$$

Para eliminar posibles errores que puedan cometer los encoders, pudiendo ocasionar desequilibrios en el controlador, aplicaremos un filtro a los datos calculados

$$(dq_i)_k = \frac{(dq_i)_k + (dq_i)_{k-1} + (dq_i)_{k-2}}{3} \qquad (7)$$

En nuestro caso utilizaremos tres conjuntos para cada variable de entrada donde "P" significa positivo, "Z" significa cero o nulo y "N" significa pequeño. Para la determinación de los valores de las constantes de cada una de las funciones vamos a utilizar la experiencia si bien, como en este caso no existe, nos apoyaremos en la experimentación, partiendo de unos valores lógicos y afinando después.

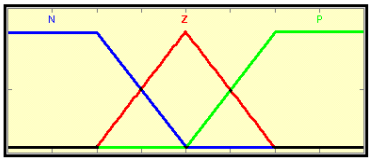


Figura 3: Conjunto borroso de las variables de entrada.

Como se puede observar, hemos definido las funciones de modo que se solapan, dando ese carácter de aproximación que hemos comentado anteriormente. Por ejemplo, si el error en la posición del eslabón 1 es 0.04 radianes, podemos decir que participa un 80% de "Z" y un 20% de "P", es decir, no es ni cero total ni positivo total. Esta relación se va a transmitir a las decisiones, participando en la salida tanto las reglas en las que aparece "si q1 es Z" como en las que aparece "si q1 es P", de modo que la salida tomará un valor intermedio entre el que tendría si fuese "Z" o "P" siendo la participación de cada una de ellas función de esos porcentajes.

En cuanto a la variable de salida vamos a definirla con cinco conjuntos: "PG" positivo grande, "PP" positivo pequeño, "Z" cero, "NP" negativo pequeño y "NG" negativo grande, determinando los parámetros experimentalmente como en caso anterior.

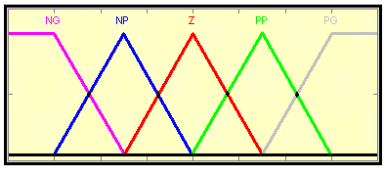


Figura 4: Conjunto borroso de las variables de salida.

3.2. LAS REGLAS DE INFERENCIA

El segundo paso a dar es la definición de las reglas de inferencia. Aquí la experiencia y la lógica entran en escena, diciendo que debe hacer el controlador ante las diferentes combinaciones de entradas. En nuestro caso particular aplicaremos la lógica para mantener el segundo eslabón verticalmente, tratando a la vez que el primero esté en la posición deseada. La estrategia de control que vamos a seguir va a ser dar prioridad al segundo escalón y una vez estabilizada su verticalidad tratar de llevar el primer escalón hasta la posición deseada.

Para el primer objetivo nos fijaremos en lo que haríamos para mantener un eslabón verticalmente en la palma de la mano. Si el eslabón se cae a la derecha instintivamente movemos la mano hacia la derecha y en función de la velocidad de caída lo hacemos más o menos velozmente; si cae de forma lenta, movemos la mano lentamente y viceversa. En el caso de estar vertical y quieto, no movemos la mano. Todas estas reglas van a componer nuestro Motor de Inferencia que nos determinará el valor borroso de la variable de salida. En sentencias si...entonces..., las reglas son las expuestas en la tabla 1

Tabla 1: Reglas de inferencia.

- *Si* **q_2** *es* **N** *y* **dq_2** *es* **N** *entonces* **u** es **NG**
- *Si* **q_2** *es* **N** *y* **dq_2** *es* **Z** *entonces* **u** Es **NP**
- *Si* **q_2** *es* **N** *y* **dq_2** *es* **P** *entonces* **u** Es **Z**
- *Si* **q_2** *es* **Z** *y* **dq_2** *es* **N** *entonces* **u** es **NP**
- *Si* **q_i** *es* **Z** *y* **dq_2** *es* **Z** *entonces* **u** es **Z**
- *Si* **q_2** *es* **Z** *y* **dq_2** *es* **P** *entonces* **u** es **PP**
- *Si* **q_2** *es* **P** *y* **dq_2** *es* **N** *entonces* **u** es **Z**
- *Si* **q_2** *es* **P** *y* **dq_2** *es* **Z** *entonces* **u** es **PP**
- *Si* **q_2** *es* **P** *y* **dq_2** *es* **P** *entonces* **u** es **PG**

O bien de forma matricial

		Q_2		
		N	*Z*	*P*
dq2	*N*	**NG**	**NP**	**Z**
	Z	**NP**	**Z**	**PP**
	P	**Z**	**PP**	**PG**

En cuanto a la forma gráfica, la representación consiste en crear una superficie tridimensional con los valores de salida para las diferentes entradas. En nuestro caso la superficie es la siguiente

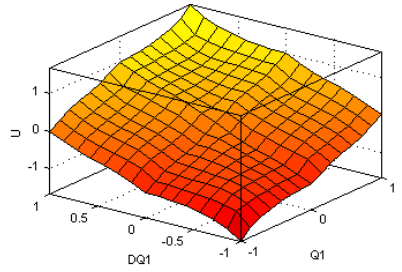


Figura 5: Superficie de control del segundo eslabón.

Ambas representaciones nos van a permitir observar que el conjunto de reglas de inferencia, es completo, consistente y continuo, propiedades que nos permiten considerarlo un buen motor de inferencia. Actuando de la misma manera para el primer eslabón, obtenemos La matriz de inferencia de este segundo motor va a ser la siguiente

		Q_1		
		N	*Z*	*P*
dq1	*N*	**PG**	**PP**	**Z**
	Z	**PP**	**Z**	**NP**
	P	**Z**	**NP**	**NG**

si bien las combinaciones extremas nunca se van a dar pues si el segundo eslabón está estable el movimiento del primer eslabón va a ser casi nulo. Para completar la descripción de este segundo conjunto de reglas vamos a representar la superficie de control.

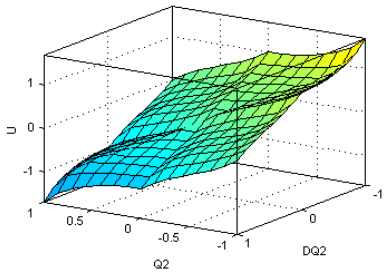


Figura 6: Superficie de control del primer eslabón.

Una vez definidas las reglas que van a constituir nuestro motor de inferencia deberemos traducirlas al lenguaje matemático, para así poder escribir los algoritmos de control que implementaremos en la computadora.

Dentro de la lógica borrosa vamos a tener tres operaciones básicas: Unión, Intersección y Complemento. La elección del conjunto de operaciones va a ser función del tiempo de cómputo que tenemos y de la naturaleza de las variables antecedentes. En nuestro caso las variables son puramente físicas, ángulos y velocidades, por lo que las operaciones que mejor se adecuan son

Unión: $\forall u \in X \quad \mu_{A\cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)$

Intersección: $\forall u \in X \quad \mu_{A\cap B}(x) = \mu_A(x) \cdot \mu_B(x)$

Complemento: $\forall u \in X \quad \mu_{A'}(x) = 1 - \mu_A(x)$

Además, este tipo de operaciones permite reducir bastante el tiempo de cálculo pues nos va a permitir juntar en la programación los algoritmos de borrosificación y decisión en un solo bloque de sentencias.

El proceso de cálculo de decisiones, consistirá en el cálculo del resultado de cada una de las reglas de inferencia y la Unión de todas esas conclusiones para la obtención del resultado final.

3.3. DESBORROSIFICACION

Este último paso va a consistir en la conversión del valor borroso de la variable de salida en un valor numérico que será el que se aplicará al elemento de control sobre el que actúa el controlador, es decir, el motor conectado al primer eslabón. En nuestro caso utilizaremos el método de las alturas dada la necesidad de velocidad en el controlador para poder conseguir realizar los cálculos dentro del periodo de muestreo deseado.

4 EXPERIMENTACION.

El último paso a dar es la comprobación empírica. En primer lugar, aprovechando que en este caso el sistema, aunque complejo, es tratable matemáticamente, haremos uso de la simulación, utilizando la herramienta Simulink de Matlab, para comprobar el funcionamiento del controlador y ajustar los parámetros del mismo. A continuación realizaremos además los siguientes cuatro experimentos:

- Control del péndulo en el punto q1= $\pi/2$.
- Comportamiento en un punto diferente al utilizado para el cálculo de los parámetros del controlador, como por ejemplo *q1*=1.3 radianes.
- Comportamiento ante perturbaciones (golpeando el péndulo para desestabilizarlo).
- Comportamiento ante cambios de referencia durante el control comprobando la capacidad de seguimiento de una referencia que tiene nuestro controlador borroso.

4.1. CONTROL EN EL PUNTO Q1= $\pi/2$

El primer objetivo es mantener verticalmente en equilibrio el sistema, ajustando los parámetros que utilizaremos en el resto de los experimentos y que nos permitirá demostrar la idoneidad del control en las diferentes situaciones a diferencia de los controladores "clasicos", como el adaptativo que debe calcularse para cada punto.

Los resultados para el primer experimento, como se observa en la figura 8, nos muestran el buen comportamiento que tiene el controlador borroso para mantener el sistema en equilibrio vertical, observando en la gráfica de la derecha, correspondiente a la señal de control, el poco esfuerzo que debe realizar.

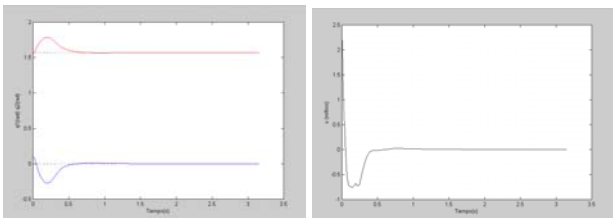


Figura 8: Control en q1= $\pi/2$.

4.2. COMPORTAMIENTO EN Q1=1.3 RADIANES.

En este caso trataremos de controlar el péndulo en un punto diferente al utilizado para ajustar el controlador. Los resultados podemos observarlos en la figura 9,

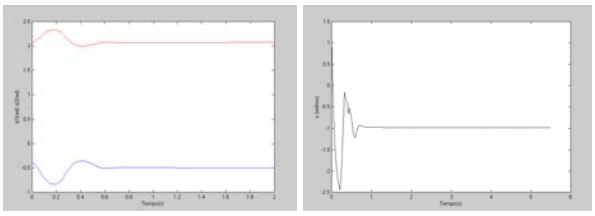


Figura 9: Control en q1=1.3 radianes.

pudiendo concluir que el comportamiento es tan bueno como el demostrado en el punto de equilibrio donde ajustamos los parámetros.

4.3. PERTURBACIONES

Someteremos ahora el sistema a perturbaciones en el punto de equilibrio, golpeando el eslabón 2 para sacarlo de la vertical. El resultado es el siguiente:

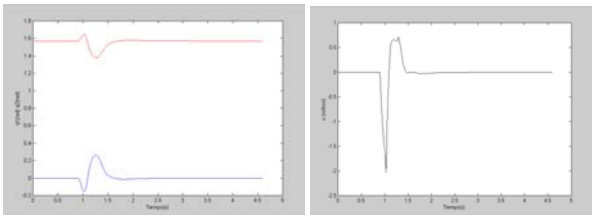


Figura 10: Control ante perturbaciones.

Volvemos a tener un buen comportamiento con, además, una señal de control adecuada.

4.4. CAMBIOS DE REFERENCIA DURANTE EL CONTROL

Por último variaremos la referencia y veremos el comportamiento que tiene para seguirla,

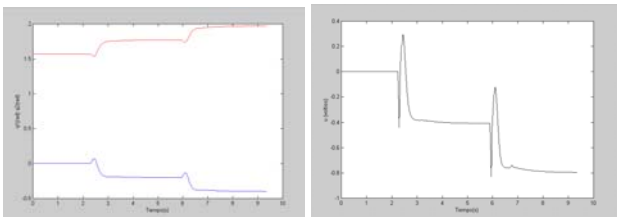


Figura 11: Control ante perturbaciones.

obteniendo un resultado satisfactorio como en los anteriores experimentos.

Una vez comprobado el comportamiento del controlador en las simulaciones pasaremos a realizar los experimentos en el sistema real. Para ello construimos el programa de control mediante la herramienta de programación Visual C++ y realizamos los mismos experimen-

tos. Los resultados obtenidos se pueden observar en la figura 12

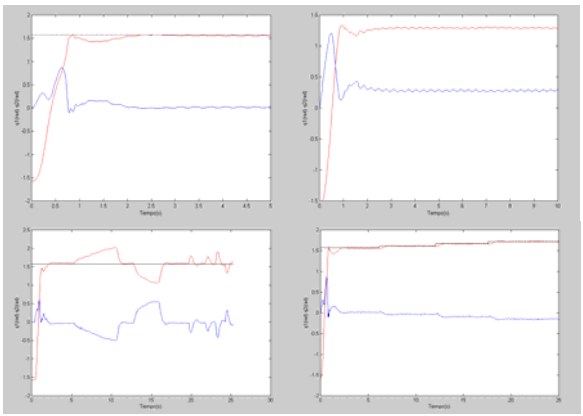


Figura 12: Experimentación real.

donde, al igual que en las simulaciones, el controlador responde de manera adecuada en cada uno de las diferentes situaciones. La única diferencia que presenta es una pequeña oscilación de los eslabones debida a que en el sistema real, a diferencia que en el simulado, tenemos una componente de rozamiento que obliga al control a aplicar un mayor esfuerzo inicial y le hace ser menos preciso. Ajustando un poco más los parámetros o utilizando variables de entrada y salida más complejas conseguiríamos eliminarlo.

5 CONCLUSION

Se ha conseguido un controlador borroso, basado en la experiencia, con unas reglas muy sencillas que consigue objetivos similares o mejores que el adaptativo propuesto en la práctica. El controlador es robusto en todos los puntos de equilibrio del sistema a diferencia del adaptativo que debe ir cambiando de estructura. En suma, constituye un buen ejemplo del poder de la lógica borrosa para captar la experiencia humana en el control de sistemas altamente no lineales.

Referencias

[1] D. J. Block, “Mechanical design and control of the Pendubot”, M.S thesis, Department of General Engineering, University of Illinois at Urbana-Champaign, 1996.

[2] M. Galende G. I. Sainz M. J. Fuente A. Herreros. Interpretability-accuracy improvement in a neuro-fuzzy ART based model of a DC motor. *Proceedings of the 17th World Congress The International Federation of Automatic Control* Seoul, Korea, July, 2008

[3] K. Passino, S. Yurkovich Fuzzy Control, Adisson Weskey Longman Inc, 1998.

[4] G. I. Sainz, L. A. Campillo, and M. J. Fuente. Accurate-interpretable based fuzzy modeling of a DC motor. In *Proceedings of the European Control Conference 2007*, pages 5169-5174, Kos, Greece, July 2007.

[5] J. C. Sanz Villanueva. Aplicación de Técnicas Borrosas en Sistemas Altamente no Lineales. Proyecto fin de carrera (Universidad de Valladolid).

[6] J.J Slotine L. Weiping Applied Nonlinear Control. Prentice Hall 1991

[7] X. Q. Ma, “Fuzzy control for an under-actuated robotic manipulator: Pendubot”, M.S thesis, Department of Mechanical Engineering, Concordia University, August 2001.

[8] X. Q. Ma and C.-Y. Su A New Fuzzy Approach for Swing up Control of Pendubot. *Proceeding of the American Control Conference, Anchorage*, May 2002

[9] L.A. Zadeh: Towards a theory of fuzzy systems. En: R.E. Kalman, N. DeClaris (Eds.), *Aspects of Network and System Theory*, Holt, Rinehart and Winston, Nueva York, pp. 469-490, 1971.

Takagi-Sugeno fuzzy modelling for control of greenhouse climate

Meriem Nachidi[1], **Fernando Tadeo**[2], **Francisco Rodriguez**[3], **Jose Luis Guzman**[3]

[1]*Dpto. de Ingeniería de Sistemas y Automática, Universidad Carlos III, Av. Universidad 30, 28911 Leganes, Madrid, Spain*, mnachidi@ing.uc3m.es
[2]*Dpto. de Ingeniería de Sistemas y Automática, Escuela de Ingenierías Industriales, Dr. Mergelina s/n, 47005 Valladolid, Spain*, fernando@autom.uva.es
[3]*Ingeniería de Sistemas y Automática, Universidad de Almería, Crta. de la Playa, s/n, 04120 Almería, Spain*, frrodrig@ual.es

Abstract

In this study, Takagi Sugeno fuzzy models are developed to describe the inside air climate of greenhouses that are artificially heated at night, and naturally ventilated during the day. Thus, based on physical models, a methodology is developed to derive Takagi-Sugeno models suitable for fuzzy controller design. This makes possible to include first-principle models, well developed in greenhouse systems, to precisely derive fuzzy controllers that ensure stability. To illustrate the technique, data from a greenhouse placed in Almería will be used to parameterize the model and to show that the obtained Takagi-Sugeno model describes the greenhouse climate dynamic satisfactorily, and then can be used to design controllers for climate control.

Keywords: Greenhouse, Takagi-Sugeno fuzzy models, modelling for control.

1 Introduction

The greenhouse climate is a very complex system in which the variables depend strongly on the outside climate conditions (solar radiation, temperature, wind speed and direction, humidity, etc), and the greenhouse design, with the internal conditions related through heat and mass transfer phenomena. The study and design of efficient greenhouse environmental controllers require to have a priori knowledge of the greenhouse climate models. It is clear that the scientific understanding of the greenhouse climate mechanism has advanced a lot by the availability of computers that allow to simulate the climate by means of static and dynamic greenhouse climate models. Whereas the pioneering works (for instance, [1], [5] and [6]) were rather simple by modern standards, the actual deterministic models allow the greenhouse climate dynamics to be studied in a detailed way. Consequently, the calibration and validation of such models requires that many parameters are determined experimentally. For most climate control purposes, the application of such models is rather difficult and time consuming, and success is dependent on the accuracy of some parameters.

Consequently, in recent decades several black and grey box models have been developed [2, 3, 6, 10, 11]. Black box models merely describe statistically what happens with a given input over a limited range. This implies that they are fully based on empiricism. Grey box models have a structure that is based partly on physical, chemical or biological laws (like deterministic models) and partly on empiricism. Grey box models often show relatively accurate high-frequency responses, which make them especially suitable for control purposes. Some of these models, for instance the case with fuzzy models, have the advantage that they do not require explicit evaluation of transfer coefficients or model formulation. Using these techniques, [7, 13] were able to model the greenhouse climate satisfactorily.

Takagi-Sugeno (T-S) models [4, 8, 9] make possible built a nonlinear model interpolating various models that describe local dynamics in a compact set using premise variables. Application of this technique to the greenhouse modelling and controller design show that it may be an efficient tool to handle the high nonlinearities inherent to greenhouses [14].

Environmental control of greenhouses includes control and modification of day and night temperatures, relative humidity, and carbon dioxide levels for optimum plant growth. However, most of the published greenhouse climatic models are adequate for heated and ventilated greenhouses [10, 7, 11, 14]. Indeed, the purpose of the heating system is to replace energy lost from the greenhouse when outside temperatures are lower than desired in the greenhouse growing area, whereas, the greenhouse ventilation plays an important role to control high temperatures during days caused by the influx of solar radiation, to main-

tain relative humidity at acceptable levels during nights, to provide uniform air flow throughout the entire greenhouse, and to maintain acceptable levels of gas concentration in the greenhouse [10, 7]. The objective of this paper is to extend the work of [14] to a multi-variable greenhouse model which includes both control signals: heating and ventilation. The proposed model considers the natural ventilation effect, which involves high nonlinearities caused by the coupling between the greenhouse inside air, the vents opening and the wind speed. The obtained T-S model is given in the state space representation which is more systematic for control design problems.

This paper is organized as follows: Section 2 presents preliminary results for T-S Fuzzy modeling of a nonlinear dynamic system. In section 3 a T-S fuzzy model of greenhouse is derived from its original dynamic model. In sections 4 application of the technique to a specific greenhouse is proposed. Finally, section 5 presents some preliminary controller design based on the derived model.

2 Preliminaries

2.1 Takagi-Sugeno Models

The models suggested by Takagi and Sugeno are described by rules representing the local relations of linear input-outputs in various operation points of a system [4]. These local representations, called "sub-models", make possible to express the dynamics of a system around particular state space operation points. Thus, the fuzzy formalism intervenes in the determination of the contribution of each one of these sub-models to the representation of the total system.

Typically, the T-S fuzzy models in the state space are described by a set of N rules using membership functions F_l^i and fuzzy variables $z_l(k)$ as follows:

$$\begin{array}{l}\text{Rule i: IF } z_1(t) \text{ is } F_1^i \text{ and } \ldots \text{ and } z_l(t) \text{ is } F_l^i \\ \text{THEN } \left\{ \begin{array}{l} \dot{x}(t) = A_i x(t) + B_i u(t) + D_i v(t), \\ y(t) = C_i x(t), \end{array} \right. \end{array} \quad (1)$$

where A_i, B_i, D_i and C_i are constant matrices of appropriate size, $x(t) \in \Re^n$ is the state, $u(t) \in \Re^m$ is the control, $v(t) \in \Re^s$ is the perturbation and $y(t) \in \Re^p$ is the output. It has been shown in [4] that the overall global model can be structured as follows:

$$\left\{ \begin{array}{l} \dot{x}(t) = \sum_{i=1}^{N} h_i(t)(A_i x(t) + B_i u(t) + D_i v(t)), \\ y(t) = \sum_{i=1}^{N} h_i(t) C_i x(t), \end{array} \right. \quad (2)$$

where $h_i(k)$'s are the so-called *normalized activation function* in relation with sub-model i^{th}, such that:

$$h_i(t) = \frac{\prod_{j=1}^{l} F_j^i(z_j(t))}{\sum_{i=1}^{N} \prod_{j=1}^{l} F_j^i(z_j(t))}, \; h_i(k) \geq 0.$$

In order to design a controller for (2), the concept of Parallel Distributed Compensation (PDC) [8], is used: a linear controller is designed for each subsystem in each rule. Hence the nonlinear system controller shares the same fuzzy sets with the fuzzy system (2). The overall controller is then given by a fuzzy interpolation of local controllers, as follows:

$$u(t) = -\sum_{i=1}^{N} h_i(t) K_i x(t), \quad (3)$$

where $u(t)$ is the nonlinear state-feedback control law.

Combining (9) and (3) the following augmented closed-loop system is obtained

$$\dot{x}(t) = \sum_{i,j=1}^{N} h_i(t) h_j(t)((A_i - B_i K_j)x(t) + D_i v(t)), \quad (4)$$

3 Greenhouse model

3.1 Physical model

In a greenhouse, the state climate can be represented by two variables, namely, *inside air temperature* $T_a[K]$, and *absolute humidity* $w_a[kg\,H2O/kg\,air]$. Based on the analysis of the physical processes of greenhouse climate, the model equations can be written by deriving the appropriate energy and mass balances [11]:

$$\begin{array}{c} C_{cap,t} \frac{dT_a}{dt} = Q_{rad} + Q_{cvn,ss-a} + Q_{cal} - Q_{cvn-cnd,a-e} - \\ Q_{ven,a-e} - Q_{losses} - Q_{tra,crop}, \end{array} \quad (5)$$

$$C_{cap,w} \frac{dw_a}{dt} = -M_{ven,a-e} - M_{c,cnd} + M_{losses} + M_{tra,crop}, \quad (6)$$

where

$$\begin{array}{l} Q_{rad} = c_{rad} S_e, \\ Q_{cvn,ss-a} = h_{s-a}(T_s - T_a), \\ Q_{cvn-cnd,a-e} = h_{a-e}(T_a - T_e), \\ Q_{ven,a-e} + Q_{losses} = \phi_v(T_a - T_e), \\ Q_{tra,crop} = \lambda M_{tra,crop}, \\ Q_{cal} = q_{max} h_{cal} u_{cal}, \\ M_{ven,a-e} + M_{losses} = \rho_{air} \phi_v (w_a - w_e), \\ M_{c,cnd} = c_{c,cnd}(w_c - w_a), \\ C_{cap,t} = \rho_{air} C_p V_g / A_g, \\ C_{cap,w} = \rho V_g / A_g \end{array}$$

where $C_{cap,t}$ $[JK^{-1}m^{-2}]$, $C_{cap,w}$ $[Kg\,m^{-2}]$ are the heat and mass capacities of the greenhouse air, respectively.
The energy balance in the greenhouse air (5) is affected by the *energy contribution from the solar radiation* (Q_{rad}),

the *heat exchanges by convection* between air and soil surface ($Q_{cvn,ss-a}$), the energy supply (Q_{cal}) from *heating system*, the *heat exchanges by convection and condition* between the inside air and the outside air of greenhouse ($Q_{cvn-cnd,a-e}$), the heat lost by ventilation ($Q_{ven,a-e}$) and infiltrations (Q_{losses}) and finally the flux of latent heat by crop transpiration ($Q_{tra,crop}$).
The humidity balance in the greenhouse air (5) is determined by canopy transpiration ($M_{tra,crop}$), condensation of water at roof surface ($M_{c,cnd}$), and ventilation exchange with the outside air ($M_{ven,a-e}$).
To approximate the effect of ventilation (ϕ_v) we will use the following model [6], discussed in detail in [11].

$$\phi_v = \left\{ \frac{L_{vent} c_d T_e}{3 A_g c_g \Delta T} \left[\left(c_g W_{vent} u_{vent} \frac{\Delta T}{T_e} + c_w P_{v,e}^2 \right)^{3/2} - (c_w P_{v,e}^2)^{3/2} \right] \right\}, \quad (7)$$

where $\Delta T = T_a - T_e$, u_{vent} is the ventilation control signal, $P_{vv,e}$ is the wind speed, c_d is the discharge coefficient, W_{vent} and L_{vent} are the width and length of the vent, respectively. $M_{trp,cu}$ is the water vapor flux incomes from the crop [12].

Using (5) and (6), the greenhouse climate model can be written as the following nonlinear state space system:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + D(t)v(t), \quad (8)$$

where $u(t) = \begin{bmatrix} u_{cal} & u_{vent} \end{bmatrix}^T$ is the control vector composed of 2 control signals: heating u_{cal} and ventilation u_{vent}, $x(t) = \begin{bmatrix} T_a & w_a \end{bmatrix}^T$ is the state vector, and $v(k) = \begin{bmatrix} S_e & T_s & T_e & w_e & w_c \end{bmatrix}^T$ is the external disturbance vector that includes: solar radiation (S_e), soil temperature (T_s), outside air temperature (T_e), outside air humidity (w_e) and cover humidity (w_c).

If the effect of transpiration is ignored and a new variable ϕ_v^* is considered such that: $\phi_v^* = \frac{\phi_v}{1+u_{vent}}$, then the system matrices of (8) can be written as follows:

$$A = \begin{bmatrix} \frac{1}{C_{cap}}(-h_{s,a} - ha,e - \phi_v^*) & 0 \\ 0 & -\frac{1}{C_{cap,w}}(\rho_{air}\phi_v^* + c_{c,cnd}) \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{1}{C_{cap}} C_{cal} & -\frac{1}{C_{cap,t}}\phi_v^* \Delta T \\ 0 & -\frac{1}{C_{cap,w}}\rho_{air}\phi_v^* \Delta w \end{bmatrix}, \text{ and}$$

$$D = \begin{bmatrix} \frac{1}{C_{cap,t}} c_{rad} & \frac{1}{C_{cap,t}} h_{s,a} & \frac{1}{C_{cap,t}} h_{a,e} + \frac{1}{C_{cap,t}}\phi_v^* & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{C_{cap,w}}\rho_{air}\phi_v^* & \frac{1}{C_{cap,w}} c_{c,cnd} \end{bmatrix}.$$

3.2 Transformation into a T-S model

In this paper, we are interested in calculating a fuzzy model that approximates the climate inside greenhouses throughout the day. To do this, we select the bounded functions ΔT, $\Delta w = w_a - w_e$ and ϕ_v^* as premise variables for the fuzzy rules (Figure 1). Each of these variables requires two membership functions to represent it exactly in a predetermined domain, knowing their bounds in that domain. Accordingly, there are 6 membership functions required in total for the exact representation of (8) and they are represented as:

$$M_{j1} = \frac{\beta_j + f_j(.)}{\beta_j - \alpha_j}, \quad M_{j2} = 1 - M_{j1} \text{ for } j = 1, \ldots 3,$$
$$\text{where } \beta_j = \max(f_j(.)) \text{ and } \alpha_j = \min(f_j(.)).$$

The obtained T-S greenhouse model is given with 3 premise variables: ϕ_v^*, ΔT and Δw. As a consequence, we need only $2^3 = 8$ rules to exactly represent this nonlinear system. It should be noted that the derived T-S models is not unique and of course, more complicated models could be used. However, this will increase the controller complexity and might make its derivation numerically complex. As a simplified model is used, the resulting controller is simple to implement in temperature control systems for greenhouses. In addition, as there is a reduced number of rules, the number of stability analysis conditions to be satisfied will be also reduced, which decreases conservatism.

4 Application to an experimental greenhouse

The experiments took place in a greenhouse of the "Palmerillas Experimental Station", situated in El Ejido (36°48'N, 2°43'W and a height above sea level of 63 m) and belonging to the Cajamar Fundation. The nocturnal and diurnal regulation of the temperature proposed in this paper is frequent in this area during the winter. A heating system with adequate capacity is needed especially in the winter nights to maintain environmental conditions inside the greenhouse for plant growth and development. Even during the coldest part of the winter, when the heating system is running at full capacity, some ventilation is still required in the greenhouse. Fresh, outside air must be ventilated into the greenhouse to remove the warm, moisture-laden air from within the greenhouse. If moist air within the greenhouse is not removed, high humidities and excessive condensation will occur, which can damage the plants and reduce the production.

4.1 Parametrization

The structure of the greenhouse, whose climatic conditions are to be controlled, is symmetric in area and the roof runs from East to West. The surface area is 877m2 with a variable height of between 2.8m and 4.4m. The available control systems are as follows: natural ventilation (in roof and sides), shades, and heating system. The heating system consists of a hot air heater fuelled by diesel oil, with a capacity of supplying 95 Mcal, a fuel tank and a ventilator that creates a forced movement of air inside the greenhouse, thus making the temperature more uniform. The heating system was set up so that the control signal could be managed directly from a PC, to simplify controller testing. The ventilation control actuators installed in the greenhouse are automated flap roof and rolling lateral continuous

ventilators, the former with a maximum aperture angle of 45^o and the latter with 37 m length and 1.2 m aperture. Both have insect screens installed in the ventilation openings to prevent insect infestations and so reducing insecticide use.

During the experiments, several interior measurements were also taken to have a record of the state of the greenhouse, such as relative humidity, radiation, wind speed, soil temperatures at 40cm and 1cm, and temperatures of the netting that covers the soil, of the interior part of the roofing, of the leaves of the plants. In addition, during the tests, the following exterior measurements were taken: air temperature, relative humidity, global radiation, PAR radiation, rainfall and the direction and velocity of the wind. The data acquisition system used was made up of an Ethernet network with Compact Field Point modules from National Instruments. Data was recorded every minute for all variables.

First of all, the parameters of the nonlinear model in (5) and (6) were estimated based on measured data [11], in the greenhouse in which the tests were carried out, obtaining the values shown in Table 1, that also include the physical parameters of the greenhouse.

Now the exact fuzzy model of the system can be expressed as follows:

$$\dot{x}(t) = \sum_{i=1}^{8} h_i(t)(A_i x(t) + B_i u(t) + D_i v(t)), \qquad (9)$$

where A_i, B_i and D_i are the system matrices of each local model. The model (9) was compared with the measured data, collected in the experimental station: Figure 2 shows the evolution of the internal temperature and humidity using (9). It can be seen that the obtained model achieves the real temperature of the greenhouse. since the model does not take into account the culture transpiration, the model has a significant moisture error with measurable data, which were collected in the presence of tomato in the greenhouse.

Table 1: Parameters of the greenhouse climate model.

V_g	Greenhouse volume	3448.58	m^3
A_g	Greenhouse area	877	m^2
ρ	Air density	1.197	Kgm^{-3}
C_p	Specific heat capacity of air	1005	$JKg^{-1}K^{-1}$
q_{max}	Nominal energy of heating system	95000	W
c_{rad}	Heat load coefficient due to solar radiation	2.8	
h_{a-e}	Convection/conduction coefficient inside/outside	200	$Wm^{-2}K^{-1}$
h_{s-a}	Convection coefficient greenhouse air/soil	90.5	$Wm^{-2}K^{-1}$
h_{cal}	Heating coefficient (heating system)	0.287	
$c_{c,cnd}$	Condensation coefficient of greenhouse roof surface	0.1	$Kgm^{-2}s^{-1}$

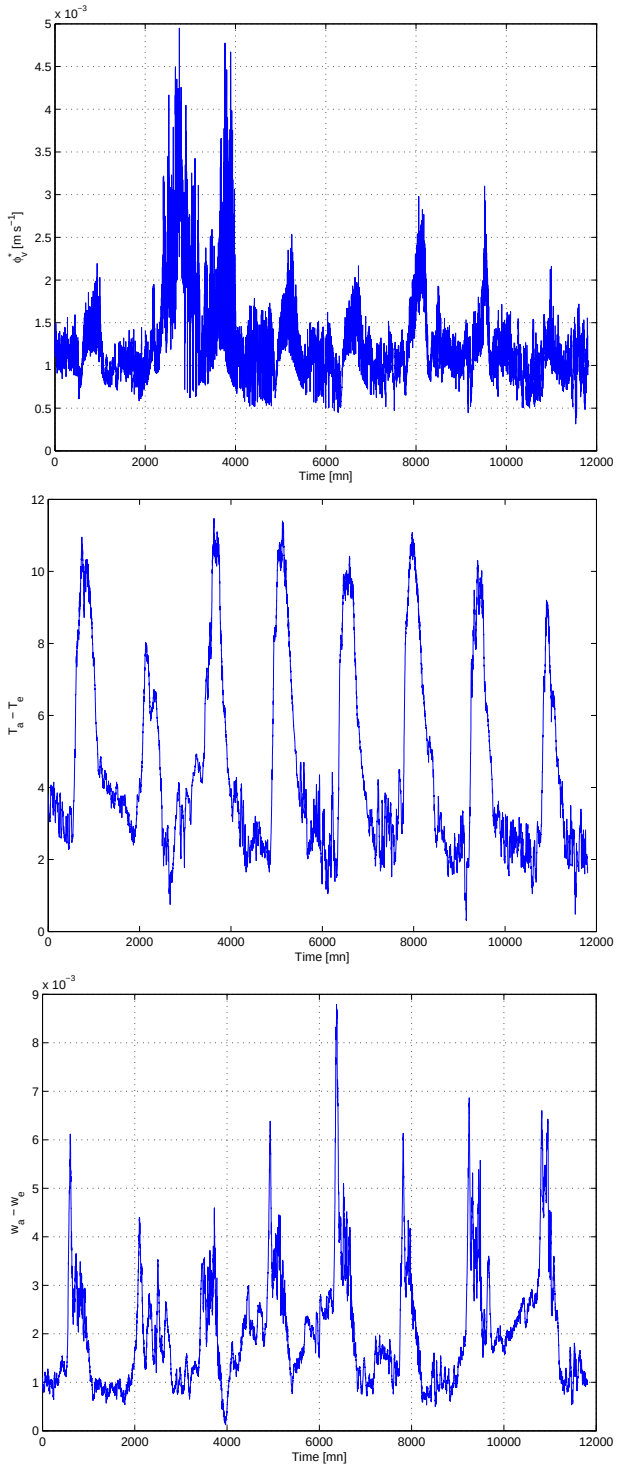


Figure 1: Measured premise variables, using data of 8 days, in experimental station of El Ejido.

4.2 Controller design

To show the applicability of the suggested model, the PDC control approach (3) for state regulation of the nominal system of (9) was applied to the proposed model deriving

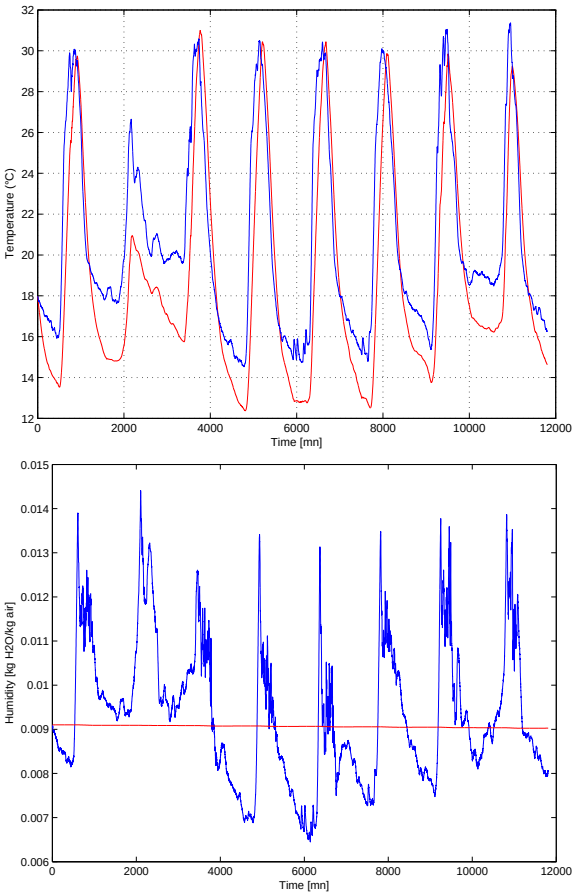


Figure 2: Evolution of internal temperature and humidity respectively using the proposed model (red), in comparison with the measured data (blue)

the controller parameters using the technique proposed in [8, 14].

The stabilizing controller gain is selected such as the closed-loop eigenvalues are $[-0.8064 - 0.6436]$. Figure 3 depicts some preliminary control results for the greenhouse climate: The inside air temperature and humidity achieve the equilibrium within 7 minutes, which show the applicability of the proposed fuzzy model.

5 Conclusion

In this paper, we proposed the using of T-S fuzzy approach to modeling and regulate the greenhouse climate using heating and ventilation systems. Further results will show the applicability of the suggested model to control a climate variables inside the greenhouses with less computing effort.

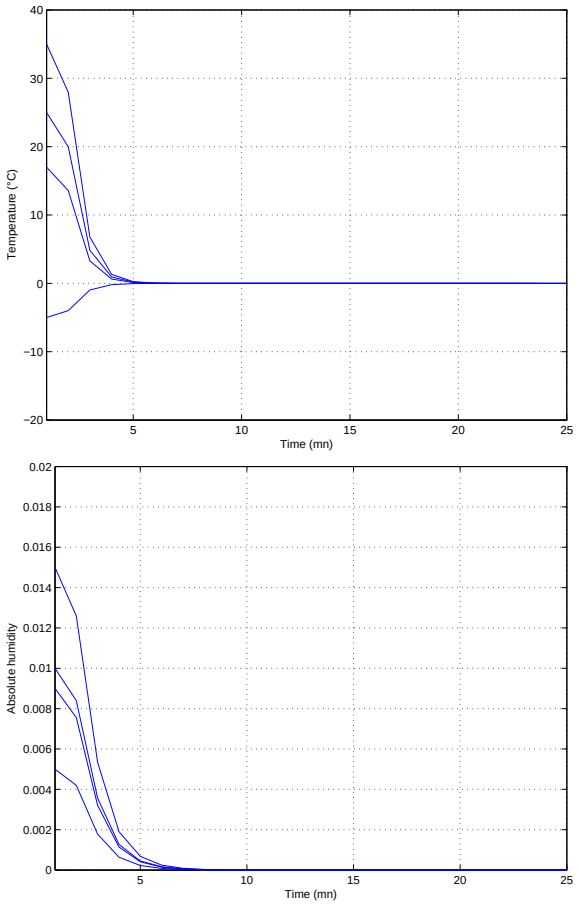


Figure 3: Control results: evolution of tracking errors

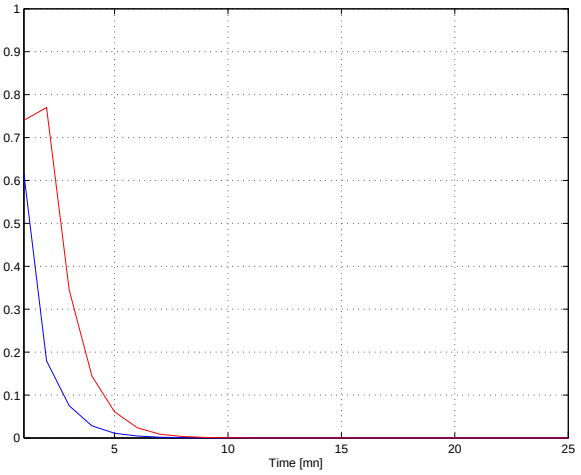


Figure 4: Evolution of the control signals: u_{cal} (red) and u_{vent} (blue) for an initial state $[10\ 0.015]$

References

[1] T. Takakura, K. A. Jordan and L. L. Boyd, Dynamic simulation of plant growth and environment in the greenhouse. *Transactions of the ASAE*, 14, pp. 964-971, 1971.

[2] H. Challa, Some remarks concerning the use of models in greenhouse climate research. *Acta Horticulturae*, 107, pp. 117-120, 1981.

[3] H-J. Tantau, Analysis and synthesis of climate control algorithms. *Acta Horticulturae*, 174, pp. 375-380, 1985.

[4] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. on Systems Man and Cybernetics*, vol. 15, 1985, pp. 116-132.

[5] A. Chotai, P. C. Young, P. David and Z. S. Chalabi, "True digital control of greenhouse systems", *Proceedings of the* 1^{st} *IFAC/ISHS Workshop on Mathematical and Control Applications in Agriculture and Horculturae*, 1991, pp. 41-45.

[6] Boulard, T. and A. Baille, Modelling the air exchange rate equipped with continuous roofs vents. *Journal of Agricultural Engineering Research*. 61, pp.37-48, 1995.

[7] R. Linker, I. Seginer, and P. O. Gutman, Optimal Control of CO_2 in a greenhouse modelled with neural networks, *Computer and Electronics in Agriculture*, 1998, pp. 289-310.

[8] K. Tanaka and H.O. Wang, *Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach*, J. Wiley and Sons Inc., 2001.

[9] M. Nachidi, A. Benzaouia, F. Tadeo and M. Ait Rami, LMI-based approach for output-feedback stabilization for discrete time Takagi-Sugeno systems, *IEEE Transactions on Fuzzy Systems*, 16, 2008, pp. 1188-1196.

[10] F. Lafont and J.-F. Balmat, Optimized fuzzy control of a greenhouse, *Fuzzy Sets and Systems*, 128, 2002, pp. 47-59.

[11] F. Rodriguez, Modelling and hierarchical control of greenhouse crop production. PhD Thesis. University of Almeria, Spain (in Spanish), 2002.

[12] E. Medrano, M.C. Sanchez-Guerrero, P. Lorenzo, Influence of vapour pressure deficit on tomato crop transpiration, *VI International Symposium on Protected Cultivation in Mild Winter Climate: Product and Process Innovation; ISHS Acta Horticulturae*, 614, pp: 613-618, Ragusa-Sicily, Italy; 2003.

[13] F. Lafont and J.-F. Balmat, Fuzzy logic to the identification and the commande of the multidimensional systems, *International Journal of Computational Cognition (ISSN 1542-5908)*, 2, 2004, pp. 21-47.

[14] M. Nachidi, F. Rodriguez, F. Tadeo and J.L. Guzman, Takagi-Sugeno control of nocturnal temperature in greenhouses using air heating, *ISA Transactions*, 50, 2011 pp 315-320.

CONTROL BORROSO PARA EL DRENAJE DE AGUAS PLUVIALES

Mercedes Ramírez Mendoza [1], Manuel González Valdez [2]

[1] *Universidad de Oriente, Cuba,* mramirez@fie.uo.edu.cu
[2] *IUT "Dr. Federico Rivero Palacios", Venezuela,* raduga535@yahoo.es

Resumen

En este artículo se presenta el diseño de un sistema de control de nivel para los tanques de drenaje de aguas pluviales en el metro de Caracas, empleando técnicas de control borroso. El sistema incluye un accionamiento constituido por un motor de inducción alimentado con un inversor (PWM). Como preámbulo se presentan las generalidades relacionadas con el proceso, luego se brinda una breve explicación de los conceptos generales del control borroso que se aplican en el diseño del sistema de control propuesto, el cual fue evaluado con las herramientas del software MATLAB©. Los resultados obtenidos permiten concluir que el control borroso garantiza el cumplimiento de los requerimientos del proceso eficazmente, siendo una alternativa importante a tener en cuenta.

Palabras Clave: Lógica borrosa**,** control inteligente, motor de inducción, control de nivel.

1 INTRODUCCIÓN

En las instalaciones del Metro de Caracas se encuentran sistemas de drenaje de aguas pluviales y de aguas negras. Estos sistemas están formados por un conjunto de bombas, tanques de almacenamiento y tableros de control emplazados entre las diferentes estaciones de pasajeros, y su principal objetivo es el desalojo permanente del agua acumulada en el subsuelo hacia el sistema público de aguas servidas, con la finalidad de evitar que se produzca la inundación de la vía férrea y que los lugares donde están situados no sufran afectaciones por acumulación de agua, [3].

En la actualidad se utiliza en dichas instalaciones un sistema de control de dos posiciones (on-off), que emplea cuatro sensores de tipo flotante y relés para ejercer el control del nivel del pozo de drenaje, mediante el encendido y apagado de dos bombas eléctricas sumergidas.

En el Metro de Caracas, una propuesta para automatizar el control del nivel en los pozos de drenaje debe satisfacer ciertos requerimientos de seguridad, ya que es necesario evitar que los tanques rebasen el nivel máximo, pues en caso de desbordamiento, las vías férreas se inundarían, lo que provocaría la interrupción del servicio. En el caso contrario, o sea, que el nivel esté por debajo del mínimo requerido (0.5 m), las bombas sufrirían daños que podrían ser irreversibles.

El sistema de control que actualmente está en uso ofrece buenos resultados siempre y cuando el proceso no se someta a fuertes lluvias o cambios bruscos en el flujo de entrada al tanque,[3], cuando se producen estas perturbaciones el sistema no garantiza la evacuación necesaria del agua lo que ocasiona las afectaciones antes referidas.

Por ello el objetivo del presente trabajo es diseñar y poner a punto, mediante simulación, un sistema de control del nivel para un tanque de drenaje, empleando técnicas de control borroso, que garantice el mantenimiento de un nivel adecuado frente a diferentes cambios en las perturbaciones, lo que permitiría evaluar la propuesta para una posible implementación en el metro.

Desde hace algún tiempo la lógica borrosa se ha estado utilizando con éxito para el control de los accionamientos de frecuencia variable, es por esto que se ha tomado esta vía para intentar el control del proceso objeto de estudio, cuyas características se advienen al uso de esta técnica por la variabilidad del flujo de entrada y la dificultad para medirlo con exactitud.

Por otra parte el control de los motores eléctricos mediante conjuntos de conmutación "Todo o Nada" aunque es una solución que se utiliza para el accionamiento de una amplia gama de máquinas conlleva limitaciones que pueden ser eliminadas con el empleo de los variadores de velocidad electrónicos los cuales garantizan la aceleración y deceleración progresivas y permiten adaptar la veloci-

dad a las condiciones de explotación de forma muy precisa. En este diseño será empleado un convertidor de frecuencia.

El trabajo está estructurado de la siguiente forma, inicialmente se describe el sistema de drenaje, a continuación se plantean los modelos que representan a los diferentes componentes del sistema de control propuesto, luego se habla sobre el diseño del sistema de control borroso, se describen algunos experimentos de simulación realizados y finalmente se ofrecen las conclusiones del trabajo.

2 METODOLOGÍA

2.1. SISTEMA DE DRENAJE DEL METRO DE CARACAS.

Anteriormente se mencionó que el sistema de drenaje del Metro de Caracas está compuesto por un conjunto de tanques de almacenamiento, bombas y tableros de control, que se encuentran emplazados entre las distintas estaciones de pasajeros. Las aguas negras y las aguas pluviales, corren hacia los tanques, evitando la inundación de la vía férrea y con ello los correspondientes fallos de diversos tipos que pudieran ocasionarse. Cada pozo luego debe drenar esta agua hacia el sistema público de aguas servidas, [3].

Cada uno de los pozos del sistema de drenaje referido posee, además de una ubicación diferente, características particulares de largo, ancho, altura, capacidad, nivel de ceba o nivel mínimo para el funcionamiento de las bombas, entre otras. Para este trabajo se eligió uno de los tanques del sistema cuyas características más significativas se muestran en la Tabla 1.

Otro dato importante es el *Tiempo de reboso* que según los especialistas del metro es de 45 minutos, de manera que el flujo de entrada estimado es *15.55 l/s* o *0.015* m^3/s.

Las dos bombas electro-sumergibles a emplear son iguales y el flujo máximo de salida de cada una es: $Q_s = 64$ *l/s* o *0.064* m^3/s.

Tabla 1: Características del tanque.

Largo	9.50 m
Ancho	2.60 m
Alto	2.15 m
Capacidad	53 110 l
Cap. Real reboso	41 990 l
Nivel de ceba	0.50 m

Se puede constatar entonces que entre ambas bombas se logra extraer un total de: $Q_{sTotal}=0.128\ m^3/s$ que es muy superior al flujo de entrada.

En la Figura 1 se muestra el esquema del pozo:

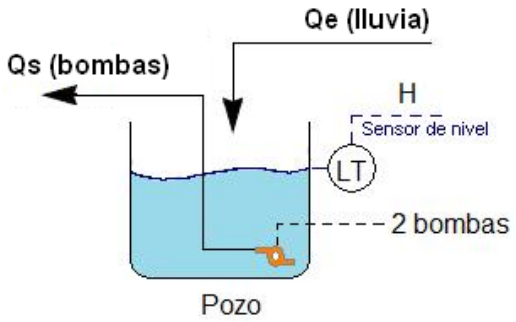


Figura 1: Esquema del proceso.

2.2. MODELO DE LOS COMPONENTES DEL SISTEMA DE CONTROL.

A continuación se muestran los modelos matemáticos de los componentes del sistema de control propuesto para el tanque de drenaje.

2.2.1. Modelo del tanque

Para modelar el tanque o pozo de drenaje se realiza un balance de masa que ofrece como resultado la siguiente expresión:

$$H(s) = \frac{Q_e(s) - Q_s(s)}{As} \tag{1}$$

Donde:

H: Nivel en el tanque.

Q_e: Flujo de entrada al tanque.

Q_s: Flujo de salida.

A: Área de la sección transversal del tanque.

2.2.2. Modelo de la bomba

Como se trata de una bomba centrífuga la ecuación que describe su comportamiento dinámico es:

$$\frac{dQ_s}{dt} = \frac{(k_p * m(t)) - Q_s}{\tau_p} \tag{2}$$

Donde:

K_p: Coeficiente de la bomba.

$m(t)$: Parámetro relacionado con la velocidad del motor de la bomba.

τ_p: Constante de tiempo de la bomba.

Q_s: Flujo de salida de la bomba.

Variando el parámetro *m(t)* es posible variar el flujo de salida mientras que la ganancia de la bomba y su constante de tiempo se mantienen constantes, [5].

2.2.3. Modelo del sensor

Para medir el nivel se propone un transmisor de nivel ultrasónico KOBOLD EchoKing NEO-5003, el cual no requiere contacto con el fluido, por lo que no requiere un mantenimiento frecuente. El mismo posee un rango de entrada de *0* a *8 m* y un rango de salida de *4 a 20 mA.* Dadas las características ofrecidas en el manual del fabricante [4] se trata de un dispositivo de comportamiento lineal con una dinámica despreciable frente a la del proceso, lo que permitió considerar su modelo como una ganancia de *2 mA/m*.

2.2.4. Modelo del accionamiento eléctrico de las bombas

Los variadores de velocidad electrónicos constan de dos módulos, normalmente integrados en una misma envolvente: un módulo de control, que gestiona el funcionamiento del aparato y un módulo de potencia, que suministra energía eléctrica al motor.

En este caso se pretende suministrar los ajustes a través de un PLC. Además se trata de un sistema dotado con amplificación de potencia que funciona a lazo abierto. La velocidad del motor queda determinada por una tensión de referencia. Para un valor dado de la consigna, la velocidad puede variar en función de las perturbaciones (variaciones de la carga). El rango de velocidad se expresa en función de la velocidad nominal.

El convertidor de frecuencia, que se alimenta de la red a tensión y frecuencia fijas, garantiza la alimentación del motor de corriente alterna con tensión y frecuencia variables, en base a las exigencias de velocidad. El circuito de potencia consta de un rectificador y de un variador que, partiendo de la tensión rectificada, produce una tensión de amplitud y frecuencia variables.

Para el accionamiento de las bombas centrífugas se propone utilizar un motor de inducción industrial de *20 HP, 460 V, 1760 rpm*, alimentado a través de una fuente inversora de tensión sinusoidal(PWM), la cual opera bajo la Teoría del Campo Orientado, [5].

En el modelo el inversor es construido completamente con bloques Simulink® Standard. La frecuencia base de la onda sinusoidal de referencia es 60 Hz, mientras que la frecuencia de la onda triangular portadora se establece en 1980 Hz. Su salida pasa a través de bloques de fuentes controladas por voltaje antes de ser aplicada a los devanados del estator de la máquina asincrónica.

El par de carga aplicado al eje de la máquina es constante y se establece en su valor nominal de *11,9 N-.m*. El motor se arranca desde el reposo y la referencia de velocidad se establece en *1760 rpm*; alcanzándose esta velocidad al cabo de *0,1 s*.

2.3. SISTEMA BORROSO

2.3.1. Generalidades sobre los sistemas borrosos

En el ámbito del control de procesos una de las alternativas más destacadas y que presenta mayor número de contribuciones en los últimos años, es el empleo de estructuras borrosas. Ésta se basa en el uso de una descripción lingüística del proceso, en vez de precisar las relaciones matemáticas entre las variables. Los modelos borrosos o Fuzzy son útiles para la descripción de los procesos complejos en los que se no se conocen todas las dinámicas involucradas, [4].

Entre las metodologías desarrolladas en el ámbito del modelado borroso se encuentran los sistemas basados en reglas, que son los más empleados debido a su fácil interpretabilidad, se basan en el mapeo del conjunto de variables de entrada sobre el conjunto de variables de salida a través de una colección de reglas Si-Entonces (If Then). Dependiendo de la estructura de las reglas, existen distintos tipos de modelos borrosos basados en reglas como pueden ser los modelos lingüísticos, los relacionales y los Takagi-Sugeno(TS), [2].

La estructura de las reglas de los modelos lingüísticos presenta la forma descrita en la ecuación (3)

$$\textbf{\textit{Si}}\ x\ \textbf{\textit{es}}\ A\ \textbf{\textit{entonces}}\ z\ \textbf{\textit{es}}\ B \qquad (3)$$

Donde la proposición *'x es A'* es el antecedente de la regla, y la proposición *'z es B'* es el consecuente. Por tanto, *x* y *z* son variables lingüísticas definidas como conjuntos borrosos con dominio $X \in R^n$ y $Z \in R^n$ respectivamente. Los términos *A* y *B* son conceptos lingüísticos constantes que definen un valor cualitativo del sistema en un punto concreto del espacio de variables, por ejemplo temperatura alta, nivel bajo, presión alta, etc., [2].

Las funciones de pertenencia para describir los conjuntos borrosos se definen como funciones matemáticas paramétricas o como tablas de valores, que se almacenan en una base de datos como parte del sistema borroso. En general, se suelen definir múltiples términos lingüísticos A_i en el dominio de una única variable, denominándose partición borrosa de una variables a la colección de los conjuntos borrosos $[A_1, A_2, , A_M]$.

En el caso de los modelos TS la estructura de la base de reglas es como se muestra en la ecuación (4).

$$\textbf{Si}\ x\ \textbf{es}\ A_i\ \textbf{entonces}\ y_i = f_i(x) \qquad (4)$$

Donde $$y_i = f_i(x) = a_i^T x + b_i \qquad (5)$$

En general, las funciones $f_i(x)$ presentan estructuras afines, tal y como muestra la ecuación (5), donde a_i es el

vector de parámetros y b_i es un escalar. Por tanto, cuando $a_i = 0$ los modelos TS son equivalentes a los modelos lingüísticos donde el consecuente es un "singleton", que es un tipo especial de conjunto difuso. En el presente trabajo se utilizará un modelo TS, [9, 10].

Los modelos TS combinan una descripción global del sistema mediante la base de reglas, y una aproximación local mediante las funciones empleadas en los consecuentes, [2].

La salida del modelo, o lo que es lo mismo, la inferencia y el desemborronado del sistema se obtiene mediante el cálculo de la media ponderada de cada una de las reglas, tal y como se indica en la ecuación (6).

$$y = \frac{\sum_{i=1}^{k} \mu_{A_i}(x) y_i}{\sum_{i=1}^{k} \mu_{A_i}(x)} \qquad (6)$$

Donde *k* es el número de reglas y μ_{A_i} es el grado de pertenencia del antecedente de cada una de las reglas.

2.3.2. Diseño del controlador borroso

El sistema de control estará formado por dos bloques borrosos en cascada con el propósito de reducir la cantidad de reglas.

El primer bloque(CLB 1) se encargará de decidir si la velocidad de los motores debe ser aumentada, mantenida o disminuida en función de los valores del nivel y del cambio en el mismo, mientras que el segundo bloque(CLB 2) a partir de la decisión tomada por el CLB 1 y de los valores de velocidad anteriores dará a la salida las velocidades de referencia Nref1 y Nref2, que serán las consignas que entren a los respectivos variadores de velocidad, los que a su vez ofrecerán a la salida valores de tensión alternos (Vabc1 y Vabc2) correspondientes a la alimentación de cada motor. La velocidad que desarrollen los motores estará en correspondencia directa con los valores de $m_1(t)$ y $m_2(t)$, (ver ec. 2) los que cambiarán el flujo de salida de las bombas, con el fin de garantizar un nivel adecuado en el pozo.

En la Figura 3 se muestra la configuración básica del sistema propuesto.

Como es usual para llevar a cabo el diseño se seleccionó además de las variables de entrada y de salida de cada uno de los bloques, los universos de discurso para cada una, la cantidad de conjuntos borrosos, el tipo de función de pertenencia y las reglas lingüísticas, [8].

Se utilizaron funciones de pertenencia triangulares y trapezoidales, tal elección se realizó atendiendo a la simplicidad de este tipo de funciones, lo que facilita la validación y corrección por parte de los expertos y provee una forma cómoda para comunicarse con los usuarios.

En la Figura 2 se muestra como ejemplo la partición realizada a la variable Nivel, en la cual se tiene en cuenta que el valor del nivel llegará al controlador a través de un transmisor de 4-20 mA.

El cambio en el nivel (dH/dt) se tomó como la diferencia entre el nivel actual y el nivel anterior, con un universo de discurso UCN = [-5, 5], y se establecieron 7 conjuntos.

.

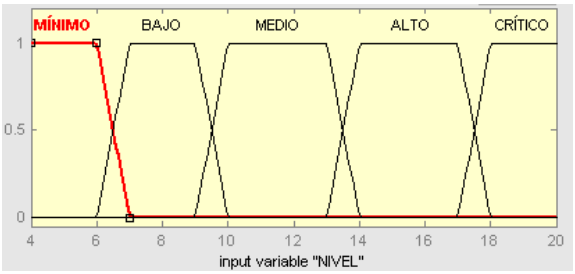


Figura 2: Partición de la variable lingüística *Nivel*.

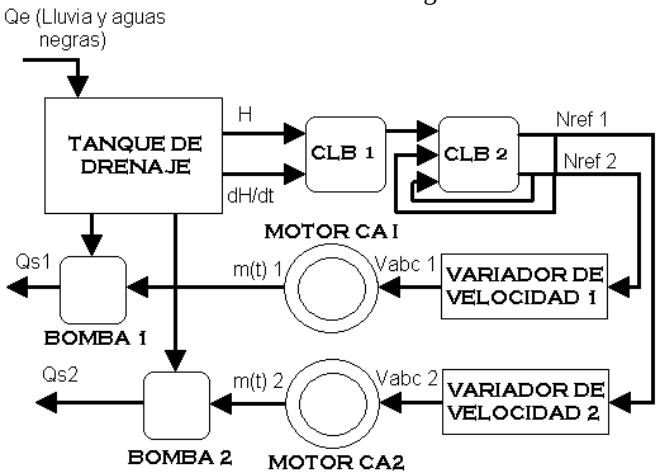


Figura 3: Esquema en bloques del sistema de control de nivel.

La variable de salida del primer controlador es llamada ACCION y toma los siguientes valores: OFF = 0, DISMINUIR = 0.25, MANTENER = 0.5, AUMENTAR = 0.75 y MAXIMO = 1. Esta variable indica al segundo controlador que hacer con las velocidades de referencia.

En el caso del segundo controlador, el mismo consta de 2 variables de salida similares llamadas Nref1 y Nref2 las cuales tomaran los valores: OFF = 0, BAJA = 0.6, ALTA = 0.8 y MAXIMA = 1 en dependencia de que se decida que los motores estén apagados o que la velocidad deseada sea el 60 %, 80 % ó 100 % de la velocidad nominal del motor respectivamente.

Las reglas borrosas utilizadas están basadas en la intuición y se plantearon buscando que el nivel del tanque no alcance el estado crítico y que para flujos de entrada grandes, el nivel se mantenga lo más bajo posible. Es decir, que no solo se actúa de acuerdo a la altura que tiene el líquido en un instante dado, sino que se trata de prever el salto a un estado superior, y que además, tenga siempre al menos el nivel mínimo de 0.5 m.

El primer controlador posee 21 reglas y el segundo 20 reglas. Para el caso del segundo CLB como Nrefant1 y Nrefant2 son iguales, en la Tabla 2 se ejemplifican las reglas que se aplican para una sola de ellas. La agregación se realizó utilizando la T-norma del mínimo.

Tabla 2: Base de reglas del segundo CLB.

REGLAS II		ACCIÓN				
		Off	D	Aum	Man	Max
Nrefant	Off	Off	Off	Baja	Off	Max
	Baja	Off	Off	Alta	Baja	Max
	Alta	Off	Baja	Max	Alta	Max
	Max	Off	Alta	Max	Max	Max

2.4. RESULTADOS DE SIMULACIÓN

A continuación se ofrecen los resultados de las simulaciones realizadas con el sistema de control propuesto. Para las mismas se utilizó el Simulink del MATLAB.

El sistema de control fue puesto a prueba con señales de entrada que simulan flujos de entrada diversos y en todos los casos se utilizó un flujo de entrada tal que $Qe > QsTo$-*tal*, durante un tiempo variable, de tal forma que el sistema es forzado a trabajar en condiciones críticas.

Los resultados obtenidos muestran que a pesar de producirse fuertes variaciones del flujo de entrada, el control borroso responde rápidamente a los cambios y el nivel no rebasa los límites establecidos, es decir, no llega a su estado crítico, lo cual evidencia el buen desempeño del lazo de control propuesto.

En la Figura 4 se observa una gráfica obtenida de una simulación en la que el flujo de entrada fue de 8 m3/min, para ver cómo se comportaba el sistema ante una entrada mayor que el caudal de salida máximo que pueden evacuar las 2 bombas juntas, funcionando a máxima velocidad. Es posible ver que aún durante una hora con una entrada que es superior al flujo de salida máximo, el tanque no se desborda.

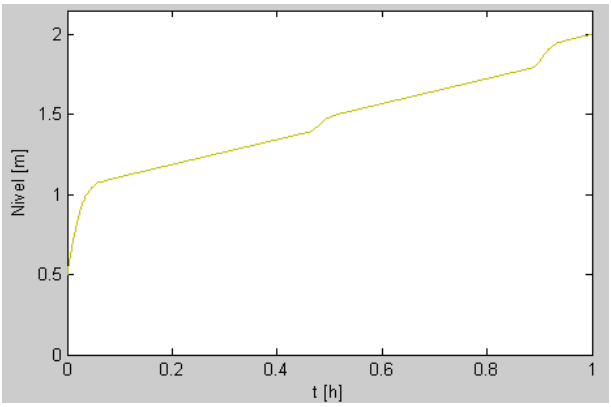


Figura 4: Comportamiento del Nivel vs. t para una entrada de 8 m^3/min.

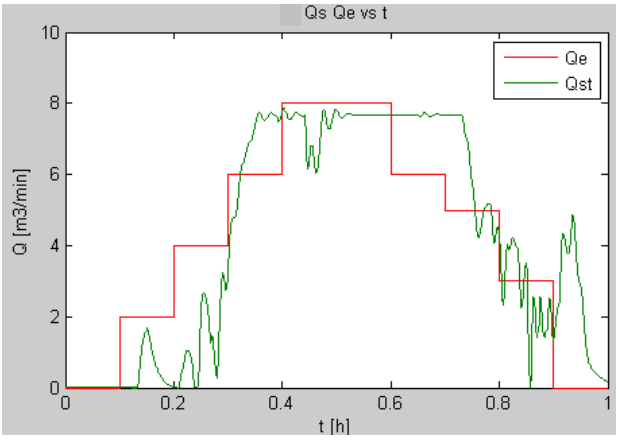


Figura 5: Gráfica de Qe y Qs vs. t para una entrada que aumenta gradualmente hasta los 8 m^3/min y luego disminuye.

En la Figura 5 se refleja la situación en que comienza a llover y el flujo de entrada se va incrementando poco a poco hasta alcanzar un valor superior al flujo máximo de salida, luego a medida que va escampando, la entrada al tanque de drenaje también disminuye gradualmente. Se

aprecia además que la curva del flujo de salida, sigue a la del caudal de entrada, y en la Figura 6 se observa como el nivel no sobrepasa los 1.6 metros a pesar de los cambios que ocurren en la entrada.

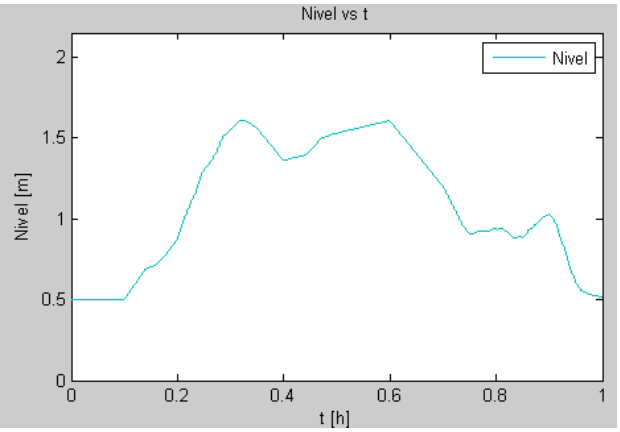


Figura 6: Comportamiento del Nivel vs. t para el flujo de entrada de la Figura 5 .

2.5. CONCLUSIONES

Se realizó el diseño y puesta a punto a través de simulación de un sistema de control de nivel para el proceso de drenaje del Metro de Caracas usando una estrategia de control borroso. Para lograr el objetivo planteado se diseñaron dos controladores borrosos que trabajan en cascada y que a partir de mediciones de nivel y de los cambios en este último, modifican convenientemente el flujo de salida de las bombas instaladas en los tanques de drenaje de aguas pluviales.

El estudio incluyó la utilización de un accionamiento eléctrico para manejar las bombas centrífugas compuesto por un motor de inducción alimentado a través de una fuente inversora de tensión sinusoidal (PWM), la cual opera bajo la Teoría del Campo Orientado.

Se concluye entonces que con la utilización de esta estrategia de control se obtienen resultados favorables, pues el sistema funciona correctamente y se observa la capacidad del mismo para adaptarse a perturbaciones diversas y donde puedan existir variaciones en los parámetros del proceso.

Se demuestra que con la propuesta realizada es posible garantizar el buen funcionamiento del sistema, y que independientemente de cómo varíe el flujo de entrada, siempre y cuando este no exceda durante mucho tiempo del flujo máximo de salida que puede alcanzarse, no habrá desbordamiento y por otra parte el tanque tendrá siempre al menos el nivel mínimo requerido.

Referencias

[1] R. Babuska, & H.B. Verbruggen,,. An overview of fuzzy modeling for control, *Control Engineering Pratice*, 4(11), 1593-1606, 1996.

[2] S. García-Nieto, Identificación y Control Predictivo Fuzzy TS en Espacio de Estados. Una Aproximación al Control No Lineal, *Tesis Doctoral, Departamento de Ingeniería de Sistemas y Automática, Universidad Politécnica de Valencia,* España, 2010.

[3] M. González, D. Guzmán, K. Insers, S. Marrero, A. Columbié: Modelación y Simulación para el Control del Drenaje de los Pozos del Metro de Caracas. *Presentada en: XIV Convención científica de ingeniería,* La Habana, Cuba, 2008.

[4] P. J. King and E. H. Mamdani. The application of fuzzy control systems to industrial processes. *Automatica*, 13(3), pp. 235–242, May 1977.

[5] KOBOLD EchoKing NEO-5003 Series. Ultrasonic Level Transmitter. *User Instructions. KOBOLD Instruments Inc. 1801* Parkway View Drive Pittsburgh PA 15205, 2008.

[6] L.D. Rojas, Accionamiento Eléctrico Asincrónico para Transporte Eficiente de Pulpa Laterítica, *Tesis Doctoral, Facultad de Ingeniería Eléctrica, Instituto Superior Politécnico José Antonio Echeverría*, Cuba, 2006.

[7] F.G. Shinskey, *Process-Control Systems. Application, design and tunning,* 4th ed. McGraw-Hill,New York 1996.

[8] W. Siler and J. J. Buckley, *Fuzzy expert systems and fuzzy reasoning*. John Wiley & Sons, Inc. Hoboken, New Jersey, 2005.

[9] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1), pp.116–132, 1985.

[10] H. Ying. Suficient conditions on uniform approximation of multivariate functions by general Takagi–Sugeno fuzzy systems with linear rule consequent. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 28(4), pp. 515–520, July 1998.

DISEÑANDO DATASETS CON DATOS DE BAJA CALIDAD PARA ALGORITMOS DE APRENDIZAJE COMPUTACIONAL

José M. Cadenas, **M. Carmen Garrido**, **Raquel Martínez**

Depto. Ingeniería de la Información y las Comunicaciones. Universidad de Murcia. Murcia. Spain,
{jcadenas,carmengarrido,raquel.m.e}@um.es

Resumen

Cuando se diseña un algoritmo de aprendizaje computacional capaz de gestionar información de baja calidad, los investigadores tienen que crear sus propios datasets para probar y ajustar su método. La herramienta NiP$_{2.0}$ facilita el proceso de crear datasets con datos de baja calidad de una manera más formal y rigurosa. Así se simplifica la comparación de resultados y se pueden crear nuevos repositorios.

Palabras Clave: Herramienta software, Datos de baja calidad, Soft Computing.

1 INTRODUCCIÓN

La explotación de la información de baja calidad en datasets es un nuevo reto en el campo de investigación de aprendizaje computacional. Sin embargo, en los trabajos en los que ciertos algoritmos de aprendizaje usan datos de baja calidad, los autores generan sus propios datasets y en algunos casos dichos datos están accesibles para su uso en posteriores trabajos en el área. Esto es así debido a la inexistencia de un repositorio de datasets con datos de baja calidad al que poder acceder y que sirva como marco de referencia común a la investigación en este campo. La generación de estos tipos de datasets suele ser una labor tediosa por lo que sería deseable disponer de alguna herramienta que facilite este trabajo.

Debido a esto, surge la idea de la creación de una herramienta que facilite la generación de datasets con datos de baja calidad. La versión estable fue denominada NiP, en su versión 1.5 [4], como abreviatura de "Nothing is Perfect". Sin embargo, desde su creación inicial hemos visto necesaria la inclusión de nuevos elementos que ayudarían a aumentar y mejorar sus prestaciones.

En este trabajo, indicamos qué aspectos van a ser recogidos en la nueva versión 2.0 de la herramienta (NiP$_{2.0}$) que en breve estará disponible en la web [4]. De forma general, los avances se centran 1) en la posibilidad de transformar los valores de baja calidad en un dataset por otros que puedan ser tratados por el algoritmo de aprendizaje computacional con el que queremos trabajar y 2) incorporar nuevos tipos de datos de baja calidad que en la versión previa no se consideraban. Estos dos aspectos principales llevan a realizar modificaciones en los módulos de entrada y salida de la herramienta.

Para ello el trabajo se estructura de la siguiente forma: en la Sección 2 se realiza una introducción a los datos de baja calidad que nos podemos encontrar, en la Sección 3 se presentan algunas herramientas usadas para el tratamiento de datos. Posteriormente, en la Sección 4, se presentan las nuevas funcionalidades de la herramienta NiP$_{2.0}$. Finalmente en la Sección 5 se presentan las conclusiones.

2 DATOS DE BAJA CALIDAD

La información de baja calidad aparece de manera general en dominios y situaciones reales. Los errores instrumentales o la corrupción debida al ruido durante la recogida de datos puede dar lugar a información incompleta para atributos específicos. En otros casos, la extracción de información exacta puede ser excesivamente costosa. En ese caso, puede ser útil aumentar el dataset con información adicional de un experto, que suele trabajar con términos lingüísticos fuzzy como *pequeño, más o menos, cercano a,* etc.

Ante estas situaciones, los datasets a partir de los cuales se debe extraer conocimiento o modelar un sistema van a contener información de baja calidad. Estos datasets deben ser utilizados posteriormente por algoritmos de aprendizaje computacional. En esta sección realizamos un estudio breve de algunos de los tipos de datos de baja calidad y su tratamiento general por los algoritmos.

2.1 VALORES MISSING

Un factor importante a considerar cuando trabajamos con algoritmos de aprendizaje computacional es el tratamiento

que dicho algoritmo realiza sobre valores missing o perdidos. Según se lleve a cabo dicho tratamiento se pueden:

- Dejar dichos valores en el dataset cuando el algoritmo es robusto a la existencia de dichos valores.
- Eliminar dicho atributo del dataset.
- Eliminar dicho ejemplo del dataset.
- Reemplazar el valor manualmente (si no son muchos) o automáticamente por un valor que preserve la media o varianza (globales o por clases/grupos), en el caso de valores numéricos, o por la moda en el caso de atributos nominales. Otra forma de estimar un valor es predecirlo a partir de otros ejemplos (imputación de valores missing), utilizando cualquier técnica predictiva de aprendizaje automático.
- Representarlo como un intervalo que abarca el dominio completo del atributo.

2.2 RUIDO

Ruido es un término muy amplio que ha sido interpretado de formas distintas. Por ejemplo, ha sido definido como un error aleatorio en una variable medida. Otra definición establece que ruido es cualquier propiedad del patrón detectado que no se debe al modelo subyacente real sino a la aleatoriedad en el mundo o los sensores [9].

En cualquiera de estas definiciones el ruido es considerado como un error ocurrido aleatoriamente. El ruido en los datos pueden aparecer por diversos motivos:

- En primer lugar debido a problemas en los instrumentos de medida o equipos.
- En segundo lugar, a que grandes datasets se obtienen mediante métodos automatizados.

Dependiendo del algoritmo de aprendizaje computacional con el que se debe trabajar, se puede:

- Si es posible detectar los valores con ruido, el tratamiento puede ser similar al caso de valores missing.
- En ciertas ocasiones es posible conocer el error cometido por un instrumento de medida (media y desviación del error). Esto permite incorporar esta información al conjunto de valores del atributo transformando cada valor en un intervalo clásico o intervalo fuzzy.

2.3 INCERTIDUMBRE

Dos tipos de información de baja calidad que podemos encontrar en nuestro entorno y en nuestro proceso de pensamiento son: la imprecisión y la incertidumbre. La imprecisión y la incertidumbre pueden ser considerados como dos aspectos complementarios de información de baja calidad.

En los últimos 10 años, se ha marcado la diferencia fundamental entre los conceptos de incertidumbre estocástica e incertidumbre epistemológica [8]. La incertidumbre estocástica está relacionada con la variabilidad aleatoria relacionada con los procesos naturales como la heterogeneidad de la población o fluctuaciones de una cantidad con el tiempo. La incertidumbre epistemológica proviene de la naturaleza imprecisa de la información disponible. Mientras la incertidumbre estocástica es considerada adecuadamente por la Teoría de la Probabilidad, hay distintas teorías que han sido desarrolladas en orden de manejar la información imprecisa como por ejemplo la Teoría de Análisis de Intervalos y la Teoría de la Posibilidad.

Algunas representaciones simples de información incierta basadas en intervalos y su generalización son:

- Mediante un intervalo $[a,b]$, de forma que suponemos que el valor del atributo está dentro de él.
- Mediante un conjunto o intervalo fuzzy que asigna un grado de posibilidad entre 0 y 1 a cada valor de dicho intervalo como posible valor del atributo.

Según se lleve a cabo el tratamiento de la incertidumbre en los distintos algoritmos de aprendizaje computacional, se puede:

- Dejar dichos valores en el dataset cuando el algoritmo es robusto a la existencia de dichos valores.
- Eliminar dicho atributo del dataset.
- Eliminar dicho ejemplo del dataset.
- Sustituir un intervalo por el valor central del mismo o en el caso de un intervalo fuzzy por el valor que se encuentra en el centro de gravedad de dicho intervalo o en el centro del α-corte mayor.
- Imputar su valor, utilizando cualquier técnica predictiva de aprendizaje automático.

3 HERRAMIENTAS SOFTWARE

Para poder comprobar la eficacia de una técnica de aprendizaje se utilizan datasets obtenidos de diferentes entornos reales, algunos de las cuales se pueden obtener de repositorios públicos [1, 2]. Estos datasets no suelen contener información de baja calidad, a excepción de algún valor missing o, algunos de ellos, valores en intervalos. Por otro lado, cuando se diseña una técnica capaz de gestionar información de baja calidad, los autores tienen que crear sus propios datasets para probar y ajustar su método. Con $\mathrm{NiP}_{2.0}$ se facilita dicho proceso de una manera más formal y rigurosa para simplificar asimismo la comparación de resultados con otros autores. A continuación se revisan algunas de las herramientas software de minería de datos y aprendizaje junto con sus principales características desde este punto de vista.

El funcionamiento de *Sodas2* [7] consiste en construir datasets simbólicos que resuman la información del dataset inicial. El uso de datos simbólicos le permite introducir diferentes tipos de datos de baja calidad: atributos multivaluados, intervalos y atributos multivaluados con pesos. Este tipo de atributos permiten representar otros tipos de datos de baja calidad, como datos fuzzy, imprecisos e inciertos. No obstante, *Sodas2* no permite el uso directo de tecnología fuzzy ni la introducción de valores missing, ruido, etc.

Weka [14] proporciona un conjunto de herramientas para el preproceso de datos, clasificación, etc. Para el preproceso ofrece una amplia variedad de técnicas. En el caso de datos de baja calidad, el número de técnicas disminuye, ofreciendo pocas posibilidades. *Weka* gestiona únicamente valores missing, y proporciona herramientas para sustituir dichos valores por la moda o media. Además, permite la adición de ruido en atributos nominales.

Keel [1] contiene una gran colección de algoritmos de extracción de conocimiento, preproceso y algoritmos de aprendizaje, modelos híbridos, etc. *KEEL* permite el uso de diferentes formatos de entrada/salida como CSV, XML o ARFF. Su módulo de *Gestión de Datos* proporciona una variedad de técnicas para selección de instancias, selección de características, discretización, etc. Una vez más, las técnicas proporcionadas para gestión de datos de baja calidad no son muchas, ya que sólo se permite el uso de valores missing mediante métodos de imputación (borrado de instancias, imputación de k-vecinos más cercanos, etc).

Rapid miner [11] (antiguamente *YALE*) ofrece la posibilidad de usar diferentes formatos de entrada/salida. El número de técnicas de preproceso es grande pero, de nuevo, disminuye si nos centramos en técnicas que gestionen datos de baja calidad. *Rapid miner* permite gestionar valores missing usando métodos de imputación: reemplazo con la media, moda, máximo, mínimo o el valor predicho mediante una técnica de minería de datos definida por el usuario. También permite la introducción de ruido en atributos numéricos y/o nominales.

MiningMart [12] ofrece diversas técnicas de agregación, discretización, limpieza de datos, tratamiento de valores nulos y selección de atributos relevantes. El único soporte que ofrece para datos de baja calidad es sobre los valores missing, eliminando las instancias que los contienen o bien usando algún método de imputación.

Existen otras herramientas centradas en el proceso de extracción de conocimiento, como *ADaM* [13], *D2K* [10] or *KNIME* [3], entre otras, pero no prestan mucha atención al tratamiento de datos de baja calidad.

La mayoría de estas herramientas centran su funcionalidad en el tratamiento de los datos, teniendo sólo en cuenta la existencia de datos de baja calidad de valores missing. En cuanto a la introducción de imperfección, son sólo algunas herramientas las que permiten introducir valores missing, menos las que introducen ruido, y no hay ninguna que trabaje con conjuntos fuzzy. NiP$_{2.0}$ es capaz de generar datasets con esos tipos de datos de baja calidad que el resto de herramientas no soporta o soporta parcialmente.

4 NIP 2.0

NiP$_{2.0}$ [4] es una herramienta que permite generar y gestionar datasets con datos de baja calidad, concebida para su uso en entornos de investigación debido a la carencia de un software de características similares que ayude a establecer un marco común de trabajo con estos tipos de datos. En esta sección mostramos las nuevas funcionalidades del sistema, describiendo las diversas opciones. La Figura 1 muestra, de manera muy general, las funcionalidades del software, destacando las que son adaptadas y mejoradas.

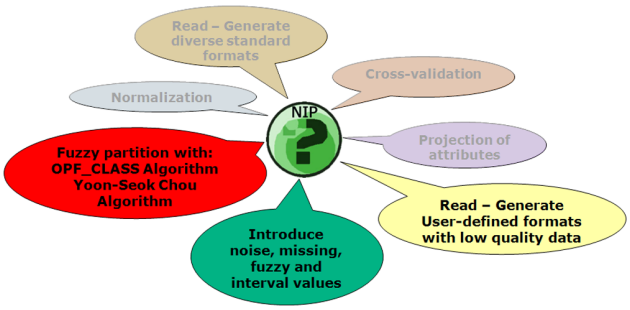


Figura 1: Funcionalidades de NiP$_{2.0}$

4.1 FORMATO DE ENTRADA

NiP$_{2.0}$ ha sido actualizado para permitir que en los formatos de entrada predefinidos por el usuario (Custom Input Format) se pueda incorporar la entrada de datos de baja calidad, en concreto, valores expresados mediante una función de pertenencia fuzzy y valores expresados mediante un intervalos (Figura 2).

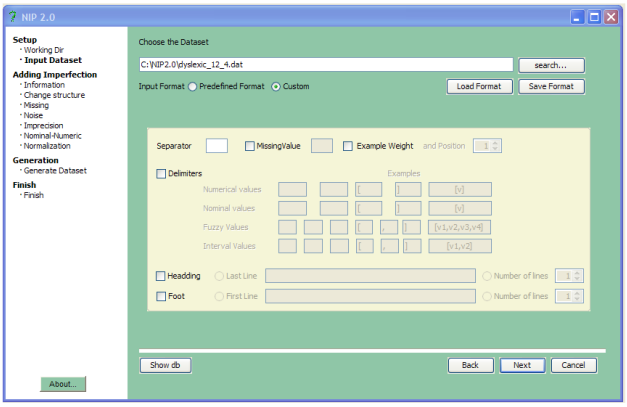


Figura 2: Formato entrada

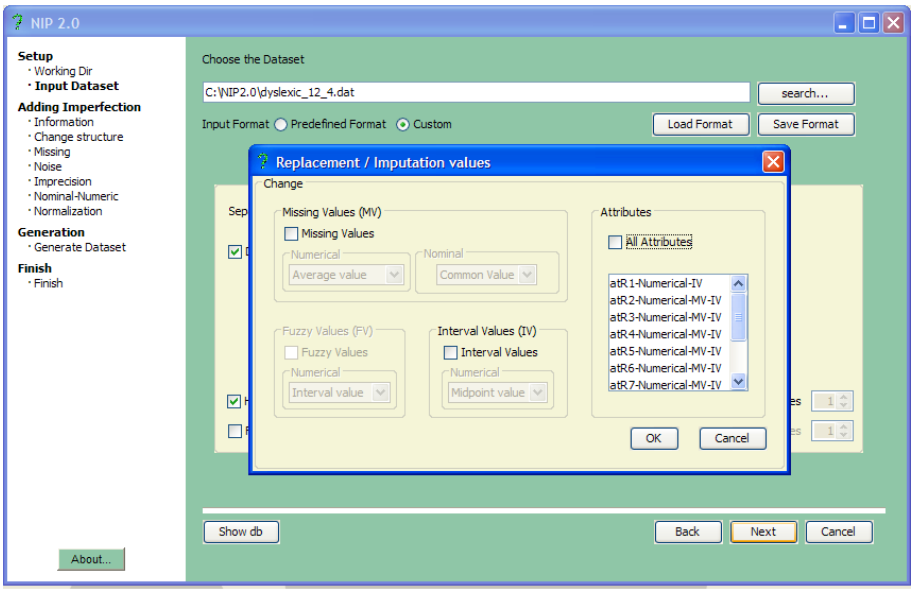


Figura 3: Sustitución de valores

Cuando la herramienta carga un dataset, automáticamente detecta, si los hubiera, los tipos de datos de baja calidad que contiene. En estas ocasiones y si la técnica de aprendizaje computacional con la que se trabaja no permite el tratamiento de estos datos, NiP$_{2.0}$ ha sido ampliado para permitir que el usuario pueda sustituir los valores de baja calidad del dataset por otro que pueda ser manipulado por la técnica en cuestión (replacement and imputation values). De esta forma, aquellos valores de baja calidad que no puedan ser manipulados pueden ser sustituidos por otro tipo de valor mediante las opciones permitidas en la Figura 3.

4.2 INTRODUCCIÓN DE DATOS DE BAJA CALIDAD

Uno de los aspectos más importantes de la herramienta NiP$_{2.0}$, es la introducción de porcentajes de valores de baja calidad en los datasets. Esta funcionalidad es la que permite la generación de experimentos sintéticos que permitan medir la robustez de las distintas técnicas de aprendizaje computacional ante datos de baja calidad. Con ella pretendemos generar repositorios de datasets que ayuden en este campo de investigación.

En esta subsección veremos las nuevas opciones que ofrece NiP$_{2.0}$ para introducir datos de baja calidad en los datasets, centrándonos en la introducción de valores imprecisos que es donde se ha producido una mayor ampliación de su funcionalidad (Figura 5).

4.2.1 Añadiendo valores de baja calidad

NiP$_{2.0}$ permite introducir datos de baja calidad (Figura 4) al dataset en forma de:

- Valores missing.
- Valores con ruido. Permite añadir ruido gaussiano en atributos numéricos y ruido en atributos nominales. En el caso de los atributos numéricos, NiP$_{2.0}$ permite introducir ruido gaussiano indicando la media μ y la desviación típica σ de la distribución para cada atributo. En el caso de atributos nominales, el ruido cambia el valor original del atributo por otro valor que esté en el conjunto de posibles valores de dicho atributo.
- Valores imprecisos que analizaremos más detenidamente en la siguiente sección.

Dichos valores se pueden introducir en todos los atributos o en un subconjunto de ellos. El número de valores de baja calidad a añadir depende de un porcentaje establecido por el usuario, que puede ser distinto para cada atributo.

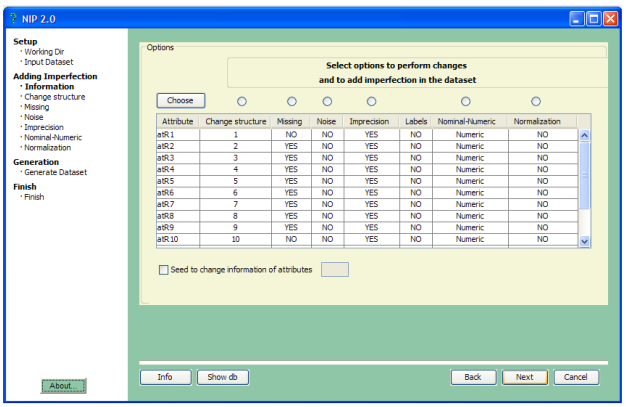


Figura 4: Introducción de valores de valores de baja calidad

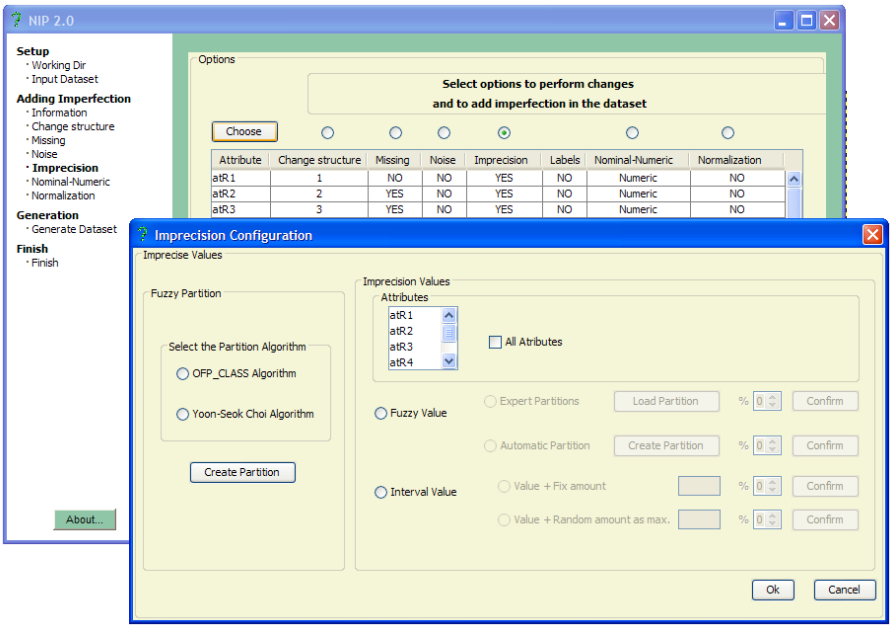


Figura 5: Introducción de valores imprecisos

4.2.2 Imprecisión

Como hemos visto en la sección 2.3, la Teoría de la Posibilidad y la Teoría de Intervalos sirven para representar información imprecisa. NiP$_{2.0}$ permite añadir imprecisión al datasets sustituyendo los valores de un atributo numérico por un valor fuzzy o un intervalo que contenga dicho valor. El usuario decide qué atributos incluyen este tipo de imprecisión y en qué porcentaje (Figura 5).

Cuando introducimos valores intervalares, podemos generar un intervalo añadiendo una cantidad fija al valor que se quiere modificar ([value-fix amount,value+fix amount]) o añadiendo una cantidad aleatoria ([value-random amount,value+random amount]). En ambos casos, los intervalos tienen como cotas las definidos por el usuario limitadas por el dominio de dicho atributo.

Cuando introducimos valores fuzzy, será necesario disponer de una partición fuzzy del atributo, de forma que un determinado valor crisp será sustituido por el valor fuzzy de la partición al que pertenece con mayor grado. Esta partición fuzzy podrá ser proporcionada por un experto en el dominio (expert partition) o ser generada por un algoritmo de particionamiento automático. En este último caso, la herramienta incorpora dos algoritmos:

- El algoritmo OFP_CLASS [5].
- El algoritmo de Yoon-Seok Choi [6].

4.3 FORMATO DE SALIDA

NiP$_{2.0}$ tambien ha ampliado el formato de salida definido por el usuario permitiendo guardar con distintos formatos los valores fuzzy que proceden de un algoritmo de particionamiento automático y aquellos que proceden de una partición proporcionada por un usuario experto (Figura 6). Los valores fuzzy podrán ser representados en el dataset de salida en uno de los siguientes formatos:

- Con el nombre de la etiqueta lingüística (label).
- Con los valores del valor fuzzy con función de pertenencia trapezoidal (V_1, V_2, V_3, V_4).

5 CONCLUSIONES

La falta de datasets con representación de la imperfección inherente a cualquier recogida de datos hace difícil el desarrollo de técnicas de extracción de conocimiento que gestionen datos de baja calidad. En este trabajo hemos descrito una herramienta software Java, llamada NiP$_{2.0}$, diseñada como una herramienta de ayuda a la generación y gestión de datasets que sirve como marco común para la prueba y comparación de técnicas y algoritmos de minería de datos y aprendizaje.

Hemos presentado las principales características de esta herramienta software y hemos destacado que NiP$_{2.0}$ es capaz de incluir en un dataset más tipos de imperfección que las herramientas actuales. Además NiP$_{2.0}$ acepta formatos de entrada/salida predefinidos y personalizados, con lo cual la convierte en un software más flexible.

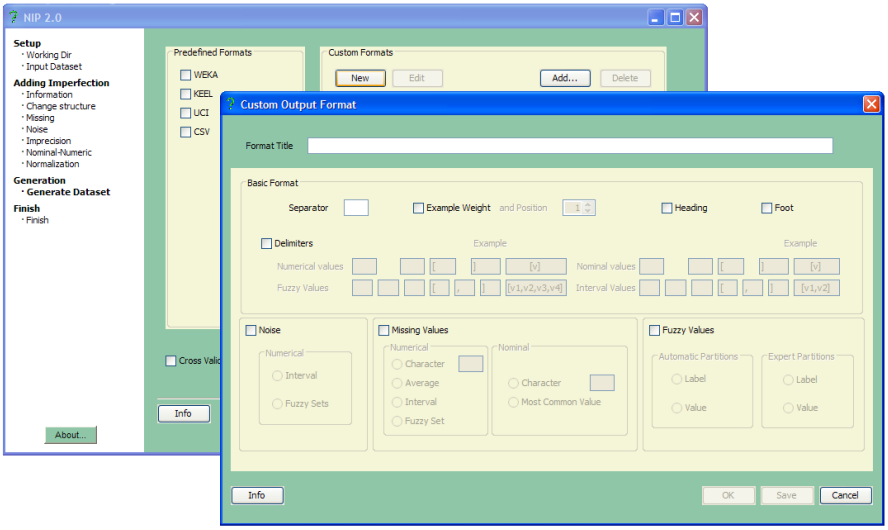


Figura 6: Formato de salida

Agradecimientos

Los autores agradecen la financiación proporcionada por los proyectos TIN2008-06872-C04-03 y TIN2011-27696-C02-02 del MICINN, y por el proyecto 04552/GERM/06 y al programa de FPI de la "Agencia de Ciencia y Tecnología" de la Región de Murcia, de España.

Referencias

[1] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández and F. Herrera, KEEL: A Software Tool to Assess Evolutionary Algorithms to Data Mining Problems. *Soft Computing* **13**(3) (2009) 307–318.

[2] A. Asuncion, and D.J. Newman, UCI Machine Learning Repository, http://www.ics.uci.edu/∼mlearn/MLRepository.html, Irvine, University of California, School of Inf. and Computer Science, 2007.

[3] M.R Berthold, N. Cebron, F. Dill, T.R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel and B. Wiswedel, KNIME: The Konstanz Information Miner, in *Data Analysis, Machine Learning and Applications, eds. C. Preisach, H. Burkhardt, L. Schmidt-Thieme and R. Decker (Springer, Berlin, 2008)*, pp. 319–326.

[4] J.M. Cadenas, J.V. Carrillo, M.C. Garrido, R. Martínez-España, E. Muñoz, NIP 1.5 - A tool to handle imperfect information on datasets. Murcia University, 2008. *http://heurimind* .inf.um.es/heurimind/NIP-Tool.shtm

[5] J.M. Cadenas, M.C. Garrido, R. Martínez-España, P.P. Bonissone, OFP_CLASS: A hybrid method to generate Optimized Fuzzy Partitions for Classification, *Soft Computing* (2011) (in press).

[6] Y-S. Choi, B.R. Moon, Feature Selection in Genetic Fuzzy Discretization for the Pattern Classification Problems, *IEICE Transac.* **90**(7) (2007) 1047–1054.

[7] E. Diday and M. Noirhomme-Fraiture, *Symbolic Data Analysis and the SODAS Software*, (Wiley-Interscience, New York, 2008).

[8] D. Dubois, D. Guyonnet, Risk-informed decision-making in the presence of epistemic uncertainty, *I. J. of General Systems* **40**(2) (2011) 145–167.

[9] R. Hickey, Noise Modeling and Evaluating Learning from Examples, *Artificial I.* **82**(1-2) (1996) 157–179.

[10] X. Llorà, E2K: Evolution to knowledge, *SIGEVOlution* **1**(3) (2006) 10–16.

[11] I. Mierswa, M. Wurst, R. Klinkenberg, M. Scholz and T. Euler, YALE: Rapid Prototyping for Complex Data Mining Tasks, in *Proc. 12th ACM SIGKDD Int. Conference on KDD-06* (2006) 935–940.

[12] K. Morik, M. Scholz, The MiningMart Approach to Knowledge Discovery in Databases, in *Intelligent Tech. for Information Analysis*, eds. N. Zhong, J. Liu (Springer, Berlin, 2004) pp. 47–65.

[13] J. Rushing, R. Ramachandran, U. Nair, S. Graves, R. Welch and H. Lin, ADaM: a data mining toolkit for scientists and engineers, *Computers & geosciences* **31** (5) (2005) 607–618.

[14] I.H. Witten, E. Frank *Data Mining: Practical machine learning tools and techniques, 2nd Edition*, (Morgan Kaufmann, San Francisco, 2005).

UN API PARA MODELADO DIFUSO APROXIMATIVO

Mercedes Valdés-Vela, **Fernando Terroso-Sáénz**, **Antonio Gómez-Skarmeta**

Departamento de Ingeniería de la Información y de las Comunicaciones, Universidad de Murcia,
{mdvaldes,fterroso,skarmeta}@um.es

Resumen

En este trabajo exponemos los principales aspectos de una estructura de clases Java para la programación de experimentos de modelado difuso aproximativo de una manera relativamente sencilla. La herramienta favorece la incorporación de nuevos tipos de modelos difusos y de nuevos algoritmos. Además, proporciona los mecanismos para serializar y hacer persistentes los modelos difusos para su uso como parte de posteriores experimentos de modelado, favoreciendo así la hibridación de técnicas, la cual, es un elemento clave en el área del *Soft Computing*.

Palabras Clave: Modelado difuso aproximativo, modelos *TSK*, java, XML

1 Introducción

El modelado difuso es un enfoque para la construcción de modelos de sistemas utilizando un lenguaje descriptivo basado en la lógica difusa con predicados difusos [6]. En un modelo difuso se expresan relaciones entre variables por medio de reglas difusas $SI - ENTONCES$. Cuando existen datos entrada-salida sobre el sistema se pueden utilizar técnicas de identificación de sistemas para realizar el proceso de modelado.

Sin embargo, según el Principio de Incompatibilidad que Zadeh introdujo en [12], las características de interpretabilidad y precisión deseables para cualquier modelo de un sistema son mutuamente contrarias. Por lo tanto, dependiendo de cuáles sean los fines para los que se realiza la tarea de modelado, se utilizarán técnicas que favorecen una característica frente a la otra.

En el *modelado difuso aproximativo o preciso* (AFM, del inglés *approximative fuzzy modeling*), el objetivo principal es la obtención de modelos que aproximen lo mejor posible al sistema objetivo, sin intención de que los conjuntos difusos que lo forman sean interpretables, en contraposición con el llamado *modelado difuso descriptivo*.

El proceso de AFM está compuesto por varias fases. Así, deben ser resueltas las tareas de identificación de la estructura e identificación de los parámetros. La identificación de la estructura consiste en la identificación de las relaciones entre las variables de entrada y las de salida que dirigen el comportamiento del sistema. La identificación de la estructura, puede dividirse a su vez en dos subtareas: la selección de variables y la generación de reglas. La primera trata de determinar las variables de entrada al sistema. La generación de reglas consiste en encontrar las relaciones entrada-salida que dirigen el comportamiento del sistema. A su vez se divide en la identificación del número de reglas difusas que formarán el modelo y, por otro lado, en la determinación de la estructura de los antecedentes y consecuentes de las reglas difusas. Por último, la optimización de los parámetros consiste en determinar los valores de los parámetros de las funciones de pertenencia (MF del inglés *membership function*) paramétricas que definen los conjuntos difusos de las reglas. Hay que puntualizar que existen representaciones paramétricas para algunos de los operadores difusos que otorgan significado al mecanismo de razonamiento difuso (FRM, del inglés *fuzzy reasoning mechanism*) de un modelo de reglas difusas. Por tanto, en presencia de modelos difusos con FRM paramétrico la fase de optimización de parámetros también se puede referir al ajuste del FRM.

La naturaleza modular del proceso de AFM ha llevado, en los ultimos años, a la aparición de numerosas estrategias híbridas que abordan el problema de AFM mediante combinaciones de técnicas procedentes del *Soft Computing* [7, 5]. De este modo, cada una de las técnicas componentes, aborda una o varias de las tareas que integran el proceso de AFM.

En este trabajo exponemos los aspectos fundamentales de una estructura de clases Java desarrolladas para la programación de experimentos de AFM en los que intervienen varias técnicas de una manera relativamente sencilla. De modo que la funcionalidad perseguida es incorporar y combinar técnicas de AFM para la realización de experimentos de modelado y la reutilización de los modelos aprendidos. La herramienta está basada en los diseños propuestos en [9, 11] de manera que favorece la extensibilidad en cuanto a la incorporación de tipos de modelos difusos no contemplados previamente y en cuanto a la integración de nuevos algoritmos de AFM, ya sea partiendo de implementaciones en Java previas como programándolos ad hoc. La herramienta proporciona los mecanismos para que los modelos difusos se almacenen en ficheros XML, de modo que pueden ser utilizados posteriormente para realizar inferencias o para formar parte de nuevos experimentos de modelado. En el trabajo [8], se pueden consultar los detalles de la metodología en la que nos basamos para el presente desarrollo. En dicho trabajo se puede ver que se trata de una metodología que emplea UML para la estructura de clases, interfaces y casos de uso. En cuanto a la descripción de los métodos que ofrecen estas clases e interfaces se emplea OCL (*Object Constraint Language*).

Debemos mencionar que, a día de hoy, esta API es para el uso interno del grupo de investigación ya que, todavía, se deben perfilar muchos aspectos para mejorar la usabilidad de la herramienta.

En cuanto a los trabajos relacionados cabe destacar el presentado en [11], donde se realiza un estado del arte de los algoritmos de AFM y se hace un estudio de los elementos que intervienen en los modelos de reglas difusas. Partiendo de tales estudios, en los trabajos [10, 9] se sientan las bases para definir algoritmos de modelado difuso compuestos por varias técnicas de una forma constructiva a partir de técnicas monolíticas. Así, dado un conjunto de datos, se le aplica la combinación de tecnicas definida y se obtiene un modelo final.

Tras las fases de análisis y diseño presentadas en los trabajos mencionados, la presente contribución está dedicada a la fase de implementación. No obstante, cabe destacar que la hibridación se consigue aquí a través de la serialización de los modelos en ficheros XML que, posteriormente, pueden ser recuperados para aplicarles subsecuentes técnicas.

Este trabajo está estructurado de la siguiente manera: la sección 2 está dedicada a analizar los principales tipos de modelos que se utilizan en procesos de AFM. Este análisis es la base del diseño de modelos difusos en el que se basa nuestra herramienta que también se explica en dicha sección, junto con la manera en que éstos son serializados en documentos XML. Después, la sección 3 está dedicada al diseño de las clases necesarias para representar técnicas de AFM, describiendo casos de uso de la herramienta a la hora de hacer experimentos de modelado. Posteriormente, la sección 4 expone recomendaciones para incorporar algoritmos nuevos a la herramienta. Finalmente, la sección 5 está dedicada a las conclusiones.

2 Taxonomía de Modelos Difusos

En la figura 1 se muestran las principales entidades que forman este API. Debido a las limitaciones de espacio nos vamos a centrar en explicar aquellas que representan modelos así como las que representan algoritmos de AFM. En esta sección nos centramos en la entidad

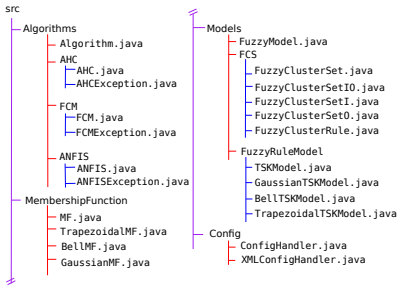


Figura 1: Estructura de clases.

para representar modelos. Como se expone en [10], un modelo es básicamente una forma de describir un *sistema* en base a *relaciones entrada-salida* que determinan su comportamiento. A partir de dichas relaciones y para una entrada dada, el modelo es capaz de inferir una salida simulando al sistema al que describe.

Desde una perspectiva de AFM, consideramos que existen diferentes representaciones del modelo dependiendo de cómo se expresan las relaciones entrada-salida que lo constituyen y del modo en que se infieren las salidas a partir de las entradas [9]. Para ello se tiene en cuenta cómo se realiza la inferencia y lo que se llama el *nivel organizativo*, es decir, la forma en la que las relaciones entrada-salida están definidas.

Un conjunto de datos entrada-salida sobre un sistema objetivo se considera como un modelo del mismo. Un conjunto donde los datos entrada-salida están estructurados en torno a clusters difusos en el espacio entrada-salida, se considera una definición por puntos de una relación entrada-salida.

En el siguiente nivel organizativo encontramos un modelo difuso paramétrico en el que las relaciones entrada-salida son reglas $SI - ENTONCES$ con conjuntos difusos definidos funciones de pertenencia paramétricas cuyos parámetros son ajustables.

El FRM de, por ejemplo, un modelo TSK es fijo. Sin embargo, un modelo difuso puede venir dado en una forma

totalmente paramétrica en la que tanto las funciones de pertenencia como el FRM son susceptibles de ajuste. Es el caso, por ejemplo, de un FRM cuya interpretación venga dada por los operadores de Schweizer y Sklar [3]. De esta forma, el FRM puede ser ajustado mediante alguna técnica de optimización de parámetros.

Los modelos mencionados se diferencian en el nivel organizativo y en la manera en que realizan la inferencia. Sin embargo, es conveniente que las técnicas de AFM los manejen de forma transparente. De modo que, dado un modelo, operaciones muy comunes en el proceso de modelado como, realizar inferencias sobre él, proyectarlo en un subconjunto de sus variables, etc., deben ser posibles para todos los tipos de modelo. Al poder manejar los modelos de forma transparente, se favorece la flexibilidad en la hibridación.

En la figura 2 podemos ver la interfaz `FuzzyModel`, la cual representa el concepto de modelo y reune las funcionalidades que cualquier tipo de modelo tiene que ofrecer. No se muestran las signaturas completas en aras de la claridad. El método `makeInference` genera una salida a partir de unas entradas dadas. El método `getXML` serializa el modelo en XML. En contraposición, el método `getFromXML` crea una instancia de un modelo a partir de un fichero XML. El método `getChart` genera el texto correspondiente a un script para que el modelo sea visualizado con la herramienta `gnuplot`. El método `projectInto`, a partir del modelo actual, genera un modelo igual pero donde se han eliminado las cláusulas correspondientes a un conjunto de variables de entrada pasadas por parámetro.

Obviamente, las clases que implementen esta interfaz deberán proporcionar la manera en que cada uno de estos métodos es llevado a cabo. Así, las clases `FuzzyClusterSet` e `IODataSet` implementarán el método `makeInference` de manera acorde a como se ha expone en [10]. Puesto que nuestro trabajo se centra en el AFM los modelos difusos más usados en este enfoque son los modelos *TSK*. Así que, aunque no se muestra en la figura proporcionamos la clase `TSKModel` genérico para implementar la interfaz `FuzzyModel`.

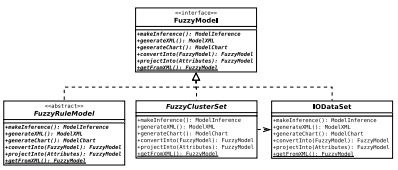


Figura 2: Diagrama de clases de los modelos difusos.

2.1 Serialización y Persistencia de Modelos Difusos

En esta sección describimos los formatos XML utilizados para serializar el modelo. Los ficheros XML que se generan, además de ser válidos, están bien formados dado que su estructura y sintáxis debe ser acorde con determinado DTD (*documento type definition*). Proporcionamos varios DTDs dependiendo del tipo de modelo que se esté representando (conjunto de clusters difusos, modelo *TSK*, datos entrada-salida).

En el ejemplo siguiente se ve un extracto de la serialización en XML de un modelo de clusters difusos. Vemos que, en este caso, se trata de un conjunto con 8 centroides extraído a partir de 998 datos compuestos por 3 atributos de entrada. Trabajamos con modelos MISO (*multiple inputs simple output*), por tanto, se supone siempre un sólo atributo de salida. La distancia a partir de la cual se han calculado los centroides es la euclídea y el valor de *fuzziness* también hay que almacenarlo como parte del modelo. Luego aparece un elemento para cada centroide yun elemento para cada ejemplo de los que componen el conjunto de datos a partir de los cuales se ha conseguido el modelo de clusters.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE FCS_c_8 SYSTEM "fcsmodel.dtd">
<FCSMODEL
    FUZZINESS = "2.05"  NUMCLUSTER = "8"
    NUMEXAMPLES = "998" NUMATRIBUTES = "3"
    PSPACE = "INPUT/OUTPUT" DISTANCETYPE = "EUCLIDEAN">
 <JAVACLASS>FuzzyModel.FuzzyClusterSet</JAVACLASS>
  <CENTROIDS>
    <CENTROID INDEX = "0">
      <VALUE INDEX = "0">3.55</VALUE>
      <VALUE INDEX = "1">8.75</VALUE>
      <VALUE INDEX = "2">1.34</VALUE>
      <VALUE INDEX = "3">1.35</VALUE>
     </CENTROID>
       ...//tantos CENTROID como centroides
   </CENTROIDS>
   <EXAMPLES>
    <EXAMPLE INDEX= "0">
      <VALUE INDEX= "0">210.00</VALUE>
      <VALUE INDEX= "1">202.00</VALUE>
      <VALUE INDEX= "2">3.00</VALUE>
      <VALUEOUTPUT>3.0</VALUEOUTPUT>
    </EXAMPLE>
       ...    //tantos EXAMPLE como ejemplos
  <EXAMPLES>
</FCSMODEL>
```

A continuación se muestra el DTD que expresa la estructura de un modelo *TSK*.

```
<!ELEMENT USER (#PCDATA)>
<!ELEMENT KEY (#PCDATA)>
<!ELEMENT JAVACLASS (#PCDATA)>
<!ELEMENT CONTENTS (INPUTVARIABLE+,OUTPUTVARIABLE, FUZZYRULE+)>
<!ELEMENT INPUTVARIABLE EMPTY>
<!ELEMENT OUTPUTVARIABLE EMPTY>
<!ELEMENT FUZZYRULE (ANTECEDENT,CONSEQUENT)>
<!ELEMENT ANTECEDENT (FUZZYPROPOSITION+)>
<!ELEMENT CONSEQUENT (PVALUE+)>
<!ELEMENT FUZZYPROPOSITION (FUZZYSET)>
<!ELEMENT FUZZYSET (PARAMETERMF+)>
<!ATTLIST TSKMODEL NUMBEROFRULES CDATA              #REQUIRED
                   NUMBEROFINPUTS CDATA             #REQUIRED
                   CONSEQUENTTYPE (SINGLETON|LINEAR) #REQUIRED
                   MFTYPE (GAUSSIAN|BELL|TRAPEZOIDAL)#REQUIRED>
<!ATTLIST INPUTVARIABLE INDEX ID #REQUIRED
                        NAME CDATA #REQUIRED>
<!ATTLIST OUTPUTVARIABLE
                        NAME ID #REQUIRED>
<!ATTLIST FUZZYRULE     INDEX ID #REQUIRED>
<!ATTLIST FUZZYPROPOSITION      RULE IDREF\ #REQUIRED
                                INPUTVARIABLEINDEX IDREF #REQUIRED>
<!ATTLIST PARAMETERMF   NAME CDATA #REQUIRED>
<!ATTLIST PVALUE        RULE IDREF #REQUIRED
                        INPUTVARIABLEINDEX IDREF #REQUIRED>
```

Se puede ver que un modelo *TSK* tiene un número de reglas, un número de entradas, un tipo de consecuente. Además se contemplan varios tipos de funciones de pertenencia para los conjuntos difusos implicados.

Puede verse también, que contiene elementos para las variables de entrada y salida implicadas y para cada una de las reglas. Cada regla está compuesta por un antecedente y un consequente. Los antencedentes contienen proposiciones difusas que involucran a conjuntos difusos. Los consecuentes contienen los coeficientes que acompañan a cada variable de entrada en el polinomio, uno de los cuales corresponderá al término independiente.

A continuación se muestra un segmento de la serialización de un modelo *TSK* lineal con dos variables de entrada llamadas *acceleration* y *speed* y una de salida llamada *maneuver*. Vemos que el término independiente del consecuente de la única regla que aquí se muestra es aquel que tiene valor *NULL* en el atributo que expresa el índice de la variable de entrada.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE FuzzyRuleModel_numrule_05
  SYSTEM "FRModel.dtd">
<FRMODEL
  NUMBEROFRULES = "5"   NUMBEROFINPUTS = "2"
  NUMBEROFOUTPUTS = "1" CONSEQUENTTYPE = "LINEAR"
  MFTYPE = "GAUSSIAN">
 <JAVACLASS>FuzzyModel.FRModel.GaussianFRModel</JAVACLASS>
 <CONTENTS>
   <INPUTVARIABLE INDEX = "0" NAME = "acceleration"/>
   <INPUTVARIABLE INDEX = "1" NAME = "speed"/>
   <OUTPUTVARIABLE NAME = "maneuver"/>
   <FUZZYRULE INDEX= "0">
     <ANTECEDENT>
      <FUZZYPROPOSITION RULE = "0" INPUTVARIABLEINDEX = "0">
        <FUZZYSET>
          <PARAMETERMF NAME ="a">0.17</PARAMETERMF>
          <PARAMETERMF NAME ="c">-0.08</PARAMETERMF>
        </FUZZYSET>
      </FUZZYPROPOSITION>
      <FUZZYPROPOSITION RULE = "0" INPUTVARIABLEINDEX = "1">
        <FUZZYSET>
          <PARAMETERMF NAME ="a">18.87</PARAMETERMF>
          <PARAMETERMF NAME ="c">6.40</PARAMETERMF>
        </FUZZYSET>
      </FUZZYPROPOSITION>
     </ANTECEDENT>
     <CONSEQUENT>
      <P_VALUE RULE= "0" INPUTVARIABLEINDEX="0">0.52</P_VALUE>
      <P_VALUE RULE= "0" INPUTVARIABLEINDEX="1">0.02</P_VALUE>
      <P_VALUE RULE= "0" INPUTVARIABLEINDEX= "NULL">
        1.14
      </P_VALUE>
     </CONSEQUENT>
   </FUZZYRULE>
   ... //tantos FUZZYRULE como reglas
 </CONTENTS>
</FRMODEL>
```

Este formato XML facilita que los modelos serializados en ficheros puedan ser mejorados incorporando conocimiento experto manualmente. Obviamente, también se pueden recuperar dando lugar a instancias de modelos difusos que se podrán utilizar para hacer inferencias o para formar parte de otros experimentos de AFM a través del método `getFromXML`. Para ello, la implementación de dichos métodos en cada modelo concreto debe analizar el texto recuperado de cada fichero. Nuestra herramienta utiliza la librería `Xerces Java Parser` [1] para la generación de dicho analizador.

3 Proceso de Modelado Difuso con el API de AFM

La figura 3 muestra el diagrama de secuencia del sistema para el escenario en el que un cliente, que bien puede ser

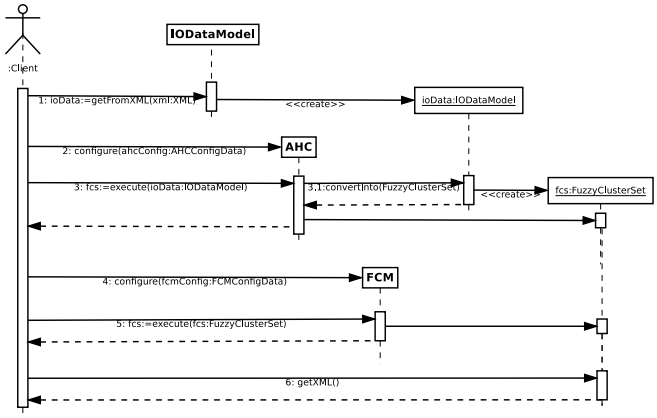


Figura 3: Diagrama de secuencia del sistema para la generación de un conjunto de clusters difusos.

el usuario final o una clase intermedia configurada por el mismo, solicita al sistema un `FuzzyModel` generado a partir de un conjunto de datos determinado. Dicho escenario muestra la capacidad del sistema de, a partir de unos datos en *crudo*, generar un modelo difuso sobre el que se pueden realizar diferentes tipos de operaciones.

Para ello, en primer lugar, el cliente carga los datos de entrada en una instancia de la clase `IODataSet` (mensaje 1 en la figura 3). Dicha instancia representa en el sistema un conjunto de tuplas de entrada con su salida asociada. Como el cliente quiere generar un modelo difuso con un

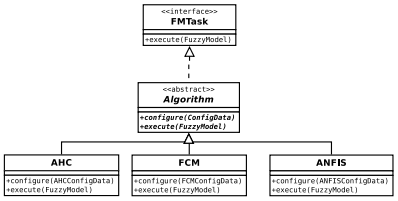


Figura 4: Diagrama de clases de los algoritmos.

determinado número de centroides, hace uso del algoritmo de clustering auto-organizativo AHC (*Agglomerative Hierarchical Clustering*) [3]. Tal y como se expone en la figura 4, este algoritmo se implementa en el sistema por la clase `AHC` que, al igual que el resto de algoritmos, cumple con la interfaz `FMTask`. Esta interfaz se compone básicamente de un único método (`execute`) el cual, tomando como entrada un `FuzzyModel`, devuelve como salida otra instancia de dicha clase que es el resultado de aplicar el algoritmo sobre la instancia de entrada. Dado que la clase `FuzzyModel` es la superclase de todos los modelos difusos contemplados en el sistema, vemos que cualquier algoritmo que implemente dicha interfaz debería ser capaz, en principio, de procesar cualquier modelo difuso, y en caso contrario, generar la correspondiente excepción De este modo, teniendo en cuenta que la clase `IODataSet`

hereda de FuzzyModel, tal y como se indica en la figura 2, el cliente puede pasar su IODataSet como entrada al AHC (mensaje 3), no sin previamente configurar dicho algoritmo indicándole el número de centroides a obtener entre otros parámetros (mensaje 2). Posteriormente, la

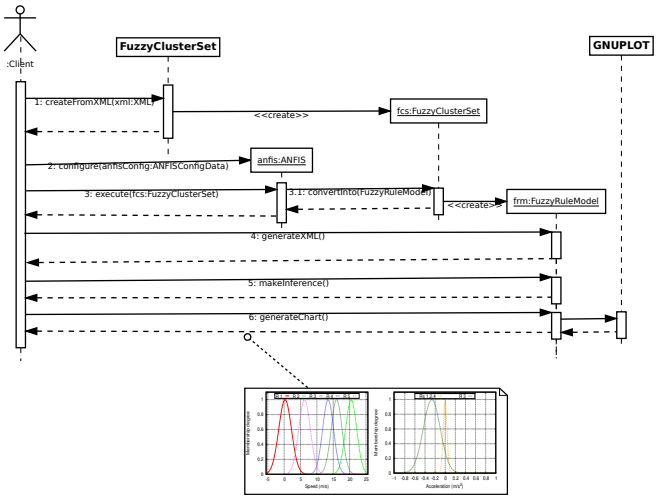


Figura 5: Diagrama de secuencia del sistema para el ajuste de parámetros de un FuzzyModel y su posterior uso.

instancia de AHC intenta transformar el IODataSet pasado como entrada en una instancia de FuzzyClusterSet ya que dicho algoritmo opera sobre conjuntos difusos. Para ello llama al método convertInto de la instancia de IODataSet (mensaje 3.1). Este método permite convertir el FuzzyModel origen al indicado como parámetro y en caso de no poder realizarse la transformación, arroja una excepción. En el presente escenario la conversión es posible, y el IODataSet inicial da origen a una instancia de FuzzyClusterSet (llamada fcs en la figura 3) que contiene tantos centroides como tuplas exiten en la instancia de IODataSet. A continuación, el AHC es aplicado sobre el FuzzyClusterSet obtenido y, finalmente, devuelve al usuario el FuzzyClusterSet con el número de centroides indicado en la configuarción del algoritmo. Sin embargo, debido a la especificación del método execute, el cliente solo verá el que la instancia de AHC le devuelve una instancia de la superclase FuzzyModel.

Una vez realizados todos los pasos anteriores, el cliente decide optimizar los centroides del FuzzyModel obtenido mediante el algoritmo de las C medias (*Fuzzy C-Means*, FCM) [2]. Para ello configura una instancia de la clase FCM (mensaje 4) la cual implementa la interfaz FMTask (ver figura 4). Así, el cliente invoca al método execute de la instancia pasándole como parámetro el FuzzyClusterSet devuelto por el AHC (mensaje 5). En esta ocasión, el FCM no tiene que realizar ningún tipo de conversión sobre el FuzzyModel de entrada ya que en realidad es una instancia de FuzzyClusterSet que es la clase sobre la que puede operar el FCM. Finalmente, cuando el cliente recibe el FuzzyModel devuelto por el FCM, el cual es una instancia de FuzzyClusterSet, decide obtener su representación en XML mediante la llamada a getXML (mensaje 6). Así, el cliente podrá guardar dicha representación en el soporte que mejor le convenga (un fichero de texto plano en el sistema de directorios local o un servidor remoto, etc.) y recuperar el FuzzyModel en un futuro para poder realizar diferentes inferencias con él. De esta forma, la generación y el uso de los modelos difusos queda completamente desacoplada.

Por tanto, vemos que a lo largo de todo el escenario expuesto en la figura 3 el cliente siempre ha ido utilizando los diferentes algoritmos sin necesidad de preocuparse por cuál era el modelo difuso subyacente, ya que eran los algoritmos los encargados de establecer el FuzzyModel concreto que debía ser utilizado en cada momento mediante las conversiones oportunas.

La figura 5 muestra el diagrama de flujo del sistema en el que un cliente decide usar un modelo difuso para realizar una inferencia, no sin antes ajustar sus parámetros mediante el modelo de red adaptativa neuro-difusa ANFIS [4]. Además, también decide guardar el modelo ajustado y obtener la representación gráfica del mismo. Para ello, el cliente en primer lugar recupera a partir de su serializacion en XML el FuzzyModel que desea usar (mensaje 1). Dicho modelo podría ser perfectamente el guardado en un experimento anterior. A continuación, configura una instancia de ANFIS (mensaje 2), que en el sistema es considerado una subclase de Algorithm, e invoca a su método execute pasándole como parámetro el FuzzyModel que acaba de recuperar (mensaje 3). Internamente, la instancia de ANFIS necesita transformar el modelo de entrada en una instancia de FuzzyRuleModel (mensaje 3.1), ya que la instancia que el cliente ha pasado al algoritmo tiene a la clase FuzzyClusterSet como clase subyacente, sobre la cual el algoritmo ANFIS no puede operar directamente. Una vez que la instancia de ANFIS devuelve el modelo difuso con sus parámetros ajustados, el cliente decide obtener su serialización en XML (mensaje 4), para poder almacenarla en algún lugar. Así, vemos que llegados a esta punto, el cliente dispone de la representación del model difuso inicial (que usó en el mensaje 1) y la del obtenido de ajustar sus parámetros mediante ANFIS para poder usar ambas en un futuro. Después, el cliente decide usar el modelo ajustado para realizar una serie de inferencias (mensaje 5) mediante la llamada al método makeInference proporcionado por la interfaz FuzzyModel. Finalmente, el cliente obtiene la representación gráfica del modelo ajustado llamando al método generateChart (mensaje 6). Dicho método hace uso del programa para la generación de gráficas GNUPlot, obteniéndose como resultado una gráfica con el mismo formato que la mostrada como comentario de la figura 5.

4 Incorporación de una Técnica de AFM en el API

La interfaz `FMTask` constituye el envoltorio con el que se tiene que recubrir una técnica para que se considere integrada en el API. Desde la perspectiva del diseño OO, esto se traduce en la obtención de una clase que implemente la interfaz `FMTask`, ofreciendo así todas las funcionalidades que reune. Existen dos posibilidades: la integración partiendo de una implementación existente y la integración de una técnica desde cero. El primer caso se supone que existe una clase `XImpl` que contiene el código de `X`. De modo que tenemos que crear otra clase, llamémosla `XTask`, que hereda de `XImpl` el código del algoritmo `X` y que, por otro, lado implementa las funcionalidades descritas en la interfaz `FMTask`. El segundo caso se quiere integrar un

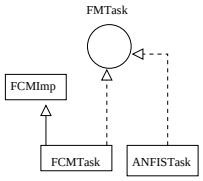


Figura 6: Incorporación de un algoritmo ya existente.

algoritmo X desde cero y hay que programarlo en una nueva clase `XTask` que implemente la interfaz `FMTask`. Así, `XTask` ofrecerá todas las funcionalidades de `XTask`.

En la figura 6 se ilustra la incorporación de dos algoritmos sobradamente conocidos como el FCM y el método ANFIS. En el primero se supone que se parte de una clase `FCMImpl` que codifica en algoritmo. La clase `FCMTask` debe heredar de ella e implementar la interfaz `FMTask`. La figura muestra que la integración de ANFIS se hace desde cero, de modo que la clase `ANFISTask` directamente codifica el algoritmo a la vez que implementa `FMTask`.

5 Conclusiones

En este documento hemos abordado algunas de las cuestiones más importantes de una API Java que utilizamos en nuestro grupo de trabajo para la implementación de experimentos de modelado difuso aproximativo. Esta API se ha implementado siguiendo algunas de las especificaciones procedentes de trabajos anteriores de los autores. La herramienta posibilita la incorporación de nuevos tipos de modelos y de nuevos algoritmos. Por otra parte, permite almacenar los modelos aprendidos en formatos basados en XML y estructurados mediante DTDs. De esta manera, tales modelos pueden ser recuperados posteriormente para mejorarlos, ya sea de forma automática, mediante la aplicación de nuevos algoritmos de modelado, o ya sea mediante la incorporación manual de conocimiento experto dada la legibilidad del formato.

Referencias

[1] Xerces java parser. `http://http://xerces.apache.org/`.

[2] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum, New York, USA, 1981.

[3] M.R. Emami, I.B. T*ü*rksen, and A.A. Goldenberg. Development of a systematic methodology for fuzzy-logic modeling. *IEEE Transactions on Fuzzy Systems*, 6(3):346–361, August 1998.

[4] J.S. Jang. Anfis: Adaptive-network-based fuzzy inference systems. *IEEE Trans. on Systems, Man, and Cybernetics*, 23(03):665–685, 1993.

[5] M. Setnes and H. Roubos. GA-fuzzy modeling and classification: Complexity and performance. *IEEE Transactions of Fuzzy Systems*, 8(5):509–522, 2000.

[6] M. Sugeno and T. Yasukawa. A fuzzy logic based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*, 1(1):7–31, August 1993.

[7] F. Terroso-Sáenz and M. Valdés-Vela. Fuzzy modeling for vehicle maneuver detection in a scene. In *IEEE World Congress on Computational Intelligence*, pages 1688–1695, July 2010.

[8] M. Valdés. *Análisis y Diseño de un Marco para la Definición de Técnicas Híbridas de Modelado Difuso (METHOD)*. PhD thesis, Universidad de Murcia, 2003.

[9] M. Valdés, J. A. Botía, and A. F. G. Skarmeta. Towards a formal framework for the specification of fuzzy modeling. In *Proceedings of the International Conference on Fuzzy Systems*, St. Louis, MO. USA., May 2003.

[10] M. Valdés, A. Gómez-Skarmeta, and J. Botía. Method: Un enfoque constructivo para modelado difuso híbrido. In *Actas del XII Congreso Español Sobre Tecnologías y Lógica Fuzzy*, pages 339–344, 2004.

[11] M. Valdés, A. F. Gómez-Skarmeta, and J. A. Botía. Toward a framework for the specification of hybrid fuzzy modeling. *International Journal of Intelligent Systems*, 20(2):225–252, 2005.

[12] L. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions of Systems, Man and Cybernetics*, SMC-3:28–44, 1973.

SÍNTESIS AUTOMÁTICA DE SISTEMAS DIFUSOS MEDIANTE XFUZZY

Santiago Sánchez Solano[1], María Brox Jiménez[2]

[1] *Instituto de Microelectrónica de Sevilla, IMSE-CNM-CSIC*, santiago@imse-cnm.csic.es

[2] *Dpto. Arquitectura de Computadores, Universidad de Córdoba*, mbrox@uco.es

Resumen

En esta comunicación se describen las dos herramientas para síntesis hardware de sistemas de inferencia difusos incluidas en el entorno *Xfuzzy*. Ambas herramientas comparten una arquitectura de implementación común, aunque están basadas en metodologías de diseño diferentes. *Xfvhdl* (*Xfuzzy to VHDL*) genera código VHDL estándar que puede ser posteriormente implementado mediante un ASIC o una FPGA. *Xfsg (Xfuzzy to SysGen)* se apoya en el entorno Matlab/Simulink y las herramientas de desarrollo de sistemas de procesado digital de señal sobre FPGAs de Xilinx. Como demuestran los ejemplos considerados, el empleo de estas herramientas permite acelerar considerablemente las etapas de descripción, verificación y síntesis de sistemas de control basados en lógica difusa.

Palabras Clave: Sistemas de inferencia difusos, Herramientas de CAD, Implementación Hardware.

1 INTRODUCCIÓN

La capacidad de la lógica fuzzy para modelar el comportamiento de sistemas complejos, así como para describir estrategias de control basadas en reglas similares a las empleadas en el lenguaje natural, ha motivado su empleo en numerosas aplicaciones relacionadas con la automatización industrial, la robótica y otros campos de la ingeniería [7]. Las primeras realizaciones de sistemas de inferencia se realizaron mediante software ejecutado en procesadores de propósito general. Posteriormente, en la década de los 90, se desarrollaron una serie de 'coprocesadores difusos' cuyo objetivo era acelerar las operaciones necesarias para llevar a cabo el proceso de inferencia. Sin embargo, como consecuencia de los continuos avances en las tecnologías de fabricación de circuitos integrados, en los últimos años se han propuesto numerosas alternativas para la implementación hardware de sistemas difusos mediante ASIC o FPGAs [3], [4].

El importante éxito de la lógica fuzzy en las últimas décadas del siglo pasado motivó el desarrollo de numerosas herramientas dedicadas al diseño de este tipo de sistemas. En los primeros años, la mayoría del sofware disponible comercialmente estaba dirigido hacia la generación de código optimizado para las diferentes familias de microcontroladores de la época. Posteriormente aparecieron numerosas propuestas para la síntesis automática de sistemas difusos mediante hardware específico basado en técnicas de diseño analógicas y digitales [5], [6], [8]. Más recientemente, han surgido diferentes alternativas que aprovechan las herramientas incluidas en el entorno Matlab/Simulink para facilitar el diseño e implementación de sistemas difusos [1], [2], [9].

En este trabajo se describen las últimas versiones de las dos herramientas de síntesis hardware de sistemas difusos integradas en el entorno de desarrollo *Xfuzzy* [14]. Una de ellas, denominada *Xfvhdl*, constituye una actualización de la herramienta incluida en versiones anteriores del entorno, y traslada la especificación del sistema difuso en una descripción VHDL que puede ser implementada en las tecnologías de destino disponibles en las herramientas estándar de síntesis para ASICs o FPGAs. La segunda herramienta, *Xfsg*, que utiliza el software SysGen para desarrollo sobre FPGAs de Xilinx, permite completar el flujo de diseño sobre el entorno Matlab/Simulink y aprovecha todas las opciones de configurabilidad y flexibilidad que ofrece dicho entorno.

En la Sección 2 se introducen las principales características del entorno *Xfuzzy* y sus diferentes herramientas. La estructura de circuitos en la que se basan las dos herramientas de síntesis es descrita en la Sección 3. Las funcionalidades y opciones de ambas herramientas se detallan en las Secciones 4 y 5. La Sección 6 ilustra un ejemplo de aplicación a un problema clásico de control. Por último, las principales conclusiones obtenidas tras la realización del trabajo son resumidas en la Sección 7.

2 EL ENTORNO *XFUZZY*

El entorno de desarrollo de sistemas difusos *Xfuzzy* proporciona un conjunto de herramientas que facilitan las sucesivas fases de descripción, verificación y síntesis de sistemas de inferencia basados en lógica difusa. Un sistema difuso se describe en *Xfuzzy* mediante una especificación XFL3 que combina bases de reglas fuzzy y módulos crisp (para realizar tareas de inferencia e implementar bloques lógicos y aritméticos, respectivamente). Las bases de conocimiento pueden ser definidas directamente con ayuda de la interfaz gráfica de usuario (*Xfedit*), o extraídas a partir de datos numéricos usando herramientas de minería de datos (*Xfdm*). Tanto las funciones de pertenencia como las bases de reglas pueden ser simplificadas para mejorar su interpretabilidad lingüística (*Xfsp*). También es posible ajustar los parámetros del sistema usando diferentes algoritmos de aprendizaje supervisado (*Xfsl*). La verificación funcional del sistema puede llevarse a cabo mediante dos herramientas del entorno. La primera de ellas permite analizar la relación entrada/salida del sistema (*Xfplot*), mientras que la segunda permite simular su comportamiento en lazo cerrado en combinación con un modelo Java de la planta (*Xfsim*). Una vez validada, la especificación puede trasladarse a código C, C++ o Java para obtener una implementación software del sistema de inferencia.

La Figura 1 ilustra la interfaz gráfica de usuario de *Xfuzzy* cuando se utiliza para definir un controlador difuso para el problema del doble integrador. El sistema está formado en este caso por dos bases de reglas y un bloque crisp que realiza la operación aritmética de sustracción.

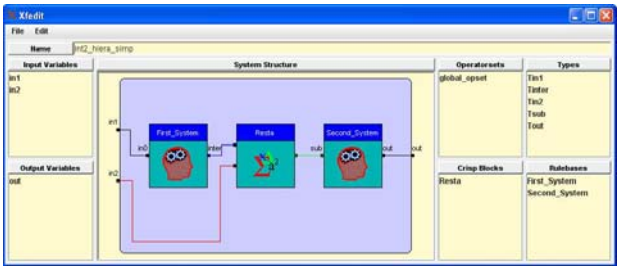


Figura 1. Descripción en *Xfuzzy* de un controlador difuso para el problema del doble integrador.

3 ARQUITECTURA BASADA EN REGLAS ACTIVAS

Las herramientas de síntesis hardware incluidas en *Xfuzzy* facilitan la implementación microelectrónica de sistemas difusos de acuerdo con la arquitectura mostrada en la Figura 2. Las características que definen esta arquitectura son la limitación del grado de solapamiento de las funciones de pertenencia de los antecedentes, la utilización de

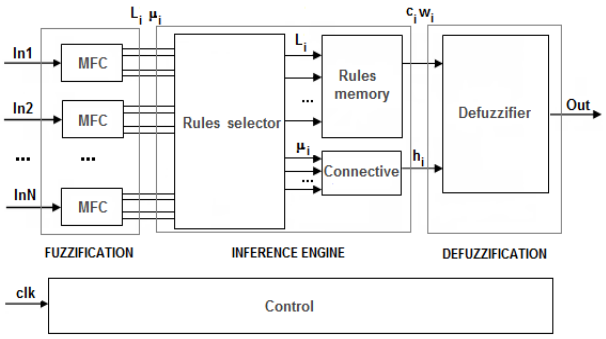


Figura 2. Arquitectura de reglas activas para la implementación hardware de sistemas difusos.

métodos de defuzzificación simplificados y el empleo de un esquema de procesado basado en reglas activas [11]. En la Figura 2 aparecen las tres etapas típicas en el cálculo de una inferencia difusa: fuzzificación, inferencia y defuzzificación. En la etapa de fuzzificación los circuitos generadores de funciones de pertenencia (MFC) suministran las parejas de datos etiqueta-grado de pertenencia (L_i, μ_i) para cada valor de las entradas del sistema. Las distintas combinaciones de estas etiquetas determinarán las reglas activas. En la etapa de inferencia se procesan secuencialmente cada una de las reglas activas, utilizándose para ello un circuito selector de reglas activas formado por un conjunto de multiplexores controlados por un contador. En cada ciclo de reloj los grados de pertenencia μi de cada una de las entradas son combinados a través del conectivo de antecedentes para calcular el grado de activación de la regla (α'), mientras que las respectivas etiquetas de los antecedentes direccionan la posición de memoria que contiene los parámetros (c_r, w_r) que representan su correspondiente consecuente. En la etapa de defuzzificación se calcula el valor de salida procesando los consecuentes de cada regla con sus diferentes grados de activación de acuerdo con el método de defuzzificación elegido. Finalmente un bloque de control se encarga de regular la temporización del proceso de inferencia.

Esta arquitectura se caracteriza por ser altamente configurable debido a la disponibilidad de diferentes opciones para implementar cada uno de los bloques que la forman. Los MFCs pueden ser implementados mediante almacenamiento en memoria o cálculo aritmético. La estrategia basada en almacenamiento en memoria almacena en cada localización los valores de las etiquetas y los grados de pertenencia correspondientes a cada punto del universo de discurso. Por otro lado, la estrategia basada en técnicas aritméticas emplea circuitos aritméticos para implementar familias de funciones triangulares normalizadas (correspondientes a los tipos *triangular* y *sh_triangular* definidos en XFL3) usando una memoria que almacena los puntos de corte y pendientes para cada antecedente. El diseñador puede elegir también implementar los conectivos de antecedentes mediante operadores mínimo o pro-

ducto. Finalmente, en la etapa de defuzzificación, dependiendo de si las bases de reglas corresponden a aproximadores o elementos de toma de decisiones, se pueden emplear diferentes defuzzificadores: FuzzyMean, Weighted-FuzzyMean, Takagi-Sugeno de orden uno o MaxLabel.

4 HERRAMIENTA *XFVHDL*

Para automatizar el proceso de diseño con la arquitectura descrita en la sección anterior, la técnica basada en VHDL utiliza una librería de bloques funcionales parametrizables y sintetizables descritos mediante lenguaje VHDL (la librería *XfuzzyLib_src*). La librería incluye bloques correspondientes a los distintos elementos de la arquitectura mostrada en la Figura 2, además de un conjunto de bloques crisps que implementan operaciones aritméticas y funciones lógicas. La construcción de un sistema difuso mediante esta técnica requiere seleccionar e interconectar los bloques necesarios, así como generar las descripciones VHDL que definen los parámetros de las bases de conocimiento. El uso del lenguaje VHDL, y el hecho de que las descripciones cumplen con las restricciones impuestas por las principales herramientas de síntesis de circuitos integrados, permite que con esta librería puedan diseñarse tanto ASICs como FPGAs siguiendo el flujo de desarrollo habitual en este tipo de herramientas.

La herramienta de síntesis *Xfvhdl* facilita la traslación automática de una especificación XFL3 en una descripción VHDL acorde con esta técnica de diseño y que puede ser implementada directamente o incluida como componente de un sistema empotrado. La interfaz gráfica de usuario de esta herramienta se muestra en la Figura 3. La parte superior de la ventana recoge información sobre los ficheros y directorios utilizados. La zona central izquierda muestra la base de conocimiento estructurada en componentes que son agrupados bajo las categorías “RuleBases” y “CrispBlocks”. Cuando se selecciona una base de reglas, la zona de la derecha de la interfaz muestra información extraída de la especificación XFL3 sobre las funciones de pertenencia y las reglas. En esta misma zona, el usuario puede definir los parámetros relacionados con la dimensión del sistema, así como seleccionar la estrategia de implementación de los antecedentes.

Las herramientas de diseño de FPGA permiten normalmente inferir el uso de memoria a partir de descripciones VHDL genéricas en la etapa de síntesis y seleccionar el tipo de memoria a utilizar en la etapa de implementación. Esta característica permite considerar distintas opciones de implementación para los antecedentes y la base de reglas de un sistema difuso. Por este motivo, en la parte derecha de la interfaz también puede seleccionarse el tipo de memoria (RAM, ROM o bloque lógico) utilizada en ambos casos y la técnica empleada en los MFCs.

Cuando las opciones arquitecturales y los parámetros relacionados con el tamaño de los buses han sido definidos para todos los componentes del sistema, aparece una marca verde en el nivel superior de la jerarquía y se activan los diferentes botones de la zona inferior de la ventana. En particular, es posible generar los ficheros de salida presionando el botón “*Generate VHDL code*”. Básicamente se genera una descripción VHDL y un testbench

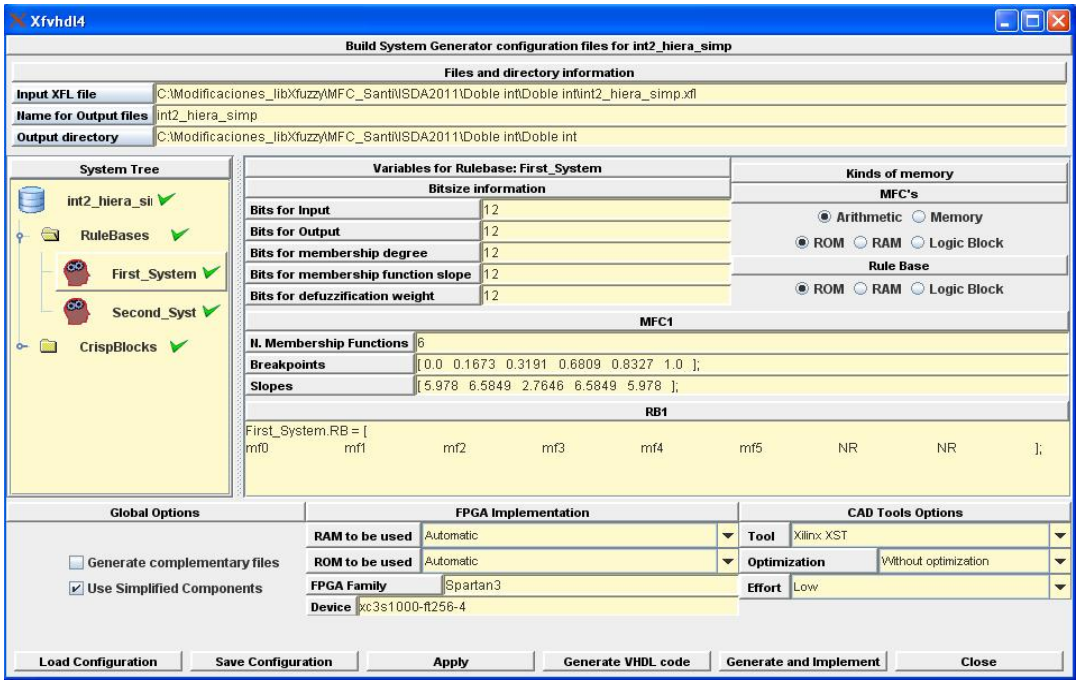


Figura 3. Interfaz gráfica de la herramienta *Xfvhdl*.

(descrito también en VHDL para verificar la funcionalidad del módulo) para cada una de las bases de reglas, junto con la descripción VHDL y un testbench para el nivel superior de la jerarquía.

Para cada una de las bases de reglas, los bloques correspondientes al método de defuzzificación o el operador usado como conectivo se extraen directamente de la especificación XFL3. Sin embargo, el estilo de la descripción VHDL y los componentes que incluye dependen de las opciones arquitecturales escogidas. Concretamente, si se selecciona la opción de bloque lógico para las memorias de antecedentes o las de reglas, la descripción generada utiliza una estructura tipo CASE. Cuando la opción escogida es ROM, se genera una estructura tal que, si el usuario selecciona la opción correspondiente, la herramienta de síntesis puede extraer la memoria ROM para implementar estas descripciones. Finalmente, si la opción seleccionada es RAM, se incluye un bloque de la librería *XfuzzyLib_src* que será identificado como una memoria RAM por la mayoría de herrarmienta de síntesis.

Para el caso de implementación en FPGAs de Xilinx, es posible seleccionar el tipo de memoria (BRAM o CLBs) que va a ser utilizada en el dispositivo programable. Esta elección, junto con otras opciones relativas al dispositivo y herramientas de síntesis empleadas, puede llevarse a cabo mediante los campos que aparecen en la zona inferior de la ventana principal de *Xfvhdl*. Además de las descripciones VHDL, la herramienta genera una serie de ficheros que facilitan la síntesis e implementación del sistema sobre FPGAs, y permite lanzar ambos procesos desde la interfaz de usuario mediante el botón "*Generate and implement*".

5 HERRAMIENTA *XFSG*

La segunda técnica de síntesis considerada está basada en SysGen, la herramienta para desarrollo de sistemas de procesado digital de señal sobre FPGAs de Xilinx. Esta herramienta está integrada en el entorno Matlab e incluye una librería de módulos de Simulink, denominada "Xilinx Blockset", así como el software necesario para trasladar un modelo Simulink en una descripción de hardware que puede ser implementada en una FPGA [13].

Haciendo uso de esta librería y utilizando las facilidades proporcionadas por SysGen, se ha generado una librería de módulos especializados, *XfuzzyLib*, compuesta por bloques que cubren las tres etapas de fuzzificación, inferencia y defuzzificación de la arquitectura de sistemas difusos descrita en la Sección 3. A partir de los bloques de esta librería, la construcción de un sistema difuso requiere la elección, interconexión y parametrización de los bloques necesarios. Los distintos módulos de la librería "Xilinx Blockset" admiten una serie de parámetros que definen aspectos tales como su funcionalidad, tamaño o tipo de aritmética utilizada. Del mismo modo, una vez realizado el diagrama de bloques de un nuevo módulo de la librería *XfuzzyLib*, se encapsula como un "subsistema" y se añade una "máscara" para identificar los distintos parámetros que caracterizan al módulo. De esta forma, cuando dicho subsistema es usado en un nivel jerárquico superior, estos parámetros pueden asignarse mediante valores numéricos o mediante variables a las que se les asigna un valor numérico en la ventana de comandos de Matlab o a través de un fichero ".m". La funcionalidad del diseño puede verificarse en todo momento con ayuda del simulador Simulink y sus diferentes facilidades para generar señales de excitación y para capturar y representar gráficamente los datos de salida.

Además de los bloques que constituyen los elementos constructivos básicos de un sistema difuso, la librería *XfuzzyLib* incluye bloques con descripciones de arquitecturas tipo de módulos de inferencia difusos (FLCs) que se diferencian entre sí en el número de entradas, el conectivo que usan para calcular el grado de activación de las reglas y el método de defuzzificación empleado. Es posible combinar de forma jerárquica estos módulos de inferencia simples para definir sistemas difusos de mayor complejidad. Los bloques que describen arquitecturas de FLCs son totalmente parametrizables, por lo que el diseñador de sistemas difusos puede adaptar la funcionalidad del módulo a los requerimientos del problema planteado sin más que definir los parámetros adecuados. Básicamente hay dos tipos de parámetros: los que corresponden al dimensionado del sistema de inferencia, como son el número de bits de las entradas, salidas y grados de pertenencia, y los que definen la base de conocimiento del mismo, es decir, las funciones de pertenencia y la base de reglas. Para facilitar su uso por parte del diseñador, los distintos parámetros corresponden a variables y estructuras de datos a las que pueden asignarse valores a través de la propia interfaz gráfica de Simulink o utilizando un fichero ".m" de Matlab. También a este nivel es posible aprovechar toda la funcionalidad del entorno Matlab para verificar la funcionalidad del sistema de inferencia [12].

La herramienta *Xfsg*, incorporada al entorno de desarrollo *Xfuzzy*, permite generar los ficheros requeridos para automatizar el flujo de diseño descrito anteriormente. La interfaz gráfica de usuario de esta herramienta es muy similar a la de *Xfvhdl*. Una vez identificados todos los componentes del sistema difuso, *Xfsg* genera un fichero ".mdl", que contiene el modelo Simulink del controlador difuso, y un fichero ".m", que contiene los parámetros que definen el tamaño y la funcionalidad de los diferentes componentes del sistema de inferencia. El modelo obtenido incluye un bloque "System Generator" que facilita la síntesis hardware del diseño descrito en Simulink mediante la traslación del modelo a diferentes tipos de netlists. También puede generar el fichero de programación de la FPGA, e incluso permite la verificación funcional del controlador en lazo cerrado a través de la cosimulación hardware/software de la implementación del sistema de control junto a un modelo de la planta.

6 APLICACIÓN A UN PROBLEMA DE CONTROL

Las herramientas de síntesis descritas anteriormente han sido utilizadas para la implementación de sistemas difusos aplicados a un problema típico dentro de la ingeniería de control, el doble integrador. La metodología de diseño descrita en esta sección combina el uso de las herramientas de síntesis hardware del entorno *Xfuzzy*, herramientas de síntesis e implementación de FPGAs de Xilinx, y herramientas de simulación y modelado del entorno Matlab/ Simulink. El flujo de diseño se ilustra en la Figura 4.

La descripción del controlador en *Xfuzzy*, correspondiente al sistema mostrado en la Figura 1, constituye la entrada a las dos herramientas de síntesis. La funcionalidad de los resultados obtenidos con ambas herramientas pueden ser verificadas usando las facilidades gráficas y de simulación de ModelSim (Figura 5b) y Matlab (Figura 5c), respectivamente, y comparadas con la prevista en el entorno *Xfuzzy* (Figura 5a). Como se muestra en la Figura 4 las metodologías de ambas técnicas de diseño siguen a priori flujos diferentes. Sin embargo, Simulink permite desarrollar un modelo en el que el código VHDL generado por *Xfvhdl* puede ser incluido en un bloque "Black Box" junto con un bloque ModelSim, ambos proporcionados por la librería "Xilinx Blockset", para realizar una cosimulación HDL.

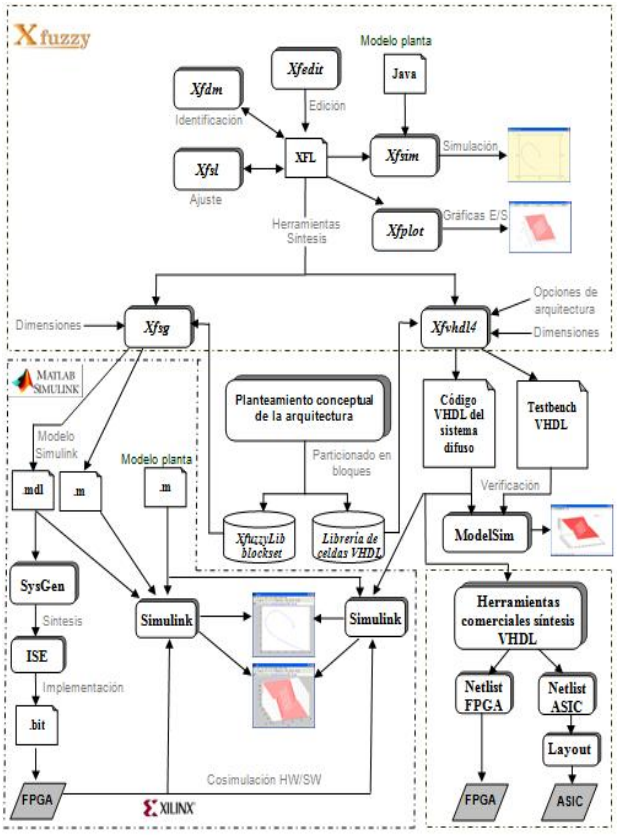


Figura 4. Flujo de diseño soportado por las dos herramientas de síntesis hardware de *Xfuzzy*.

Finalmente también es posible realizar en Simulink la verificación funcional en lazo cerrado de cualquiera de las dos implementaciones del controlador mediante la cosimulación de un modelo software de la planta descrito en Matlab y la implementación hardware del controlador en una FPGA.

La realización hardware de este controlador usando técnicas aritméticas en los antecedentes y memoria ROM de tipo distribuida (con 12 bits de precisión en las entradas y salidas para todas las bases de reglas) consume 185 Slices con *Xfvhdl* y 259 con *Xfsg* (aproximadamente el 3% y 4%, respectivamente, de los Slices disponibles en una FPGA Spartan 3A de Xilinx). El controlador también emplea para ambas técnicas 4 de los 20 multiplicadores disponibles en la FPGA.

7 CONCLUSIONES

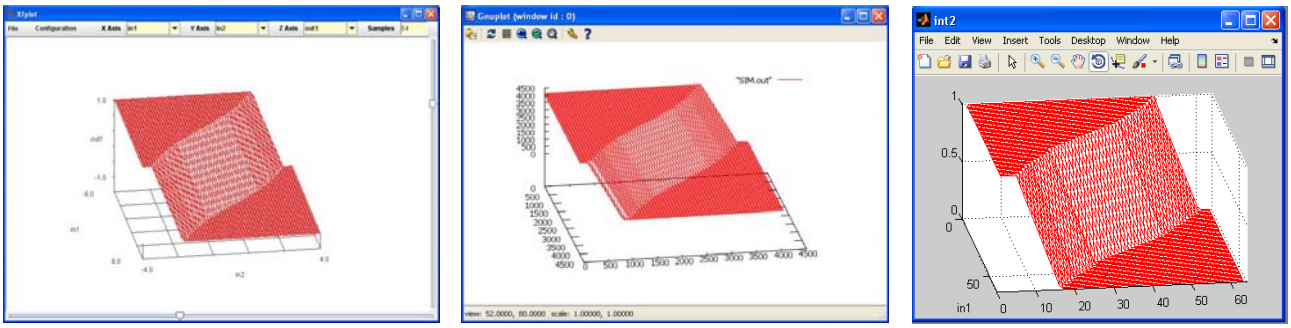
Las dos herramientas de síntesis automática de sistemas de inferencia fuzzy presentadas en esta comunicación permiten poner de manifiesto que la disponibilidad de flujos de diseño, soportados por el uso de librerías de celdas parametrizadas y herramientas de CAD, acelera considerablemente la implementación hardware de sistemas difusos, facilitando la exploración del espacio de diseño para una determinada aplicación. Una de las herramientas descritas está enfocada hacia la implementación hardware de sistemas difusos sobre FPGAs de Xilinx, mientras que la otra proporciona código VHDL sintetizable sobre cualquier tecnología (FPGA o ASIC) disponible para las herramientas de síntesis comerciales.

Comparadas con las versiones previas de las herramientas de síntesis de *Xfuzzy* [10], las nuevas herramientas proporcionan una funcionalidad mejorada de la mayoría de los componentes incluidos en las librerías VHDL y Simulink, la implementación de algunos de los tipos de familias de funciones de pertenencia definidos en el lenguaje de especificación XFL3, nuevos operadores y métodos de defuzzificación y, como avance más significativo, la implementación directa de sistemas difusos complejos compuestos por una combinación de bases de reglas y módulos crisp.

Las ventajas de las herramientas y técnicas de diseño propuestas han quedado demostradas mediante su aplicación en un ejemplo de diseño.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la Comunidad Europea con el proyecto MOBY-DIC FP7-IST-248858, el Ministerio de Ciencia e Innovación con el proyecto TEC2008-04920, y la Junta de Andalucía con el proyecto P08-TIC-03674 (con soporte FEDER).



(a) (b) (c)

Figura 5. Superficies de control obtenidas por: a) *Xfuzzy*. b) *Xfvhdl*. c) *Xfsg*

Referencias

[1] I. H. Altas, A. M. Sharaf: A Generalized Direct Approach for Designing Fuzzy Logic Controllers in Matlab/Simulink GUI Environment. *International Journal of Information Technology and Intelligent Computing*, n.4 vol.1, 2007.

[2] A. Bakhti, L. Benbaouche: Simulink-DSP Co-Design of a Fuzzy Logic Controller. In *Proc. Industrial Electronics Society Annual Conference*, vol.1, pp. 4587-4592. Nov. 2006.

[3] K. Basterretxea, I. del Campo: Electronic hardware for fuzzy computation, in Laurent, A., and Lessot, M. J. (Eds.): *Scalable Fuzzy Algorithms for Data Management and Analysis: Methods and Design* (Information Science Reference, 2009), pp. 1-30.

[4] I. Baturone, A. Barriga, S. Sánchez-Solano, C. J. Jiménez, D. López: *Microelectronic Design of Fuzzy Logic-Based Systems*, CRC Press, 2000.

[5] R. G. Carvajal, A. Torralba, L.G. Franquelo: AFAN: a tool for the automatic synthesis of neural and fuzzy controllers with architecture optimization. In *Proc. International Symposium on Circuits and Systems*, vol. 1, pp. 637-640. 1997.

[6] T. Hollstein, S. K. Halgamuge, M. Glesner: Computer-Aided Design of Fuzzy Systems Based on Generic VHDL Specifications. *IEEE Transactions on Fuzzy Systems*, vol. 4, n. 4, pp. 403-417. 1996.

[7] J. Jarris: *Fuzzy logic applications in engineering science*. Springer Verlag, 2006.

[8] D. Kim, In-Hyun Cho: FADIS: An Integrated Development Environment for Automatic Design and Implementation of FLC. In *Proc. 1997 Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 33-39. Sep. 1997.

[9] O. Kobyrynka, Y. Stekh, O. Markelov: Comparison analysis of methods implemented in MATHLAB for fuzzy logic algorithms: In *Proc. 2011 CAD Systems and Microelectronics*, pp. 239-240. Feb. 2011.

[10] E. Lago, C. J. Jiménez, D. R. López, S. Sánchez-Solano, A. Barriga: XFVHDL: A Tool for the Synthesis of Fuzzy Logic Controllers, In *Proc. Design, Automation and Test in Europe*, pp. 102-107, Feb. 1998.

[11] S. Sánchez-Solano, A. Cabrera, I. Baturone, F. J. Moreno-Velo, M. Brox: Implementation of Embedded Fuzzy Controllers for Robotic Applications, *IEEE Trans. on Industrial Electronics*, vol. 54, n. 4, pp. 1937-1945. Aug 2007.

[12] S. Sánchez-Solano, E. del Toro, M. Brox, I. Baturone, A. Barriga: A Design Environment for Synthesis of Embedded Fuzzy Controllers on FPGAs" In *Proc. IEEE World Congress on Computational Intelligence*, pp. 44-51, Jul. 2010.

[13] System Generator for DSP User Guide, v10.1, Xilinx Inc., Available: http://www.xilinx.com

[14] *Xfuzzy*: Fuzzy Logic Design Tools, IMSE-CNM. Available: http://www.imse-cnm.csic.es/Xfuzzy

XFSML: UN LENGUAJE DE MODELADO DE SISTEMAS DIFUSOS BASADO EN XML

Ramón Santano Ruiz [1], Francisco José Moreno Velo [2]

[1] *Universidad de Huelva*, ramón.santano@alu.uhu.es

[2] *Dpto. Tecnologías de la Información, Universidad de Huelva*, francisco.moreno@dti.uhu.es

Resumen

Este trabajo presenta un nuevo lenguaje de modelado de sistemas difusos llamado XFSML. Se trata de un lenguaje basado en XML que se propone como punto de partida para desarrollar un lenguaje de modelado estándar en la comunidad fuzzy. La principal característica del lenguaje es su gran capacidad expresiva, así como su independencia respecto a plataformas, herramientas o lenguajes de programación concretos.

Palabras Clave: XML, Fuzzy Modeling, Fuzzy Software.

1 INTRODUCCIÓN

En el ámbito científico los sistemas difusos han sido utilizados con éxito en múltiples campos como la ingeniería de control, el tratamiento de imágenes, el reconocimiento de patrones, la minería de datos o los sistemas de toma de decisiones, entre otros. Esto ha generado una floreciente comunidad científica dedicada al desarrollo teórico y práctico de los sistemas difusos, con un nivel de producción de cientos de artículos científicos al año.

La situación es bastante diferente en cuanto a la implantación de los sistemas difusos en el ámbito industrial. La mayor parte de las aplicaciones de sistemas difusos en la industria están dedicadas a resolver problemas de automática y control. En el resto de campos de aplicación, los sistemas difusos tienen, hoy por hoy, una presencia en la industria casi residual. Esto contrasta con la situación de otros paradigmas de la Inteligencia Artificial, como las redes neuronales o los sistemas basados en reglas clásicas, que sí se han incorporado con éxito a aplicaciones del ámbito empresarial (*business intelligence*), como los sistemas de extracción de conocimiento (KDD), los sistemas expertos o los sistemas de ayuda a la toma de decisiones.

Algunos autores, como D. Nauck [8], consideran que los sistemas difusos no han alcanzado el mismo grado de implantación que otros paradigmas de la Inteligencia Artificial debido entre otras razones a la falta de una herramienta de diseño de sistemas difusos que sea aceptada como estándar dentro de esta área. Dicha herramienta estándar, de existir, contribuiría a difundir el uso de sistemas difusos en sectores más amplios de la industria.

En nuestra opinión, es complicado que la comunidad fuzzy se decante abiertamente por el uso de una herramienta en concreto, debido fundamentalmente a la variedad de aplicaciones existentes y a sus diferentes campos de aplicación. Por ejemplo, entre las aplicaciones de ámbito comercial podemos citar a las herramientas TILShell, de Togai Infralogic, FIDE, de Aptronix, o FuzzyTech, de Inform. También de ámbito comercial son los módulos de lógica fuzzy (*toolboxes*) de entornos como MatLab o Mathematica. En el ámbito académico se han desarrollado librerías de funciones difusas, como FFLL[12] o FuzzyJ [11], y numerosas herramientas dedicadas a campos como la minería de datos (p.e. KEEL [4]), la generación de datos imperfectos (p.e. Nip1.5 [3]), la generación de clasificadores difusos (p.e. NEFCLASS [9]), controladores difusos (p.e. NEFCON [10]), o el diseño de sistemas difusos (p.e. FisPro[5], GUAJE [1] o Xfuzzy [2]).

Consideramos que es muy difícil que una nueva herramienta consiga sustituir a las aplicaciones existentes, dada la diversidad de enfoques, campos de aplicación, lenguajes de programación y plataformas de ejecución que utilizan estas aplicaciones. Nuestro punto de vista es que resultaría mucho más adecuado avanzar hacia un lenguaje de representación de sistemas difusos que se aceptara como estándar dentro de la comunidad fuzzy. De esta forma, los desarrolladores de software sólo tendrían que adaptar sus aplicaciones al uso de dicho lenguaje de representación y se mantendría la diversidad de enfoques y plataformas existentes en la actualidad.

Siguiendo esta filosofía, este trabajo presenta un lenguaje de modelado de sistemas difusos que pretende ser lo suficientemente expresivo como para poder ser utilizado

como lenguaje de representación por la mayoría de herramientas existentes en el ámbito fuzzy.

2 EL LENGUAJE XFSML

XFSML es un lenguaje de modelado de sistemas difusos basado en XML. El nombre es el acrónimo de compleX/eXtensible Fuzzy System Modeling/Marckup Language. El punto de partida en el diseño del lenguaje ha sido la experiencia acumulada en el diseño del lenguaje XFL3 [6], que se definió como lenguaje de representación del entorno Xfuzzy 3.

El lenguaje está centrado en la descripción de la estructura de los sistemas difusos y no en su descripción funcional. Esto quiere decir que el lenguaje no incluye la definición de las funciones asociadas a los operadores difusos (como funciones de pertenencia, t-normas, t-conormas, o métodos de *defuzzificación*), sino la forma en que estos operadores se combinan para describir un determinado sistema. Estas funciones son referenciadas por su nombre y se asume que la descripción funcional debe encontrarse en alguna librería externa. Esto permite ampliar libremente las funciones disponibles para cada operador difuso (por ejemplo, definiendo nuevos tipos de funciones de pertenencia).

La característica fundamental del lenguaje es su alta capacidad expresiva. Esto quiere decir que se han incluido numerosas formas de representar sistemas difusos, de manera que el lenguaje permita modelar la mayor parte de los sistemas difusos utilizados por la comunidad científica y técnica.

El uso de XML permite además que XFSML sea independiente de herramientas, plataformas o lenguajes de programación concretos. De esta forma se permite que los desarrolladores adapten sus aplicaciones a este nuevo lenguaje de representación, sea cual sea el lenguaje de programación utilizado o el sistema operativo del equipo en el que se pretenda ejecutar.

2.1. ESTRUCTURA GLOBAL

La definición de un sistema difuso en XFSML consiste en un fichero XML formado por una única etiqueta (*fuzzysystem*). Esta etiqueta tiene el atributo "*name*" con el nombre del sistema difuso y contiene otras cuatro etiquetas que desarrollan las descripción del sistema:

- ***domains*** : describe los dominios (universos de discurso) de las diferentes variables que formen parte del sistema.
- ***partitions*** : describe las particiones difusas (es decir, los conceptos lingüísticos) asociados a cada dominio.
- ***relations*** : describe las relaciones difusas entre dominios. Esto permite expresar relaciones como X > Y en términos difusos.
- ***modules*** : describe los módulos (bases de conocimiento) que forman el sistema. La descripción global del sistema corresponde al módulo llamado "*system*".

A continuación se muestra un esquema del contenido de un sistema difuso descrito en XFSML.

```
<?xml version="1.0"?>
<fuzzysystem name="Example" xmlns:xsi= … >
        <domains>
        ...
        </domains>
        <partitions>
        ...
        </partitions>
        <relations>
        ...
        </relations>
        <modules>
        ...
        </modules>
</fuzzysystem>
```

Figura 1: Esquema general de un especificación XFSML.

2.2. DOMINIOS

Los dominios describen los universos de discurso de las variables que aparecen en el sistema. En el caso de los controladores difusos las variables suelen ser de tipo real. Sin embargo, en otro tipo de aplicaciones, como la minería de datos, es frecuente trabajar con variables de tipo entero o con valores enumerados. Por tanto, XFSML se ha diseñado para admitir tres tipos de dominio: reales, enteros y enumerados. Para definir estos dominios se utilizan tres etiquetas: *real*, *integer* y *enum*. Los dominios reales y enteros se describen por medio de los atributos *min* y *max*. Por su parte, los dominios enumerados se describen por medio de etiquetas *elem*.

```
<domains>
 <real name="dom1" min="-5.0" max="5.0" />
 <integer name="dom2" min="1" max="10" />
 <enum name="dom3">
   <elem value="Iris-setosa" />
   <elem value="Iris-virginica" />
   <elem value="Iris-versicolor" />
 </enum>
</domains>
```

Figura 2: Ejemplo de definiciones de dominios.

```
<partition name="part1" domain="dom1" >
  <freelabel name="SMALL" library="std"
             function="trapezoid">
    <param value="-10.0" />
    <param value="0.0" />
    <param value="40.0" />
    <param value="60.0" />
  </freelabel>
  <freelabel name="BIG" library="std"
             function="trapezoid">
    <param value="40.0" />
    <param value="60.0" />
    <param value="100.0" />
    <param value="110.0" />
  </freelabel>
</partition>
```

```
<partition name="part2" domain="dom2" >
  <family name="fam1" library="std"
          function="FAMtriangular">
    <param value="-10.0" />
    <param value="0.0" />
    <param value="10.0" />
  </family>

  <familylabel name="NEG" family="fam1"
               index="0" />
  <familylabel name="ZE" family="fam1"
               index="1" />
  <familylabel name="POS" family="fam1"
               index="2" />
</partition>
```

Figura 3: Ejemplo de definición de particiones usando funciones de pertenencia libres y familias de funciones

2.3. PARTICIONES

Las particiones describen los conceptos lingüísticos que se pueden asociar a los distintos dominios. Corresponden a la definición de las funciones de pertenencia asociadas a las variables en los sistemas difusos.

Al igual que el lenguaje XFL3 [7], las funciones de pertenencia definidas en XFSML pueden ser de dos tipos: funciones de pertenencia libres (que no tienen ningún tipo de restricción con respecto a otras funciones de pertenencia) y funciones de pertenencia de una familia. Las familias permiten definir conjuntos de funciones de pertenencia que presentan restricciones entre sí. Por ejemplo, funciones de pertenencia que compartan parámetros, funciones de pertenencia que deban ser simétricas, funciones de pertenencia que deben mantener un orden, etc.

La descripción funcional de estas funciones de pertenencia no está incluida en el lenguaje XFSML. Esto hace que el lenguaje no tenga restricciones en cuanto a la forma de las funciones de pertenencia que se pueden utilizar. Toda referencia a una función de pertenencia (o a una familia) se hace utilizando el nombre de la función y el nombre de la biblioteca en la que está descrita. Por otra parte, los parámetros de definición de estas funciones se describen como cadenas (*strings*), de manera que sea la función la responsable de verificar el valor de los parámetros utilizados.

La Figura 3 muestra un ejemplo de una partición formada por dos funciones libres de tipo trapezoide y de una partición formada por una familia de tres funciones triangulares que comparten los parámetros.

2.4. RELACIONES

Las relaciones permiten describir conceptos lingüísticos que afectan a dos dominios. Esto permite expresar enunciados del tipo "*el valor de X es aproximadamente igual al de Y*" o "*el valor de X es mucho mayor que el de Y*". Matemáticamente, las relaciones se definen como funciones de pertenencia sobre un universo de dos dimensiones.

Hasta el momento, las relaciones difusas no han sido demasiado utilizadas en el diseño de sistemas difusos, probablemente porque ninguna herramienta software ha dado soporte a este tipo de concepto. La inclusión de relaciones en XFSML abre la posibilidad de utilizar estos conceptos a la hora de representar el conocimiento.

La descripción de las relaciones es bastante similar a las particiones, aunque en este caso se definen sobre dos universos de discursos y no existen las familias. La Figura 4 muestra un ejemplo de descripción de relaciones.

```
<relation domain1="dom1" domain2="dom1">
  <label name="IgualQue" library="std"
         function="RELEqual">
     <param value="0.5" />
  </label>
  <label name="MuchoMayorQue"
       library="std" function="RELMuchGT">
     <param value="8.0" />
  </label>
</relation>
```

Figura 4: Ejemplo de definición de una relación.

```
<tree name="modulo1">
 <inputs>
   <input name="X1" partition="part2" /> <input name="X2" partition="part2" />
 </inputs>
 <outputs>  <output name="Y" partition="part1"/>  <outputs>
 <and library="std" function="product" />
 <defuz library="std" function="MaxLabel" />
 <root>
  <switch var="X1">
   <case label="NEG”> <node> <decision var="Y" label="SMALL" /> </node> </case>
   <case label="ZE">
    <node><switch var="X2">
      <case label="NEG”> <node> <decision var="Y" label="SMALL" /></node> </case>
      <case label="ZE"> <node> <decision var="Y" label="BIG" /> </node> </case>
      <case label="POS"> <node> <decision var="Y" label="SMALL" /> </node> </case>
    </switch></node>
   </case>
   <case label="POS"> <node> <decision var="Y" label="BIG" /> </node> </case>
  </switch>
 </root>
</tree>
```

Figura 5: Ejemplo de definición de un árbol de decisión difuso.

2.5. MÓDULOS

Los módulos permiten representar las bases de conocimiento que forman el sistema difuso. El lenguaje XFSML ofrece diversas formas de representación, como los conjuntos de reglas difusas, los árboles de decisión o los módulos jerárquicos. En general, un sistema difuso puede estar formado por uno o varios módulos. La descripción global del sistema es aquella asociada al módulo llamado "*system*".

El lenguaje XFSML incluye una amplia variedad de representaciones de conocimiento. Esto permite optar al diseñador por el uso de estructuras de diferente grado de complejidad. De cara a los desarrolladores de software, los diferentes tipos de módulos permiten desarrollar técnicas específicas (de identificación, de optimización, de simplificación, de generación de código, etc.) para cada tipo de estructura. Por ejemplo, el código de implementación de un árbol de decisión es diferente del de un conjunto de reglas.

Los siguientes sub-apartados describen los diferentes tipos de módulos incluidos en XFSML. La Figura 5 muestra un ejemplo de uno de estos tipos de módulos.

2.5.1. Árboles de decisión difusos

Los árboles de decisión son módulos de N entradas y una salida. El conocimiento se expresa en forma de árbol donde los nodos hoja son decisiones del tipo “Y es LB” y los nodos internos reflejan preguntas sobre el valor de una variable de entrada.

La definición de un árbol de decisión difuso está contenida en una etiqueta “*tree*”. El contenido de la etiqueta está formado por otras cinco etiquetas en el siguiente orden:

- ***inputs***: Descripción de las variables de entrada. Cada variable se describe mediante una etiqueta “*input*” que incluye el nombre de la variable y su partición.
- ***outputs***: Descripción de las variables de salida. Cada variable se describe mediante una etiqueta “*output*” que incluye el nombre de la variable y su partición.
- ***and***: Descripción del operador AND utilizado en el proceso de inferencia.
- ***defuz***: Descripción del método de *defuzzificación* utilizado en el proceso de inferencia.
- ***root***: Descripción del contenido del árbol a partir del nodo raíz. Los nodos terminales se representan con una etiqueta “*decisión*”. Los nodos internos se representan mediante una etiqueta “*switch*” con un lista de hijos formada por etiquetas “*case*” que a su vez pueden tener como hijo a un nodo terminal o a un nodo interno.

2.5.2. Matrices de reglas difusas

Las matrices permiten describir de manera compacta las bases de conocimiento formadas por dos variables de entrada, una variable de salida y reglas del tipo “IF X1 es LB1 AND X2 es LB2 THEN Y es LB3”. La matriz está

formada por todas las combinaciones posibles entre los valores de las etiquetas de la primera variable y los de la segunda.

La definición de una matriz de reglas difusas está contenida en una etiqueta "*array*", que a su vez contiene otras cinco etiquetas: *inputs*, *outputs*, *and*, *defuz* y *cells*. Las cuatro primeras son idénticas a las utilizadas en los árboles de decisión. La etiqueta "*cells"* describe el contenido de las celdas de la matriz. Cada fila corresponde a un valor (una función de pertenencia de la partición asociada) de la primera variable de entrada y está descrita mediante la etiqueta "*row*". El contenido de cada fila es un conjunto de celdas, correspondientes a cada valor (función de pertenencia de la partición asociada) de la segunda variable de entrada. Las celdas se describen con etiquetas "*cell*" con el atributo "*label*" que expresa el valor de la variable de salida para esa celda de la matriz.

2.5.3. Tablas de reglas difusas conjuntivas

Las tablas de reglas difusas definen módulos de N entradas y M salidas. Las filas corresponden a reglas del tipo "IF X1 es LB1 AND X2 es LB2 AND X3 es LB3 THEN Y1 es LB7 AND Y2 es LB8". Las tablas se definen mediante la etiqueta "*table*", que está formada por las etiquetas *inputs*, *outputs*, *and*, *defuz* y *rows*. Esta última etiqueta contiene la descripción del contenido de las reglas de la tabla. Cada fila se describe por una etiqueta "row", formada a su vez por etiquetas "*cell*" que describen las celdas de la tabla.

2.5.4. Conjuntos de reglas difusas conjuntivas

Los conjuntos de reglas difusas conjuntivas describen reglas similares a las expresadas en forma de tabla, con la diferencia de que los antecedentes de las reglas pueden no incluir valores de algunas variables de entrada. Se describen por medio de la etiqueta "*ruleset*" que contiene las etiquetas *inputs*, *outputs*, *and*, *defuz* y *rules*. La etiqueta *rules* está compuesta por una lista de etiquetas "*rule*", formada por un antecedente (etiqueta "*antecedent*") y un consecuente (etiqueta "*consequent*"). Tanto el antecedente como el consecuente está formado por un conjunto de términos (etiquetas "*term*") que expresan información de tipo "X es LB". Por ejemplo <term var="X1" label="Positive"/>.

2.5.5. Listas ordenadas de reglas difusas conjuntivas

Las listas ordenadas de reglas difusas conjuntivas definen módulos de N entradas y M salidas donde las reglas son del tipo "IF ... ELSEIF ... ELSEIF ... ELSE...". La definición de una lista ordenada de reglas difusas conjuntivas está contenida en una etiqueta "*rulelist"*. El contenido del módulo está formado por las etiquetas *inputs*, *outputs*, *and*, *defuz* y *rules*. La etiqueta *rules* está compuesta por una lista de etiquetas "*rule*", formada por un antecedente (etiqueta "*antecedent*") y un consecuente (etiqueta "*consequent*"). Tanto el antecedente como el consecuente está formado por un conjunto de términos (etiquetas "*term*"). Las reglas deben entenderse en el orden en el que están escritas (IF ... ELSEIF ... ELSEIF). La etiqueta *rules* puede contener opcionalmente una última etiqueta "*else*" que está formada sólo por el consecuente (no tiene la etiqueta "*antecedent*").

2.5.6. Conjuntos de reglas difusas complejas

Los conjuntos de reglas complejas se diferencian de los conjuntos de reglas conjuntivas en la forma de los antecedentes de las reglas. El antecedente de las reglas conjuntivas se interpreta como la operación AND sobre la lista de términos que forma el antecedente. Por su parte, las reglas complejas tienen antecedentes formados por combinaciones de operadores lógicos (AND, OR y NOT), por modificadores lingüísticos (STRONGLY, MOREORLESS), términos complejos (por ejemplo, "X > LB" o "X <LB") y por relaciones entre variables (por ejemplo, "MayorQue(X,Y)".

Los conjuntos de reglas complejas se definen por medio de la etiqueta "*complexruleset*", que a su vez contiene las etiquetas *inputs*, *outputs*, *operators*, *defuz* y *rules*. La etiqueta *operators* contiene la definición de las funciones asociadas a cada modificador lingüístico utilizado en las reglas. La etiqueta *rules* contiene la lista de reglas (etiquetas *rule*), cuyo antecedente está formado por una expresión compleja.

2.5.7. Listas ordenadas de reglas difusas complejas

Las listas ordenadas de reglas complejas permiten definir reglas del tipo "IF ... ELSEIF ... ELSEIF ... ELSE...", donde el antecedente de las reglas puede ser una expresión compleja como las explicadas en el apartado anterior.

Para definir una lista ordenada de reglas difusas complejas se utiliza la etiqueta "*complexrulelist*", que está formada por las etiquetas *inputs*, *outputs*, *operators*, *defuz* y *rules*. La etiqueta *rules* puede contener opcionalmente una última etiqueta "*else*" que está formada sólo por el consecuente.

2.5.8. Sistemas difusos jerárquicos

Los módulos jerárquicos son módulos formados por llamadas a otros módulos Estos módulos se definen por medio de la etiqueta "*hierarchy*". El contenido de la etiqueta está formado por otras tres etiquetas en el siguiente orden: *inputs*, *outputs* y *calls*. Las etiquetas *inputs* y *outputs* se diferencian de las de otros módulos en que las variables se definen con su nombre y su dominio en vez de su nombre y partición. La etiqueta *calls* está formada por una lista de etiquetas "*call"* que describen las llamadas que forman la estructura jerárquica. Cada etiqueta *call* tiene un atributo "*module*" (que referencia al módulo al que se llama), una etiqueta *inputs* (que indica las variables

de entrada del módulo) y una etiqueta *outputs* (que indica cuales son las variables de salida del módulo).

2.5.9. Módulos no difusos

Los módulos no difusos son módulos dedicados a operaciones aritméticas entre datos no difusos (suma, multiplicación,...). Se suelen utilizar como módulos dentro de una estructura jerárquica. Para definirlos se utiliza la etiqueta "*crisp*", que incluye a su vez tres etiquetas: *inputs*, *outputs* y *description*. Las etiquetas *inputs* y *outputs* definen las variables del módulo con su nombre y dominio. La etiqueta *description* contiene la referencia a la librería, función y parámetros que definen el módulo no difuso.

3 CONCLUSIONES

Los sistemas difusos tienen un gran desarrollo a nivel académico pero su uso en otros ámbitos es aún escaso. En nuestra opinión, la adopción por parte de la comunidad *fuzzy* de un lenguaje estándar de modelado de sistemas difusos contribuiría a difundir el uso de estos sistemas y facilitaría su implantación fuera del ámbito académico.

La definición de este lenguaje estándar debería ser promovida y aprobada por el Comité Técnico de Sistemas Difusos (*Fuzzy Systems TC*) de la IEEE Computational Inteligence Society, como órgano más influyente dentro de la comunidad *fuzzy*. Dicho Comité Técnico cuenta actualmente con un equipo de trabajo (*Task Force on Fuzzy Systems Software*) que podría constituir un marco adecuado para la definición de dicho estándar.

El lenguaje XFSML, presentado en este trabajo, puede ser un buen punto de partida para construir dicho estándar. Este lenguaje está basado en XML y se apoya en la experiencia acumulada en el desarrollo del entorno Xfuzzy 3 y de su lenguaje de especificación XFL3. Además, en este momento es fácil introducir modificaciones ya que se encuentra en los primeros pasos de su definición.

Agradecimientos

Este trabajo ha sido financiado parcialmente por los Proyectos de Investigación TEC2011-24319, TEC2008-04920 y TIN2008-06681-C06-06 del Ministerio de Ciencia y Tecnología y el Proyecto de Excelencia P08-TIC-03674 de la Junta de Andalucía.

Referencias

[1] J. M. Alonso, L. Magdalena: GUAJE - A Java environment for generating understandable and accurate models. En: *Actas del XV Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF-2010)*, Huelva, pp. 399-404, 2010.

[2] I. Baturone, F. J. Moreno-Velo, S. Sánchez-Solano, A. Barriga, P. Brox, A. Gersnoviez, M. Brox: Using Xfuzzy Environment for the Whole Design of Fuzzy Systems. En: *Proc. IEEE International Conference on Fuzzy Systems*, Londres, 2007.

[3] J. M. Cadenas, J.V. Carrillo, M. C. Garrido, E. Muñoz: Nip1.5 : una herramienta software para la generación de conjuntos de datos con imperfección para minería de datos. En: *Actas del XV Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF-2010)*, Huelva, pp. 411-416, 2010.

[4] J. Derrac, A. Fernández, J. Luengo, S. García, L. Sánchez, J. Alcalá, F. Herrera: KEEL: una herramienta software para el análisis de sistemas difusos evolutivos. En: *Actas del XV Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF-2010)*, Huelva, pp. 417-422, 2010.

[5] S. Guillaume, B. Charnomordic: Learning interpretable Fuzzy Inference Systems with FisPro. En: *Information Sciences* 181(20), pp. 4409-4427, 2011.

[6] F. J. Moreno-Velo, S. Sánchez-Solano, A. Barriga, I. Baturone, D.R. López: XFL3: a new fuzzy system specification language, En: *Advanced in Scientific Computing, Computational Intelligence and Applications*, World Scientific and Engineering Society Press, Singapur, pp. 361-366, 2001.

[7] F. J. Moreno-Velo, I. Baturone, S. Sánchez-Solano , A. Barriga: The Parametric Definition of Membership Functions in XFL3. En: *Proc. IEEE International Conference on Fuzzy Systems*, Budapest, 2004.

[8] D. Nauck: GNU Fuzzy. En*: Proc. IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE 2007)*, Londres, pp. 1-6, 2007.

[9] D. Nauck, R. Kruse: NEFCLASS-J - a Java-based soft computing tool. En*: Intelligent Systems and Soft Computing: Prospects, Tools and Applications, ser. Lecture Notes in Artificial Intelligence*, B. Azvine, N. Azarmi, and D. Nauck, Eds. Berlin: Springer-Verlag, no. 1804, pp. 143-164, 2000.

[10] A. Niirnberger, D. Nauck, R. Kruse: Neuro-fuzzy control based on the NEFCON-model: Recent developments. En: *Soft Computing*, vol. 2, no. 4, pp. 168-182, 1999.

[11] R. A. Orchard: Fuzzy reasoning in Jess: The Fuzzy J Toolkit and Fuzzy Jess. En: *Proceedings of the Third International Conference on Enterprise Information Systems (ICEIS 2001)*, Setubal (Portugal), pp. 533-542, 2001.

[12] M. Zarozinski: An Open Source Fuzzy Logic Library. En: *AI Game Programming Wisdom*, Ed. Charles River Media, pp. 90-103, 2002.

Impacto de las relaciones jerárquicas inter/intra-recursos en la similitud semántica entre conceptos del dominio biomédico

Israel Alonso Martínez [1], David Contreras Bárcena [1], Francisco P. Romero Chicharro [2]

[1] *ICAI - Universidad Pontificia Comillas*, {ialonso,davidcb}@dsi.icai.upcomillas.es

[2] *Escuela de Ingeniería Industrial de Toledo - Universidad Castilla-La Mancha*, franciscop.romero@uclm.es

Resumen

En el presente trabajo se realiza un estudio y evaluación de las principales métricas para el cálculo de la similitud semántica entre conceptos dentro del ámbito biomédico, tomándose como referencia un conjunto de prueba definido en trabajos previos y utilizando el Metatesauro UMLS, como recurso de conocimiento para la caracterización de la similitud en el dominio biomédico. Como aportación novedosa de este trabajo, se analiza la importancia de las relaciones inter/intra-recursos de jerarquía directa entre conceptos y se unifican las métricas utilizadas para la evaluación de los resultados obtenidos en trabajos previos. Finalmente, para evaluar el rendimiento de las métricas de similitud semántica con corpus no controlados, se ha realizado un experimento de agrupamiento sobre una colección de notas radiológicas de libre disposición.

Palabras Clave: Similitud semántica, Ontologías, Historias clínicas electrónicas, UMLS.

1 INTRODUCCIÓN

El tratamiento de la información en el ámbito biomédico presenta una gran dificultad debido a su complejidad y ambigüedad. Es por ello que surgen diferentes recursos como modelo de representación, acceso y tratamiento del conocimiento en dicho dominio, como son MeSH, SNOMED-CT, ICD-9-CM, etc. Todos ellos, a su vez, están recogidos y relacionados entre sí en el Metatesauro UMLS (Unified Medical Language System).

Estos recursos permiten la mejora de los sistemas de recuperación de información basados en documentos clínicos, mediante la extracción y aplicación de conceptos similares o sinónimos en la reformulación de consultas. Así, por ejemplo, las consultas basadas en historias clínicas electrónicas (EHR) para la recuperación de documentos afectados y asociados a una enfermedad en controles epidemiológicos, conlleva la utilización de una gran variedad de términos que pueden estar estrechamente relacionados o ser similares con la enfermedad original de búsqueda. Por tanto, la dificultad principal estriba en definir cómo de relacionados o similares pueden ser dos conceptos entre sí, especialmente en un dominio específico y de gran complejidad como el biomédico.

El resto del artículo se organiza de la siguiente manera. En la sección 2 se presenta una revisión del estado del arte entorno a las métricas de similitud semántica. En la sección 3 se definen los objetivos de este trabajo y como se han llevado a cabo. En la sección 4 se recogen los resultados obtenidos en la evaluación de diferentes métricas de similitud semántica, obteniendo la métrica y recursos que mejor resultado aporta en la aproximación semántica entre conceptos. En la sección 5 se realiza un experimento que permita confirmar la evaluación de las distintas métricas para un conjunto de datos abierto basado en notas clínicas. Finalmente, en la sección 6 se comentan brevemente las conclusiones y trabajos futuros.

2 ESTADO DEL ARTE

En el entorno de recuperación de información, es de gran interés el estudio y definición de medidas que permitan obtener de una forma automatizada la similitud semántica entre diferentes conceptos. Ciertos trabajos lo han reflejado en dominios genéricos basados en ontologías de la lengua inglesa [3], aplicandolo en el procesamiento del lenguaje natural para la recuperación de información y/o documentos [5], [12], [15]. Dichos enfoques pierden su eficacia en dominios específicos como el biomédico [13], haciendo necesario el estudio de nuevas métricas apoyadas en recursos de dicho ámbito de conocimiento.

Algunos nuevos trabajos, combinan el conocimiento aportado por las taxonomías del dominio biomédico, junto con el tratamiento de textos médicos [15] o de ocurrencias en corpus establecidos [12]. Estas últimas aproximaciones, tienen limitaciones debidas a la falta de generalidad y cobertura de los corpus en un dominio tan extenso como el tratado aquí. Por tanto, este trabajo se centra en las diferentes métricas de similitud semántica basadas en el camino entre conceptos dentro de ontologías o taxonomías del dominio biomédico. A continuación se detallan los principales trabajos en esta área.

Ciertos estudios, se basan en la idea de la similitud semántica como el camino más corto en las interrelaciones entre dos conceptos (*p*). Así, Rada [14] define una métrica donde se calcula el camino más corto entre dos términos a partir de sus relaciones jerárquicas indirectas (RB/RN), mientras Caviedes y Cimino [2] lo aplican a las relaciones jerárquicas directas (PAR/CHD). Estas métricas utilizan relaciones extraídas de recursos del dominio biomédico como MeSH, ICD-9-CM y SNOMED-CT. Una variación normalizada de la métrica anterior, denominada "Path Length" (*path*), es definida por Pedersen et al. [13] a partir de relaciones "is-a" en SNOMED-CT. Estas métricas, omiten ciertos elementos de posible interés como la profundidad o tamaño de la taxonomía y la localización de los términos dentro de ella.

Así, Leacock y Chodorow (*lch*) [6] reflejan estas características obviadas considerando el camino más corto (*p*) entre dos conceptos (c_1, c_2), escalándolo de forma constante con respecto a la profundidad (*depth*) de la taxonomía. De esta manera, cuanto mayor es la profundidad de la taxonomía (más compleja y completa), mayor es el valor relativo de la similitud semántica entre dos conceptos. Por otro lado, Wu y Palmer (*wup*) [18], tratan de diferenciar aquellos términos que estando al mismo nivel de profundidad dentro de la taxonomía, tienen un grado de similitud diferente. Para ello obtienen la similitud entre dos términos a partir del concepto común más próximo que los relaciona (LCS - Lowest Common Subsumer), asociando la profundidad de los conceptos en la jerarquía con respecto al LCS.

Sin embargo, se observa que las métricas anteriores no tienen en cuenta en que parte de la taxonomía se encuentra el camino más corto (*p*) entre los conceptos (c_1, c_2). Esto puede afectar al grado de similitud semántica entre conceptos, ya que cuanto más próximo esté a la raíz de la taxonomía implica menor similitud al ser conceptos muy genéricos. Para resolver este problema, Nguyen y Al-Mubaid (*nam*) [11] definen una métrica que incorpora la importancia de la localización de (*p*) entre (c_1, c_2), dentro de la profundidad de la taxonomía. Esta métrica obtiene la mayor similitud semántica entre conceptos, cuando el LCS se sitúa en la parte más próxima a las "hojas" de la taxonomía y más corto es el camino entre los conceptos.

Otra aproximación se basa en la métrica Concept Match [8], restringida al conjunto de superconceptos de un término en una jerarquía de tipo "is-a". Dando lugar a la adaptación denominada por Sánchez et al. [16], como SuperConcept based Distance (SCD). En ella, se tiene en cuenta el conocimiento taxonómico aportado por los antecesores comunes y no comunes de los dos conceptos comparados para la evaluación de la similitud entre ellos.

Por último, Pedersen [13] y McInnes [10] comparan en sus trabajos la valoración de las métricas más importantes (*path, lch, nam y wup*) empleando los recursos MeSH y SNOMED-CT.

3 PROPUESTA DE ESTUDIO

A continuación se describen los objetivos del trabajo, las herramientas utilizadas y los métodos llevados a cabo para alcanzar los resultados.

3.1. OBJETIVOS

Este trabajo se centra en el estudio de las métricas más importantes de similitud semántica entre conceptos del dominio biomédico. El objetivo definido es evaluar dichas métricas para el conjunto de prueba expuesto en el apartado 3.3, con la actualización del Metatesauro UMLS versión 2010AB. Trabajos previos utilizaron la versión UMLS 2008AB para los recursos SNOMED-CT y MeSH sobre relaciones PAR/CHD [10], [13].

Además, se evaluará cómo afecta a los resultados la utilización (independiente o combinada) de diferentes recursos del dominio y la incorporación de relaciones jerárquicas indirectas (RB/RN), así como todas las relaciones existentes en el Metatesauro UMLS.

Por último, la falta de unificación en los coeficientes de correlación utilizados, mostrará grandes diferencias entre los resultados de trabajos previos [13], [16]. Por ello, en este trabajo se justifica el uso del coeficiente de correlación de Pearson como medida estándar de aplicación.

3.2. METATESAURO UMLS

UMLS (Unified Medical Language System) recoge todos los recursos de conocimiento del dominio biomédico para su representación, acceso y tratamiento de manera individual o en conjunto [9]. Su objetivo principal es definir los diferentes conceptos (más de 1 millón, almacenados en más de 100 recursos) y las relaciones entre ellos. Se agrupan en diferentes perspectivas del ámbito biomédico. Por ejemplo, la información científica (MeSH, CRISP), terminología clínica (SNOMED-CT), administrativa (ICD-9-CM, CPT-4), o de intercambio de datos (HL7, LOINC), etc.

Las relaciones entre conceptos pueden ser intra-recursos o inter-recursos, pudiéndose dividir en jerárquicas y no jerárquicas [1]. En este trabajo principalmente se han utilizado las relaciones jerárquicas, ya que son las que expresan similitud entre conceptos.

Por último, para la validación de las métricas de similitud semántica utilizadas en este trabajo (*path, wup, lch y nam*) se ha recurrido a un conjunto de módulos Perl de libre distribución que los implementan. Estos módulos están disponibles en CPAN[1] (The Comprehensive Perl Archive Network) y serán evaluados con la utilización de los recursos del dominio biomédico contenidos en la versión 2010AB de UMLS.

[1] http://www.cpan.org

3.3. MÉTODO DE TRABAJO

Las métricas expuestas anteriormente han sido evaluadas con un conjunto de prueba definido por 9 codificadores médicos (*Coders*) y 3 facultativos (*Physicians*) de la Clínica Mayo [13]. Éste lo forman 29 pares de conceptos, al que los expertos asignaron una puntuación (en el rango [4.0,0.0]) que representa la similitud semántica entre ellos. Definiendo si dos conceptos eran prácticamente sinónimos [4.0,3.0], relacionados [3.0,2.0], ligeramente relacionados [2.0,1.0] y no relacionados [1.0,0.0]. El coeficiente de correlación de Pearson para la valoración de similitud otorgado por los expertos es de 0.85.

Otras fuentes para el cálculo de similitud entre conceptos como ConceptNet [7] (de propósito general), no son válidas dentro del contexto biomédico ya que la mayoría de los términos no están incluidos.

4 RESULTADOS

A continuación se describen los resultados obtenidos en la evaluación de las diferentes métricas y el efecto de los recursos aplicados.

4.1. MÉTRICAS SOBRE RECURSOS MESH Y SNOMED-CT

En primer lugar se evalúan las distintas métricas sobre el recurso MeSH con relaciones PAR/CHD, con el objeto de contrastar nuestros resultados con los obtenidos por McInnes y Pedersen [10], [13]. Para la nueva versión MeSH (2010AB), los resultados reflejaron que 11 pares de conceptos del conjunto de prueba no devolvieron ningún resultado, optando por ello no realizar el cálculo de correlación para este recurso al no ser comparables.

Por este motivo, se han evaluado las métricas *path* y *lch*, únicas métricas calculadas en trabajos previos [10], sobre el recurso SNOMED-CT para las relaciones de tipo PAR/CHD. En este caso, se obtienen los resultados para cada uno de los 29 pares de conceptos, lo cual permite su comparación con ambas versiones de SNOMED-CT (Tabla 1). Hay que destacar que los resultados obtenidos en trabajos citados anteriormente, siempre se calculan con el coeficiente de correlación de Spearman.

Como muestra la Tabla 1, las correlaciones obtenidas para las distintas versiones de SNOMED-CT varían significativamente (8-10.5%).

Tabla 1. Valores de correlación (Spearman-mínimos) de similitud semántica con relaciones PAR/CHD para SNOMED-CT 2008AB y 2010AB.

	SNOMED-CT 2008AB		**SNOMED-CT 2010AB**	
	Physicians	**Coders**	**Physicians**	**Coders**
path	0.3500	0.5000	0.3134	0.4596
lch	0.3500	0.5000	0.3134	0.4593

4.2. IMPACTO DEL TIPO DE CORRELACIÓN

Los resultados anteriores nos han conducido al análisis de las variables empleadas en los trabajos de McInnes y Pedersen [10], [13]. Estos utilizan el coeficiente de correlación de Spearman, el cual demuestra cómo de ajustado es el orden de los resultados obtenidos por las métricas, en relación al orden de valores definidos por los expertos, sin tener en cuenta la proximidad de los propios resultados con respecto a los valores definidos por los expertos. Este hecho resulta importante para la utilización de estas métricas en sistemas de recuperación de información. Por ello, en este trabajo se plantea la utilización del coeficiente de correlación lineal de valores de Pearson, como criterio de evaluación de la proximidad entre conceptos. Para facilitar la comparación con otros trabajos, se utilizaran ambas correlaciones.

Tabla 2. Valores de correlación (Pearson) de similitud semántica con relaciones PAR/CHD para SNOMED-CT 2008AB y 2010AB.

	SNOMED-CT 2008AB		**SNOMED-CT 2010AB**	
	Physicians	**Coders**	**Physicians**	**Coders**
path	0.5400	0.7100	0.5451	0.7170
lch	0.3300	0.4500	0.3348	0.4566

Los cálculos realizados con el coeficiente de correlación de Pearson, demuestran variaciones poco significativas (1-1.5%) en los resultados obtenidos entre las versiones de 2008AB y 2010AB (Tabla 2). Esto demuestra, para Spearman, que pequeñas diferencias en los resultados entre pares de conceptos afectan en el rango de ordenamiento, reflejando mayores variaciones en la correlación.

En la Tabla 3, se muestran las correlaciones para las métricas *wup* y *nam* (no evaluadas en [10] y [13]), con el objetivo de confirmar la importancia de la localización de los conceptos en la taxonomía y su posición con respecto al LCS. Estos resultados varían con Tabla 1, ya que el criterio para el cálculo de la correlación de Spearman se basa en valores mínimos para rangos de repetición, frente a los valores medios utilizados en la formulación del modelo [17]. Este criterio es el utilizado en este trabajo, véase Tabla 3.

Tabla 3. Valores de correlación (Spearman-medias y Pearson) de similitud semántica con relaciones PAR/CHD para SNOMED-CT 2010AB.

		path	**lch**	**wup**	**nam**
Spearman	**Physicians**	0.2744	0.2744	0.3377	**0.4063**
	Coders	0.4160	0.4156	0.4190	**0.5578**
Pearson	**Physicians**	**0.5451**	0.3348	0.3372	0.4301
	Coders	**0.7170**	0.4566	0.3840	0.4456

Al igual que en trabajos anteriores [10], [13] la métrica *nam* obtiene el valor de correlación (Spearman-medias) más alto. Además, los resultados obtenidos se ajustan más

a los criterios de similitud semántica definidos por codificadores médicos frente a facultativos. Por ello, en las figuras de este trabajo se comparan los resultados obtenidos con los valores establecidos por dichos codificadores.

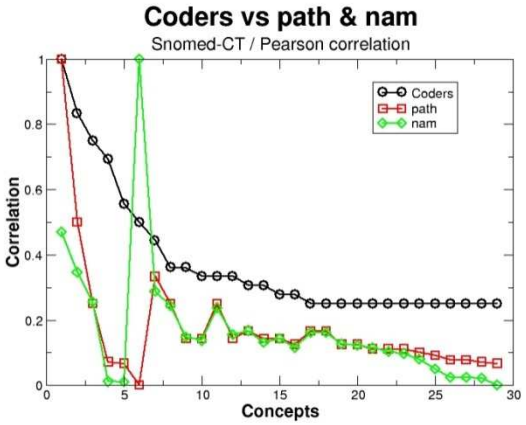


Figura 1. Correlación (Pearson) de similitud semántica para métricas *path* y *nam* (SNOMED-CT 2010AB).

Para realizar un mejor seguimiento del trabajo propuesto en este artículo, se muestran los resultados según los coeficientes de correlación de Pearson y Spearman (basado en medias). La correlación Pearson no confirma, como otros autores sugieren [13], [18] y [11], la relevancia de la localización y profundidad de los conceptos en la taxonomía para el cálculo de la similitud semántica entre términos (Figura 1).

4.3. USO DE RELACIONES PAR/CHD Y RB/RN PARA EL CÁLCULO DE CORRELACIONES

Los cálculos analizados en el capítulo 4.1, se obtuvieron con relaciones del tipo PAR/CHD. Por ello, en este apartado se evalúa el impacto de incluir las relaciones jerárquicas de tipo RB/RN a estos cálculos. En la Tabla 4 se muestran los nuevos valores de correlación obtenidos para la métrica *path* y *wup* (se reducen las métricas calculadas por el alto coste computacional requerido).

Como se observa, los resultados obtenidos para la métrica *path* ofrecen mejores valores de correlación que con los tipos de relación exclusiva PAR/CHD.

Tabla 4. Valores de correlación (Spearman-medias y Pearson) de similitud semántica con relaciones PAR/CHD y RB/RN para SNOMED-CT 2010AB.

		path	**wup**
Spearman	**Physicians**	0.3393	0.2651
	Coders	**0.4907**	**0.3637**
Pearson	**Physicians**	0.5730	0.2734
	Coders	**0.7438**	**0.3049**

Esto es debido a la reducción en los caminos más cortos entre ciertos conceptos. Así por ejemplo, para los términos "Carpal tunnel síndrome" C0007286 y "Osteoarthritis" C0029408, los expertos codificadores asignaron un valor 1.111 de similitud semántica entre ellos (0.2778 normalizado de 0 a 1). Los resultados de la métrica *path*, mejoraron pasando de 0.1250 para PAR/CHD a 0.2000 para PAR/CHD y RB/RN.

Para *wup* no se mejoran los resultados debido a la naturaleza de la métrica. Ésta prioriza la profundidad de los conceptos c_1 y c_2 hasta la raíz de la taxonomía, en relación con la profundidad del LCS que los relaciona. Sin dar importancia y valor a la distancia mínima entre conceptos.

4.4. EVALUACIÓN DE MÉTRICAS CON EL METATESAURO COMPLETO UMLS

A continuación, se evalúa la incorporación de todos los recursos del dominio biomédico contenidos en el Metatesauro UMLS 2010AB. Se realizan los cálculos para las relaciones jerárquicas directas (PAR/CHD) e indirectas (RB/RN), así como para las relaciones no jerárquicas (PAR, CHD, RB, RN, SIB, RO, RL, SY, RQ, AQ y QB).

En la Tabla 5, se muestra los resultados para todos los recursos con la métrica *path*. (Sólo se realizan los cálculos para este caso, debido al alto coste computacional).

Tabla 5. Correlaciones similitud semántica para métrica *path*, con relaciones PAR/CHD, PAR/CHD+RB/RN y TODAS las contenidas en recursos de UMLS 2010AB.

		PAR/CHD	**PAR/CHD + RB/RN**	**TODAS**
Spearman	**Physicians**	0.6382	0.5761	**0.4788**
	Coders	**0.6422**	**0.6495**	0.4338
Pearson	**Physicians**	0.7059	0.6740	0.6168
	Coders	**0.7982**	**0.8012**	**0.7046**

Para las relaciones PAR/CHD y PAR/CHD + RB/RN, los resultados muestran una mejora significativa en la correlación de los valores recogidos en la definición de la similitud semántica entre los 29 conceptos iniciales (Figura 2).

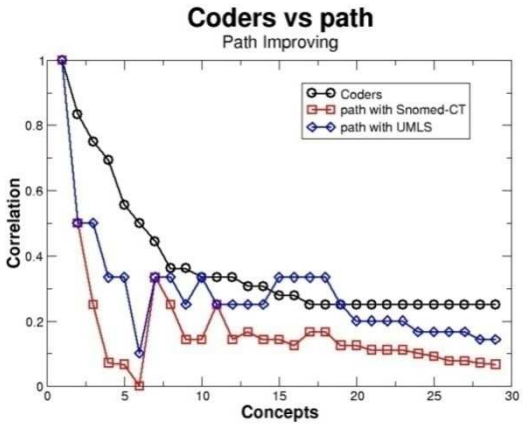


Figura. 2. Mejora de correlación para métrica *path* (UMLS 2010AB) con relaciones PAR/CHD frente a SNOMED-CT.

En ambos casos, se llega a obtener una correlación muy alta (alrededor de 0.80) con respecto a los criterios de similitud semántica definidos por los codificadores. Es conveniente recordar que la correlación entre los valores establecidos por codificadores y facultativos es de 0.85.

Como muestra la Tabla 5, el uso de todas las relaciones (jerárquicas y no jerárquicas) en UMLS 2010AB degrada notablemente los resultados respecto a los casos anteriores. La aparición de nuevas relaciones no jerárquicas entre conceptos distorsiona la similitud entre ellos.

5 EXPERIMENTO

Con el objeto de evaluar las distintas métricas para un conjunto de datos no cerrado, se realiza un experimento sobre una colección de documentos, basados en notas clínicas utilizadas en el Computational Medicine Center 2007[2]. Cada nota está asociada a un texto corto y a uno o varios códigos ICD-9 que lo describe. Para el experimento, se ha utilizado un conjunto de 894 documentos distribuidos en grupos según la codificación mayoritaria que los identifica. Cada nota clínica se ha representado utilizando un vector de términos en los que el peso de cada uno de ellos viene dado por su frecuencia de aparición. Posteriormente se construye una matriz de similitud semántica entre términos a partir de los valores de similitud extraídos de SNOMED-CT utilizando las relaciones PAR/CHD. Finalmente se utiliza esta matriz para transformar los vectores basados en frecuencia en vectores reajustados semánticamente según lo explicado en [4]. Con el objeto de medir la calidad de la agrupación de los documentos, se han utilizado dos indicadores principales: la Cohesión intra-grupos y la Cohesión corregida.

La Cohesión intra-grupo (Eq. 1) se define como la media de la similitud entre los elementos de un mismo grupo, que se puede calcular mediante la función coseno aplicada a cada elemento y el centroide de cada grupo,

$$coh_intra(C) = \frac{\sum_{d \in C} \cos(d,c)}{|C|} \quad (1)$$

siendo *d* un documento perteneciente a un cluster C, *c* el centroide (elemento central) del cluster *C* y $|C|$ su número de elementos. La cohesión de toda la estructura será la media ponderada de las cohesiones de cada uno de los grupos,

$$coh_estruc(C^E) = \sum \frac{|C|}{|C^E|} * coh_intra(C) \quad (2)$$

siendo C^E la estructura obtenida y $|C^E|$ el número de elementos de toda la estructura. Los valores de cohesión intra-grupos obtenidos, son altos y muy próximos para cada una de las distintas métricas. Sin embargo, este valor no es suficiente para determinar la calidad de un agrupamiento, ya que una mayor cohesión intra-grupos debe ir acompañada de una menor similitud inter-grupos. De esta forma, se ha corregido el valor de la cohesión intra-grupo restándole la similitud media entre todos los elementos de la colección:

$$coh_intra'(C') = coh_intra(C) - \frac{\sum_{d,d' \in C} \cos(d,d')}{|C^E|} \quad (3)$$

En este caso, se busca maximizar el valor C' a partir de una cohesión intra-grupos alta y una similitud media de toda la colección baja. Un valor más alto de cohesión corregida implicará una calidad más alta en la agrupación. Como muestra la Tabla 6, la métrica *path* es la que obtiene mejores resultados (un 20% superior al resto). Este hecho demuestra como la cohesión intra-grupo que ofrece la métrica *path*, está basada únicamente en la similitud de los contenidos que identifican cada uno de los grupos y no en la similitud del conjunto de documentos.

Tabla 6. Cohesión Intra-grupos

	path	**wup**	**lch**	**nam**
coh_intra' (C')	**0.51**	0.42	0.38	0.33

6 CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se ha presentado un amplio estudio y comparativa de las diferentes métricas propuestas para el cálculo de la similitud semántica entre conceptos, variando tanto los recursos del dominio como las relaciones jerárquicas definidas en ellos. Se ha logrado demostrar, como la métrica *path* es la que mejor aproxima los valores de similitud semántica entre conceptos. Obteniendo los mejores resultados cuanto mayor es el número de recursos que aportan información sobre el dominio, pero únicamente para las relaciones jerárquicas inter/intra-recursos de tipo PAR/CHD y RB/RN. También se ha demostrado que las relaciones no jerárquicas no aportan información de similitud, además UMLS se señala como el recurso que ofrece el mayor conocimiento del dominio, donde las nuevas relaciones definidas en sus actualizaciones afectan a los resultados de las métricas. En nuestro estudio y aplicación, los valores arrojados por las métricas son claves, ya que la representación de notas clínicas en vectores de conceptos depende de la corrección del valor de similitud obtenido. Por ello, se plantea la correlación de Pearson como la medida adecuada para reflejar la relevancia de las métricas en la aproximación de los resultados con respecto a los valores de similitud semántica expresada por los expertos y no con respecto al orden (como define la correlación de Spearman).

Por último, se han validado las métricas tanto con un conjunto cerrado de conceptos (al igual que el resto de trabajos), como con un experimento basado en un corpus

[2] http//computationalmedicine.org/catalog/downloads/2007-challenge-trining-data

de notas clínicas. Con este experimento, se ha comprobado que la métrica *path* es la que ofrece mejor formación en el agrupamiento entre grupos de documentos.

Como trabajos futuros se plantea la definición de un nuevo método para la medición de la similitud semántica entre conceptos que trate la complejidad de las diferentes relaciones inter e intra recursos contenidos en el Metatesauro UMLS, con la intención de poder optimizar los resultados para aquellos casos en que la métrica path no es tan eficiente y si los son aquellas que consideran la localización del camino más corto entre dos conceptos dentro de la ontología formada por UMLS. Para ello se pretende integrar las características especificas de distintas métricas mediante operadores de agregación borrosa (OWA) que permitan incluir en una única métrica, las propiedades de aquellas que mejor se adapten a las relaciones semánticas definidas en el Metatesauro para cada par de conceptos. Para ello será necesaria la ponderación de las diferentes relaciones jerárquicas (directas e indirectas) que permitan establecer mediante reglas borrosas el peso específico de cada relación semántica en función de los diferentes características semánticas asociadas.

Por otro lado, se plantea el uso de este método en la mejora de la recuperación de información en repositorios de historiales clínicos electrónicos, con la aplicación de procesos de expansión semántica de la consulta. Con el fin de poder evaluar estas alternativas se plantea la utilización de la colección de prueba, proporcionada en el TREC 2011 Medical Track[3], que permita de esta manera, la evaluación comparativa de las propuestas con respecto a otros métodos existentes.

Referencias

[1] Burgun A, Bodenreider O. Methods for exploring the semantics of the relationships between co-occurring UMLS concepts. S. H. Technol Inform. 2001:171-5.

[2] Caviedes JE, Cimino JJ. Towards the development of a conceptual distance metric for the UMLS. J Biomed Inform. 2004;37(2):77-85.

[3] Fellbaum C. WordNet: An electronic lexical database. The MIT press; 1998.

[4] Garcés PJ, Olivas JA, Romero FP. Concept matching IR systems vs word matching information retrieval systems: Considering fuzzy interrelations for indexing web pages. JASIS Technol. 2006;57(4):564-76.

[5] Jiang JJ, Conrath DW. Semantic similarity based on corpus statistics and lexical taxonomy. Tapei. Taiwan.; 1997. p. 19-33.

[6] Leacock C, Chodorow M. Combining local context and WordNet similarity for word sense identification. 1998;49(2):265-83.

[7] Liu H, Singh P. In: Commonsense reasoning in and over natural language. Knowledge-based intelligent information and engineering systems; Springer; 2004. p. 293-306.

[8] Maedche A, Zacharias V. Clustering ontology-based metadata in the semantic web. Principles of Data Mining and Knowledge Discovery. 2002:383-408.

[9] McCray AT, Aronson AR, Browne AC, Rindflesch TC, Razi A, Srinivasan S. UMLS knowledge for biomedical language processing. Bull Med Libr Assoc. 1993;81(2):184.

[10] McInnes BT, Pedersen T, Pakhomov SVS. In: UMLS-interface and UMLS-similarity: Open source software for measuring paths and semantic similarity. AMIA annual symposium proceedings; American Medical Informatics Association; 2009. p. 431-35.

[11] Nguyen HA, Al-Mubaid H. In: New ontology-based semantic similarity measure for the biomedical domain. May 10-12; Atlanta, USA. IEEE; 2006. p. 623-8.

[12] Patwardhan S, Pedersen T. Using WordNet-based context vectors to estimate the semantic relatedness of concepts. Making Sense of Sense: Bringing Psycholinguistics and Computational Linguistics Together. 2006:1.

[13] Pedersen T, Pakhomov SVS, Patwardhan S, Chute CG. Measures of semantic similarity and relatedness in the biomedical domain. J.B.I. 2007;40(3):288-99.

[14] Rada R, Mili H, Bicknell E, Blettner M. Development and application of a metric on semantic nets. Systems, Man and Cybernetics, IEEE Transactions on. 1989;19(1):17-30.

[15] Resnik P. In: Using information content to evaluate semantic similarity in a taxonomy. Proc. 14th international joint conference on artificial intelligence vol 1; M. Kaufmann Pub. Inc.; 1995. p. 448-53.

[16] Sanchez D, Batet M, Valls A. Web-based semantic similarity: An evaluation in the biomedical domain. Int. Journal of Soft. and Inform. 2010;4(1):39-52.

[17] Silva C, Salinas M. Modelos de regresión y correlación IV. Spearman. Cienc Trab. 2007:143-5.

[18] Wu Z, Palmer M. In: Verbs semantics and lexical selection. Proceedings of the 32nd annual meeting on association for computational linguistics; Association for Computational Linguistics; 1994. p. 133-8.

[3] http://www-lpir.nist.gov/projects/trecmed/2011/tm2011.html

Revisión sistemática de estrategias para el modelado automático de ontologías borrosas

Emilio Fdez-Viñas[1], **Jesús Serrano-Guerrero**[2], **Francisco P. Romero**[2], **Jose A. Olivas**[2], **Victor H. Menendez-Dominguez**[3]

Departamento de Tecnologías y Sistemas de la Información, Escuela Superior de Informática,
Paseo de la Universidad, 4, 13071, Ciudad Real, Universidad de Castilla-La Mancha.
Universidad Autónoma de Yucatán Periférico Norte 97110 Mérida, México
[1]*Dirección1*, Fdez.Emilio@gmail.com
[2]*Dirección2*, {Jesus.Serrano,FranciscoP.Romero,JoseAngel.Olivas}@uclm.es
[3]*Dirección3*, mdoming@uady.mx

Resumen

La información que el ser humano utiliza para la comunicación está representada por una estructura que sólo puede ser entendida por él mismo. Una posible solución para proporcionar a los sistemas informáticos la capacidad de entender texto en lenguaje natural es la incorporación de etiquetas que aporten semántica o etiquetas que permitan la interpretación correcta de los contenidos. Una posible solución a este problema es el uso de ontologías borrosas que permiten modelar la incertidumbre que existe en la comunicación humana de forma que pueda ser interpretada por máquinas. En este trabajo se presenta una revisión de distintas aplicaciones de la lógica borrosa en el mundo de las ontologías y se comentarán otras definiciones de estructuras que representen el conocimiento junto con distintas estrategias de modelado sugeridas en los últimos años.

Palabras Clave: ontología, lógica borrosa, estrategias de modelado.

1 Introducción

Es sumamente difícil calcular la cantidad de información que el ser humano a lo largo de la historia ha sido capaz de recopilar. Una gran parte de esa información se encuentra representada en lenguaje semiestructurado o en lenguaje natural, ocasionando que sólo sea entendible por el hombre. ¿Por qué no dotar a las máquinas de la capacidad de poder entender dicha información? Como respuesta a esta pregunta y centrándose en el mayor sistema de información que conoce el hombre: la Web, muchos autores han obtenido como respuesta la Web Semántica, una extensión de la Web actual a la que se añaden metadatos y estructuras que representen el conocimiento. La Web Semántica[1] propone superar las limitaciones de la Web actual mediante la introducción de descripciones explícitas sobre su significado, es decir, dotando a la información de un cierto contexto.

La Web Semántica se basa, entre otros elementos, en el uso de las ontologías, las cuales crean una jerarquía de conceptos con atributos y relaciones que definen una terminología consensuada para identificar relaciones semánticas entre entidades [9].

Las aplicaciones de las ontologías no sólo se reducen al mundo de la Web Semántica sino que se puede hacer uso de una ontología en muchas aplicaciones como por ejemplo, dotar a una máquina de la capacidad de "comprender" un texto en lenguaje natural y poder así expandir una consulta formulada a un sistema de recuperación de información en base a un contexto semántico [16] [25] [26], en sistemas recomendadores [22] [12] y demás sistemas basados en conocimiento.

En la actualidad, la Web Semántica no es una realidad palpable, entre otros problemas, por lo costoso y complicado que resulta la creación de una ontología. Incluso acotada una ontología a un dominio reducido, el tiempo y el esfuerzo por parte de los expertos que la llevan a cabo es muy elevado. Debido a este problema diversos investigadores centran sus esfuerzos en intentar automatizar este proceso de modelado de conocimiento mediante el desarrollo de herramientas como FOGA [19] [20] que utiliza Formal Concept Analysis [8] para el modelado y FOG [7] basada en las estrategia de Keyword Connection Matrix [17] y Multi-relationsip [10].

El resto del documento está compuesto por la sección 2 en la cual se introduce el concepto de ontología y la flexibilización del mismo hacia otras estructuras para la representación del conocimiento. En la sección 3 se estudian las ontologías borrosas y como ayudan para el tratamiento de la imprecisión inherente en el lenguaje natural. En la sección 4 se realiza una revisión de distintas estrategias de modelado del conocimiento. Finalmente se terminará con la sección de conclusiones.

2 ¿Qué es una ontología?

La mayoría de la información está estructurada de forma que sólo puede ser entendida por el ser humano. Para proporcionar a los sistemas informáticos la capacidad de entender texto es necesaria la incorporación de etiquetas que aporten semántica o etiquetas que permitan la interpretación correcta de los contenidos. De esta forma un valor como puede ser un número puede indicar, según el contexto, un año, una distancia, una edad, etc. Su significado depende de la etiqueta que se le haya asignado.

El etiquetado y la anotación de los contenidos almacenado en sistemas digitales deben realizarse de una forma estándar para hacer posible el entendimiento de un texto por todos los programas que hagan uso de él. Un marcado consistente implica la estructuración del dominio que se pretende representar detallando las entidades que lo componen junto con la naturaleza de las relaciones existentes entre las entidades.

Es necesario el cumplimiento de estándares para el coherente etiquetado de los contenidos Web de forma que sea posible el intercambio de información entre distintas máquinas. Las ontologías se presentan como una posible solución natural que soporta el etiquetado de la información. Las ontologías son representadas en su mayoría por una notación basada en XML lo que permite que cualquier máquina pueda utilizarlas y compartir información.

Llegados a este punto se puede definir una ontología como *"una especificación formal y explícita de una conceptualización consensuada"* [9]. Esta definición puede ser desglosada de la siguiente forma:

- Una conceptualización implica que toda ontología describe un modelo abstracto del dominio o fenómeno del mundo que representa. Dicho modelo se basa esencialmente en el empleo de conceptos, atributos, valores y relaciones.
- Una especificación explícita expresa que una ontología contiene información de un dominio de una forma coherente y precisa.
- El término formal hace referencia a que toda representación de un dominio se debe realizar de una forma que garantice el entendimiento entre máquinas ya sea para compartir información, aumentar el conocimiento, etc. Su representación debe ser de una forma única y sin ambigüedades.
- Decir que una ontología debe ser consensuada indica que la representación de un dominio debe ser validada por todos los expertos que han participado en su generación.

La definición presentada por Gruber es una de las más aceptadas. Otra definición en la misma línea sobre qué es una ontología es la presentada por Uschold y Gruninger en la cual *"una ontología es un entendimiento compartido de algún dominio de interés"*[21].

La parte más importante de la definición dada por Gruber y que comparten la definición presentada por Uschold es la conceptualización del dominio o fenómeno a modelar.

2.1 Flexibilización del concepto de ontología

Actualmente no existe una única definición formal de qué es una ontología, pero teniendo presente tanto la definición de Gruber [9] como la de Uschold y Gruninger[21] se pueden extraer puntos en común. Según Stumme et al. [18], se puede definir formalmente los elementos comunes de las definiciones de ontologías como una tupla $\mathscr{O} := (\mathscr{C}, is\ a, \mathscr{R})$, donde $\mathscr{C}$ es el conjunto de elementos que agrupa conceptos, *is a* representa un orden parcial en $\mathscr{C}$ (este tipo de relaciones son reflexivas, transitivas y antisimétricas)y $\mathscr{R}$ es el conjunto de elementos denominados relaciones.

Partiendo de esta definición presentada por Stumme sobre el núcleo de una ontologías, existen varias estructuras de datos que permiten la representación del conocimiento perteneciente a un dominio o un fenómeno de una forma precisa y completa que permite el procesamiento de dicho modelo por parte de las máquinas.

En la figura 1, se muestra la relación de similitud que se da entre la definición formal de una ontología y una red semántica [2]. Una red semántica puede representar el mismo conocimiento que una ontología permitiendo ser ambas estructuras de datos utilizadas de la misma forma tal y como se presenta en la siguiente figura[1].

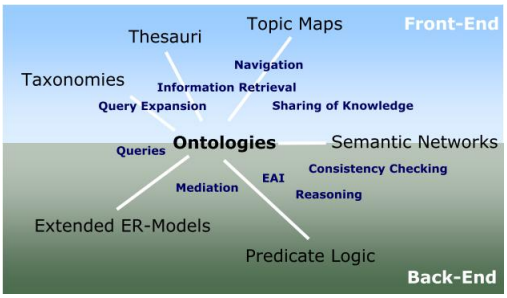


Figura 1: Estructuras para el modelado de conocimiento.

3 Ontologías borrosas

Las ontologías sirven para representar una descripción del conocimiento contenido dentro de un dominio. Este dominio está representado en lenguaje natural caracterizado por

[1] Paul Buitelaar, Philipp Cimiano, Marko Grobelnik, Michel Sintek; *Ontology Learning from text.* Tutorial at ECML/PKDD, Oct. 2005, Porto, Portugal

la imprecisión del lenguaje natural y las relaciones semánticas inherente a la utilización de conceptos en las descripciones. Por estas características las ontologías clásicas pueden no representar un dominio de una forma exacta, por lo que es necesario dotar a las representaciones ontológicas de una flexibilidad similar a la utilizada en el lenguaje natural para la descripción de dominios. La incorporación de la lógica borrosa permite establecer los grados de relación semántica entre los conceptos jerarquizados dentro de una ontología.

El grado de relación en el que dos conceptos pertenecientes a un domino se relacionan, está determinado por un valor borroso que caracteriza esa relación en base al contexto al que pertenecen. Una definición formal de ontología borrosa es la presentada por Calegari [3] en la cual una ontología es una cuádrupla $\mathbf{O}_F = (\mathbf{C}, \mathbf{R}, \mathbf{F}, \mathbf{A})$ donde:

- **C** es el conjunto de conceptos o entidades.
- **R** es el conjunto de relaciones R. Cada $R \in \mathbf{R}$ es una relación borrosa en el dominio de entidades $R : \mathbf{C}^n \rightarrow [0,1]$. En particular $\mathbf{R} = T \bigcup T_{not}$ donde T es el conjunto de relaciones taxonómicas y T_{not} es el conjunto de relaciones no taxonómicas.
- **F** es el conjunto de relaciones borrosas en el conjunto de entidades C y un dominio específico en $\mathbf{E} = \{enteros, cadena\ de\ caracteres, ...\}$.
- **A** es un conjunto de axiomas expresados en un lenguaje lógico como por ejemplo, predicados que restringen el significado de los conceptos y relaciones.

Está claro que el tipo de relaciones que se pueden establecer en una ontología clásica para representar una relación del tipo: *"Rome is a city of Italy"*, pueden ser traducidas por una instanciación de individuos y relaciones que resulta de la siguiente forma: $Rome \xrightarrow{1} Italy$. En las ontologías borrosas se puede llegar a modelar expresiones como *"Paul lives sometimes in London and sometimes in Rome"* lo que daría como resultado $Lives(Paul, London) = 0.6$ y $Lives(Paul, Rome) = 0.5$ y podría ser representado como $Paul \overset{Lives0.6}{\rightarrow} London$ y $Paul \overset{Lives0.5}{\rightarrow} Rome$.

Según Calegari S., los tipos de relaciones utilizadas en su definición de ontologías borrosas son las relaciones reflexivas, simétricas y transitivas [3].

- **Reflexiva.** Si y sólo si $\forall i \in \mathbf{C}, R(i,i) = 1$.
- **Simétrica.** Si y sólo si $\forall i, j \in \mathbf{C}, R(i,j) = R(j,i)$.
- **Transitiva.** Si y sólo si $\forall i, j \in \mathbf{C}, sup_{k \in \mathbf{c}} \{t(R(i,k), R(k,j))\} \leq R(i,j)$. Donde t es una t-norma como por ejemplo "min".

4 Estrategias de modelado automático

En la actualidad existen diversas ontologías como son la conocida Wordnet [23], EuroWordNet [6], etc. Estas ontologías son mundialmente reconocidas y de carácter general, lo que por otra parte las convierte en recursos de conocimiento no especializados y por tanto, no permiten una amplia especificación. Partiendo de esta idea, diversos investigadores desarrollaron distintas estrategias para la creación de modelos de conocimiento que puedan presentarse como ontologías y cubran la necesidades específicas de una aplicación o un uso concreto.

Elevando el punto de vista de la definición de ontología presentada por Thomas Gruber [9] y haciendo uso de la teoría de conjuntos borrosos de Zadeh [24], un modelo de conocimiento basado en una estructura ontológica para la representación semántica del contenido de una colección finita de documentos se puede definir como un conjunto de conceptos contenidos en un dominio acotado y de relaciones borrosas que caracterizan el grado de proximidad semántica de dichos conceptos.

Muchas de las ontologías existentes son generadas por la acción de expertos. Esta forma de modelar el conocimiento requiere un gran esfuerzo por parte de los humanos que intervienen en el modelado y una gran inversión de tiempo. En el caso de que el dominio a modelar no esté bien acotado, dicha tarea se puede convertir en algo inabordable. Por este motivo diversos autores comenzaron a desarrollar sistemas automáticos para el modelado de conocimiento.

Como resultado de este modelado automático de conocimiento se crea una estructura jerárquica compuesta por nodos que representan los conceptos y la relaciones borrosas que caracterizan la semántica entre los conceptos que definen un dominio. Uno de los primeros algoritmos utilizados para el modelado de conocimiento utilizado en sistemas de recuperación de información, es el presentado por Ogawa en 1991 [17], en el cual se identificaba de forma automática las palabras más relevantes en una colección. Estas palabras son conocidas como "'keywords"' y junto con las palabras se identifican las relaciones entre las mismas basándose en la frecuencia de aparición de cada par de palabras por la estrategia basada en la co-ocurrencia entre términos, dando lugar a una estrategia de modelado llamada "'keyword Connection Matrix"'

La idea lanzada por Ogawa [17] fue utilizada por Ilan Itzkovich en 1994 [11] para el desarrollo del concepto de las redes conceptuales, como estructura para el modelado de conocimiento, en el que se presentaban dos tipos de relaciones: las relaciones borrosas de sinonimia y las relaciones borrosas de generalización. Para poder desarrollar una red conceptual borrosa, primero se debe definir que es una red conceptual. Según Itzkovich [11] una red conceptual (CN) consiste en un conjunto finito de elementos C lla-

mado conceptos y un conjunto R de relaciones, donde R puede ser de dos tipos: relaciones de sinonimia R_{syn} y relaciones de generalización R_{gen} sobre C, con la extensión de la generalización de la siguiente forma: Si $(C_k,C_i) \in R_{syn}, (C_j,C_l) \in R_{syn}, (C_i,C_j) \in R_{gen}$ y $(C_j,C_i) \notin R_{gen}$, entonces $(C_k,C_l) \in R_{gen}$ y $(C_l,C_k) \notin R_{gen}$.

Formalmente se puede definir relación de sinonimia R_{syn} sobre un conjunto finito C como una relación binaria que satisface las siguientes propiedades [11]:

- **Reflexiva.** $(C_i,C_i) = 1$
- **Simétrica.** Si $(C_i,C_j) \in R_{syn}$, entonces $(C_j,C_i) \in R_{syn}$
- **Transitiva.** Si $(C_i,C_j) \in R_{syn}$ y $(C_j,C_k) \in R_{syn}$, entonces $(C_i,C_k) \in R_{syn}$

De la misma forma, la relación de generalización R_{gen} sobre un conjunto finito de C se puede definir como una relación binaria que satisface las siguientes propiedades [11]:

- **Reflexiva.** $(C_i,C_i) \in R_{gen}$
- **Antisimétrica.** Si $(C_i,C_j) \in R_{gen}$, entonces $(C_j,C_i) \notin R_{gen}$, por lo que se deduce el elemento más a la derecha es el más general
- **Transitiva.** Si $(C_i,C_j) \in R_{gen}$ y $(C_j,C_k) \in R_{gen}$, entonces $(C_i,C_k) \in R_{gen}$

Así pues, la red conceptual borrosa (FCN) propuesta por Itzkovich, parte de la idea de red conceptual (CN) y se puede definir como un conjunto finito de elementos C llamados conceptos y un conjunto de relaciones borrosas R, donde R queda definida por dos tipos de relaciones: relación de similitud R_{sim} y relación de generalización R_g con la extensión de la propiedad de generalización establecida de la siguiente forma: $\forall\ i,j,k$ y l, Si $\mu_g(C_i,C_j) > 0$ y $\mu_g(C_j,C_i) = 0$, entonces

$$\mu_g(C_k,C_l) \geq \forall_{C_i,C_j}(\exists(\mu_{sim}(C_k,C_i),\mu_g(C_i,C_j),\mu_{sim}(C_j,C_l)))$$

La relación de similitud borrosa R_{sim} utilizada en la definición de FCN se puede entender como una relación binaria sobre un conjunto finito de conceptos C que satisfaga las siguientes propiedades [11]:

- **Reflexiva.** $\mu_{sim}(C_i,C_i) = 1$
- **Simétrica.** $\mu_{sim}(C_i,C_j) = \mu_{sim}(C_j,C_i)$
- **Transitiva.** $\mu_{sim}(C_i,C_k) \geq \forall_{C_j}(\mu_{sim}(C_i,C_j) \wedge \mu_{sim}(C_j,C_k))$

De la misma forma se puede definir el grado borroso de generalización R_g sobre un conjunto finito de conceptos C como una relación binaria que satisface las siguientes propiedades [11]:

- **Reflexiva.** $\mu_g(C_i,C_i) = 1$
- **Antisimétrica.** Si $\mu(C_i,C_j) > 0$ y $\mu_g(C_j,C_i) > 0$, entonces $C_i = C_j$
- **Transitiva.** $\mu(C_i,C_k) \geq \forall_{C_j}(\mu_g(C_i,C_j) \wedge \mu_g(C_j,C_k))$

Haciendo uso de esta definición de red conceptual borrosa (FCN), Chen et al. [4] extendieron la forma de representar el grado de las relaciones que caracterizaban el vínculo de generalización y similitud borrosa. En este trabajo se propuso crear un intervalo borroso para caracterizar las relaciones de la red conceptual borrosa. La incorporación del concepto de intervalo, modifica los conceptos de red conceptual y las relaciones que lo componen.

Una relación con intervalo de similitud R_{ivsim} sobre un conjunto de conceptos C es una relación binaria borrosa que satisface las siguientes necesidades [4]:

- **Reflexiva.** $\mu_{ivsim}(C_i,C_j) = [1,1]$
- **Simétrica.** $\mu_{ivsim}(C_i,C_j) = \mu_{ivsim}(C_j,C_i)$
- **Transitiva.** Supongamos el intervalo de similitud entre los conceptos C_x y C_y representado por $\mu_{ivsim}(C_x,C_y)$, donde $\mu_{ivsim}(C_x,C_y) = [S^l(C_x,C_y),S^h(C_x,C_y)]$ y $0 \leq S^l(C_x,C_y) \leq S^h(C_x,C_y) \leq 1$ entonces $S^l(C_i,C_k) \geq \forall_{(C_j)}(S^l(C_i,C_j) \wedge S^l(C_j,C_k))$,$S^h(C_i,C_k) \geq \forall_{(C_j)}(S^h(C_i,C_j) \wedge S^h(C_j,C_k))$,

De igual forma se puede definir el intervalo de generalización R_{ivg} sobre un conjunto finito de elementos C como una relación binaria la cual satisface las siguientes propiedades [4]:

- **Reflexiva.** $\mu_{ivg}(C_i,C_i) = [1,1]$
- **Antisimétrica.** Si $\mu_{ivg}(C_i,C_j) > [0,0]$ y $\mu_{ivg}(C_j,C_i) > [0,0]$, entonces $C_i = C_j$
- **Transitiva.** Supongamos el grado de generalización entre los conceptos C_x y C_y representados por $\mu_{ivg}(C_x,Cy)$, donde $\mu_{ivg}(C_x,C_y) = [g^l(C_x,C_y),g^h(C_x,C_y)]$ y $0 \leq g^l(C_x,C_y) \leq g^h(C_x,C_y) \leq 1$. Entonces $g^l(C_x,C_y) \geq \forall_{C_j}(g^l(C_i,C_j) \wedge g^l(C_j,C_k))$, $g^h(C_x,C_y) \geq \forall_{C_j}(g^h(C_i,C_j) \wedge g^h(C_j,C_k))$

Teniendo presente esta tendencia, en 1991 Lucarella y Morara [15] presentaron un sistema de recuperación de información basado en conocimiento. Dicho conocimiento está estructurado en una red conceptual borrosa basada en la teoría de conjuntos borrosos de Zadeh[24] de forma que permita representar las relaciones entre conceptos. Con el objetivo de aumentar la velocidad de inferencia de conocimiento en la red conceptual, Chen et al.[5] propusieron un método para los sistemas de recuperación de información,

donde la matriz de conceptos es usada por la red conceptual permitiendo el cálculo de la clausura de transitividad. Con este cálculo se puede obtener un nuevo valor implícito en la relaciones entre los pares de conceptos que forman la red. Tanto en la propuesta de Lucarella como la de Chen, el peso que tenía cada relaciones en la red, era generado por la acción de expertos.

Estas redes conceptuales además de establecer las relaciones entre conceptos también establecen las relaciones entre conceptos y documentos, por tanto se identifica el peso que tiene cada concepto respecto a un documento. Lucarella y Chen, formulaban que las relaciones entre nodos fuesen únicas y fijas. En 2003, Horng et al.[10] propusieron que las relaciones entre conceptos dependieran de la perspectiva, es decir, según el contexto en el que se formule una pregunta, se puede elegir una relación u otra entre un par de nodos. Horng, establece cuatro tipos de asociaciones entre nodos: asociación borrosa positiva, asociación borrosa negativa, asociación borrosa de generalización y asociación borrosa de especificación. A cada una de estas relaciones se les atribuye un valor, proporcionando así semántica a la unión entre nodos en base a un contexto y de la unión de todas las relaciones se obtiene como resultado una red conceptual multi-relacionado. Esta red conceptual borrosa basado en multi-relaciones queda definido por una dupla MRFCN(*E,L*), donde E es un conjunto de nodos que definen un concepto o un documento y L es un conjunto de arcos dirigidos entre los nodos. Si $\ell \in L$ entonces el grafo dirigido ℓ tiene uno de los siguientes formatos:

- $C_i \overset{(\langle\mu_p,P\rangle,\langle\mu_n,N\rangle,\langle\mu_g,G\rangle,\langle\mu_s,S\rangle)}{\longrightarrow} C_j$, significa que el grafo dirigido ℓ conecta desde el concepto C_i al concepto C_j con una asociación cuádrupla $\langle\mu_p,P\rangle,\langle\mu_n,N\rangle,\langle\mu_g,G\rangle,\langle\mu_s,S\rangle$, donde μ_p indica que el conceptos C_i y el concepto C_j están relacionados con un grado de asociación borrosa positiva; μ_n indica que el concepto C_i y el concepto C_j, están relacionados con una asociación borrosa negativa; el concepto C_i está relacionado con un concepto C_j más general con un grado μ_g y a su vez, el concepto C_i está relacionado con un concepto C_j más específico con un grado μ_s
- $C_i \overset{\mu}{\rightarrow} d_j$ significa que el documento d_j contiene al concepto C_i con un grado de relación μ

Utilizando la idea propuesta por Horng, en la cual, además de proponer una estructura para representar el conocimiento, se crea una estructura para la indexación de la colección documental que define el dominio. En 2008 Leite et al.[13] [14] partiendo de la estructura propuesta por Horng, basan el modelado del conocimiento como una ontología de dos niveles: El primer nivel representa las categorías y el segundo nivel representa las palabras claves del dominio. Las categorías apuntan hacia los conceptos genéricos y las palabras claves apuntan a los conceptos específicos en un dominio. Las palabras claves y las categorías están relacionadas por una relación borrosa que determina la importancia de su asociación mediante un valor en el conjunto $[0,1]$. Los documentos están asociados a la categoría y a las palabras claves independientemente de forma que caracterizan la importancia del concepto en el documento.

Este tipo de modelado puede ser considerado como ontologías ligeras, donde cada una de las relaciones que dan forma a las categorías y a las palabras claves se vinculan por asociaciones borrosas de especificación y generalización como propone Horng y distintas ontologías pueden ser relacionados mediante una relación borrosa positiva o una relación borrosa negativa.

En la propuesta de Leite et al. el conocimiento es representado en múltiples ontologías ligeras, las cuales corresponden a distintos dominios. Cada ontología que representa un dominio es un conjunto de conceptos $D_k = C_{k1}, C_{k2}, ..., C_{ky}$ donde $1 \leq \kappa \leq K$, K es el número del dominio y $||D_k||$ es el número de conceptos en cada dominio. La composición de estas ontologías representan el conocimiento contenido en las colecciones documentales. El concepto modelado en la ontología está organizado como una taxonomía con relaciones borrosas de especificación y relaciones borrosas de generalización. Los conceptos pertenecientes a distintas ontologías pueden estar relacionados por asociaciones borrosas positivas y asociaciones borrosas negativas.

5 Conclusiones

En este artículo se ha elaborado una lista de los trabajos más representativos que existen en la actualidad de las aplicaciones de la lógica borrosa al mundo de las ontologías. La aplicación de la lógica borrosa es una posible solución para salvar el problema de rigidez que presentan las ontologías clásicas en la conceptualización de ciertos tipos de información.

En los últimos tiempos son muchos los trabajos surgidos que aprovechando las características que ofrece la lógica borrosa ayudan a cubrir aspectos como el modelado de conocimiento en ciertos contextos. Algunas de las principales aplicaciones del modelado de conocimiento se centran en los sistemas de recuperación de información, a distintas fases como la indexación y la expansión.

Los trabajos presentados en este artículo son una posible solución para automatizar el proceso de generación de ontologías de forma que se pueda agilizar el trabajo por parte de los expertos humanos que necesitan de mucho tiempo y esfuerzo para el modelado de grandes dominios y suplir las carencias que se presentan al utilizar ontologías genéricas que no proporcionan una amplia especificación dentro de un dominio.

Agradecimientos

Este proyecto está financiado gracias al proyecto FIDELIO (TIN2010-20395,MEC-FEDER), proyecto de excelencia SCAIWEB-2 (PEIC09-0196-3018,JCCM), y al proyecto HECAOT(TC20112330,UCLM).

Referencias

[1] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. *Scientific American* 248(5) 34-43, 2001.

[2] Buitelaar, P., Cimiano, P.,Magnini, B.: Ontology Learning from Text. *In conjunction with the ECML/PKDD 2005 Workshop on: Knowledge Discovery and Ontologies (KDO-2005)*, 2005

[3] Calegari, S., Sanchez, E.: Object-Fuzzy Concept Network: An Enrichment of Ontologies in Semantic Information Retrieval. *Journal of the American Society for Information Science* 59(2006), 2171-2185. 2008.

[4] Chen, S.: Finding inheritance hierarchies in interval-valued Fuzzy concept-networks. *Fuzzy Sets and System* 84(1), 75-83, 1996.

[5] Chen, S.M., Horng, Y.J., Lee, C.H.: Document Retrieval Using Fuzzy-Valued Concept Networks. *IEEE Transactions on systems, man, and cybernetics. Part B, Cybernetics: a publication of the IEEE Systems, Man, and Cybernetics Society* 31(1), 111-118, 2001

[6] EuroWordNet. *http://www.illc.uva.nl/EuroWordNet/*

[7] Fdez-Vinas, E. Serrano-Guerrero, J. Olivas, J.A.: FOG: Arquitectura Flexible para la generación automática de ontologías. *ESTYLF'10: Congreso español de tecnología y lógica fuzzy* 1-6, 2010.

[8] Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. *Springer* (1999).

[9] Gruber, T.: Toward principles for the design of ontologies used for knowledge sharing. *Internacional Journal of Human-Computer Studies* 43(5-6), 907-928, 1995.

[10] Horng, Y.J., Chen, S.M., Lee, C.H.: Automatically constructing multi-relationship fuzzy concept networks for document retrieval. *Applied Artificial Intelligence* 17(4), 303-328, 2003.

[11] Itzkovich, I.: Fuzzy extension of inheritance hierarchies. *Fuzzy Sets and Systems* 62(2), 143-153, 1994.

[12] Lee, C.S., Chang, Y.C., Wang, M.H.: Ontological recommendation multi-agent for Tainan City Travel. *Expert System with Applications* 36(3), 6740-6753, 2009.

[13] Leite, M.A., Ricarte, I.: A Framework for Information Retrieval Based on Fuzzy Relations and Multiple Ontologies. *In: Lecture Notes in Computer Science* 292-301, 2008.

[14] Leite, M.a., Ricarte, I.:Fuzzy information, Retrieval Model Based on Multiple Related Ontologies. *20th IEEE International Conference on Tools with Artifical Intelligence* 309-316, 2008.

[15] Lucarella, D., Morarar, R.: FIRST: Fuzzy Information, Retrieval sySTem. *Journal of Information Science* 17(2), 81-91, 1991.

[16] Miyamoto, S.: Information tetrieval based on fuzzy associations. *Fuzzy Sets and Systems* 38(2), 191-205, 1990.

[17] Ogawa, Y. Morita, T., Kobayashi, K.: A fuzzy document retrieval system using the keyword connection matrix and a learning method. *Fuzzy Sets and System* 39(2), 163-179, 1991.

[18] Stumme, G., Maedche, A.: FCA-MERGE:Bottom-Up Merging of Ontologies. *In: Proceeding of 17th Int'l Joint Conf. Artificial Intelligence (IJCAI'01).* Morgan Daufmann, San Francisco, 225-230, 2001.

[19] Tho, Q., Hui, S., Cao, T.: FOGA: A Fuzzy Ontology Generator Framework for Scholary Semantic Web. *Knowledge Discovery and Ontologies* Pisa, Italy, 37-48, 2004.

[20] Tho, Q., Hui, S. Fong, A.: Automatic fuzzy ontology generation for semantic Web. *IEEE Transactions on Dnowledge and Data Engineering* 18(6), 842-856, 2006.

[21] Uschold, M., Gruninger, M.: Ontologies: Principles, Methods and Applications. *knowledge Engineering Review* 11(2), 96-136, 1996.

[22] Wei, C., Qing, Y., Li, Z., Bin, W.: A Fuzzy Ontology Generation Framework and Its Application to Subjective Credit Reporting Management. *Intenational Forum on Information Technology and Applications* 377-381, 2009.

[23] WordNet. *http://wordnet.princeton.edu/*

[24] Zadeh, L.: Fuzzy Sets. *Information and Control* 8,338-353, 1965.

[25] Zhai, J. Liang, Y., Yu, Y. Jiang, J.: Semantic Information Retrieval Based on Fuzzy Ontology for Electronic Comerce. *Journal Software* 3(9), 20-27, 2008.

[26] Zhai, J., Wang, Q.:Application of Fuzzy Onotology Framework to Information Retrieval for SCM. *International Symposiums on Information Processing* (1)173-177, 2008.

Un sistema automático de Control de Autoridades para Bibliotecas Digitales

Irene Diaz-Valenzuela[1], **Maria J. Martin-Bautista**[1], **M. Amparo Vila**[1]

[1]*Departamento de Ciencias de la Computación e Inteligencia Artificial. Universidad de Granada*,
{idiazval,mbautis,vila}@decsai.ugr.es

Resumen

En el presente documento, se propone un método automático para el control de autoridades basado en técnicas de Minería de Datos que utiliza datos obtenidos de Bibliotecas Digitales en la Web. Basándonos en este método, se realizará una comparación de tres fuentes de datos Web: CiteSeerX, DBLP e INSPEC, como ejemplos de Bibliotecas Digitales dedicadas a publicaciones científicas.

Palabras Clave: Control de Autoridades, Bibliotecas Digitales Web, Minería de Datos.

1 INTRODUCCIÓN

El control de autoridades es un problema clásico de gran importancia dentro del campo de la documentación. Una definición de control de autoridades podría ser la dada en [5]: *El control de autoridades es el proceso de unificar mediante la utilización de una forma normalizada, los puntos de acceso de los catálogos automatizados y mostrar además las relaciones entre los distintos puntos de acceso. Es decir, supone la normalización de los nombres de personas, entidades, títulos uniformes o materias, que pueden constituir el punto de acceso principal o los secundarios de un catálogo automatizado.* En resumen, podría decirse que el control de autoridades es el proceso por el cual distintas representaciones de una única entidad son identificadas de forma unívoca.

Existen distintos tipos de autoridades, clasificados según las áreas a las que se dediquen. En este trabajo, nos centraremos en el control de *Autoridades de Nombre*, el cual viene definido por las menciones de autoría de una publicación. Dentro de esta disciplina pueden distinguirse dos problemas distintos:

- Incertidumbre de la Autoría
- Identificación de diferentes representaciones para un mismo nombre.

Concretamente, la *Incertidumbre de la Autoría* intenta identificar la autoría de una publicación entre los autores con el mismo nombre que el firmante de dicha publicación. En el caso de la *Identificación de diferentes representaciones para un mismo nombre*, se pretende identificar y unificar las distintas representaciones que para un nombre de autor puedan existir.

Este tipo de tareas tradicionalmente se han venido realizando por un bibliotecario, quién, en calidad de experto era capaz de distinguir la autoría y asignar identificadores unívocos válidos a las distintas entradas contenidas en su biblioteca. Con la aparición de las *Bibliotecas Digitales* en la Web, en las que sus registros generalmente han sido obtenidos de forma automática, y no suele existir esta figura, los problemas anteriormente mencionados se vuelven de difícil solución.

Para nuestro propósito, definiremos una *Biblioteca Digital* como una colección de artículos e información sobre sus referencias en un determinado dominio. Ejemplos de Bibliotecas Digitales son DBLP, CiteSeerX o INSPEC, entre otros. Una *citación* o *referencia* es un conjunto de campos bibliográficos como el autor, título, la publicación o la fecha de publicación del mismo, que caracterizan a un documento y constituyen su entidad bibliográfica.

1.1 NUESTRO ENFOQUE

En este trabajo pretendemos presentar un método automático para el *Control de Autoridades* basado en técnicas de Minería de Datos. Para ello, se tomará como entrada un apellido que será utilizado como consulta en una Biblioteca Digital, mostrándose como salida las autoridades obtenidas de dicha búsqueda.

Se pretende que este método sea capaz de resolver al mis-

mo tiempo los dos problemas clave del Control de Autoridades: la *Incertidumbre de la Autoría* y la *Identificación de diferentes representaciones para un mismo nombre*. Así mismo, se establece la necesidad de que el método trabaje con datos obtenidos de distintas fuentes en la web.

El presente documento se encuentra estructurado de la siguiente manera: en la Sección 1.2 se da un estado del arte del problema, para pasar a una descripción detallada del sistema propuesto en la Sección 2. A continuación se detallan una serie de resultados experimentales en la Sección 3 que incluyen una comparativa de fuentes de datos en la Web 3.2. Finalmente la Sección 4 ofrece unas conclusiones.

1.2 ANTECEDENTES

Existen en la literatura algunas aproximaciones a este problema. Trabajos anteriores como [4] proponen un enfoque basado en *K-way spectral clustering* para resolver el problema de la *Incertidumbre de la Autoría* sobre nombres. En [6] se realiza un estudio completo del problema del control de autoridades en Bibliotecas Digitales en la Web y proponen un sistema manual basado en la web semántica para dar solución a esta cuestión.

En [3] se proponen dos modelos para la desambiguación automática de nombres propios en citaciones (*Incertidumbre de la Autoría*), utilizando una medida basada en tres parámetros, el título del documento, nombres de coautores y el título de la publicación donde apareció. Estos dos enfoques propuestos, se basan en un modelo probabilístico naive Bayes y en Support Vector Machines(SVMs) respectivamente.

Otro modelo automático es el propuesto en [7] para determinar la autoría en partituras musicales, comparando los registros obtenidos con los ficheros de autoridades de la Biblioteca del Congreso Americano.

Sin embargo, si comparamos nuestra propuesta con las anteriores, vemos que exceptuando a [7], estas propuestas se limitan al problema de la *Incertidumbre de la Autoría*, sin cubrir la identificación de las distintas representaciones de un mismo nombre. En este campo cabe destacar la existencia de herramientas manuales como la incluida en el repositorio SCOPUS, que permite a un autor indicar cuales son las distintas representaciones existentes para su nombre. En nuestro caso, al trabajar con datos Web, debemos valernos únicamente de la información disponible, sin tener disponible ningún fichero de autoridades previamente creado.

2 DESCRIPCIÓN GENERAL DEL SISTEMA

El sistema desarrollado consta de tres partes bien diferenciadas como puede verse en la Figura 1:

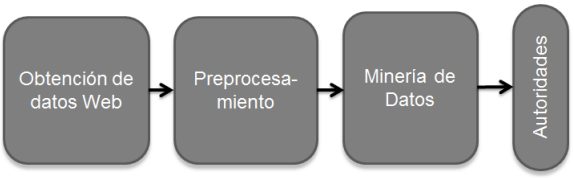


Figura 1: Esquema del Sistema completo.

- **Proceso de obtención de datos Web:** Durante esta fase inicial se obtienen los datos que serán utilizados durante el resto del proceso. Al ser datos de la Web, el sistema los localiza y extrae, almacenando la información que considera relevante en una base de datos preliminar.
- **Preprocesamiento:** En este proceso la base de datos preliminar se filtra y limpia con el objeto de representar los distintos tipos de datos contenidos en una forma canónica que permita su posterior procesamiento.
- **Proceso de minería de datos:** Esta última fase del sistema es la más importante. Comienza expresando los datos preprocesados en una forma intermedia que servirá de base para calcular las autoridades. Dicha información pasa por un proceso de clustering del cual se extraen las autoridades. Finalmente, una etapa de validación permite valorar la solución obtenida y vuelve atrás en caso de que ésta no sea lo suficientemente buena para ser aceptada.

2.1 PROCESO DE OBTENCIÓN DE DATOS

La obtención automática de autoridades que proponemos en este documento parte de información obtenida de la Web. La entrada al sistema consiste en una búsqueda por nombre en una Biblioteca Digital en la Web. En dicha búsqueda se consulta un nombre de autor y se extraen los resultados que sobre dicha consulta ofrezca la fuente. Estos resultados son almacenados en una *Base de Datos Preliminar* para su posterior tratamiento.

Considerando la definición previamente dada de *Bibliotecas Digitales*, es posible encontrar en la Web distintos ejemplos que cumplen estas características, como: CiteSeer, INSPEC, DBLP, Scopus, Google Scholar, etc. Todos estos servicios permiten la consulta de bibliografía de una forma fácil y rápida, contando con buscadores que permiten realizar consultas avanzadas, tales como búsquedas por autor, por año, por título, etc. Sin embargo, en la gran mayoría de ellos, no existe una cohesión entre nombres de autores, y en el caso de que exista como puede ser en DBLP, en bastantes ocasiones presenta errores.

Pese a que cada una de las fuentes mencionadas posee unas características específicas que las diferencian entre ellas, existe una serie de información que es común a todas. De

entre estos datos, cabe destacar la *referencia bibliográfica*, que será utilizado para obtener los datos con los que trabajará el sistema. En el caso en el que la referencia bibliográfica no ofrezca el *Abstract* del documento, este deberá ser extraído de forma independiente a la estructura, siempre y cuando esta información esté disponible.

2.2 PREPROCESAMIENTO

La información contenida en la web sufre de falta de autoridad central, lo que hace que esta información pueda estar incompleta y contener diferentes niveles de ruido. Generalmente, esto suele deberse a errores de codificación o procesos de captura automática con un nivel bajo de supervisión etc. Todo esto provoca que este tipos de información esté gravemente desectructurada, lo que provoca que su procesamiento de forma automática no sea sencillo. Esto pone de manifiesto la necesidad de buscar una representación intermedia de los datos de entrada que pueda ser utilizada en etapas posteriores del sistema.

Como paso previo a la obtención de dicha representación intermedia, se realizará un preprocesamiento de los datos con el objeto de limpiarlos de ruido y obtener una forma canónica que unifique representaciones y simplifique las comparaciones. Para dicho preprocesamiento, distinguiremos dos tipos de datos: información referente a nombres de autores e información referente a títulos y abstract.

Para el primer grupo de datos, buscaremos una representación canónica que nos permita unificar los distintos modos en los que un nombre pueda estar expresado. Para ello, utilizaremos un filtrado básico el cual elimina caracteres especiales, espacios extra y otro tipo de información no deseada, como los acentos. Del mismo modo, se transforma la entrada a una representación en mayúsculas buscando estandarizar todas las representaciones existentes de una misma cadena y conseguir de este modo esa forma canónica buscada.

Para el segundo caso, la extracción de palabras clave a partir de títulos y abstract, se aplicarán una serie de filtros, bien conocidos en la literatura, a la cadena completa con objeto de extraer dichas palabras. Dichos filtros son: *Tokenización*, *Eliminación de números*, *Eliminación de plurales*, *Eliminación de StopWords* y *Lematización*. Más detalles sobre estos filtros pueden encontrarse en [2].

2.3 PROCESO DE MINERÍA DE DATOS

Tras el proceso de *Obtención de Datos de la Web* y el *Preprocesamiento* tenemos los datos listos para trabajar. Como podemos ver en la Figura 2 el proceso seguido para obtener las autoridades a partir de los datos almacenados en la *Base de Datos Limpia* consta de tres partes:

- **Representación Intermedia:** Durante esta etapa se pretende buscar la forma más apropiada de representación de los datos. Del modo en el que se considere esta representación intermedia dependerán las futuras fases del proceso.
- **Proceso de Clustering:** Se aplica un proceso de agrupamiento (clustering) jerárquico a la representación intermedia obtenida en el paso anterior. Dicho proceso tiene como objetivo obtener grupos de elementos que se correspondan con una autoridad.
- **Validación:** El proceso de validación es fundamental en este sistema, ya que es el encargado de evaluar la solución obtenida.

Figura 2: Esquema del Proceso de Minería de Datos.

La Representación Intermedia utilizada consiste en una tabla cuyas filas representan las publicaciones obtenidas como resultados de la búsqueda y las columnas son los atributos de dichas publicaciones. Un esquema de esta representación se puede ver en la Tabla 1.

Tabla 1: Esquema de la Representación Intermedia Utilizada.

id	autor	PC_1	...	coautores

Cada una de los vectores de elementos que conforman la tabla esta compuesto de cuatro componentes bien diferenciadas:

- Un *identificador*
- El *nombre del autor* por el que se consultó en la búsqueda.
- Una lista de los pesos asociados a las *Palabras Clave (PC)* obtenidas del Título y el Abstract.
- Una lista con el resto de *coautores* de la publicación.

Estos vectores de elementos tomarán los valores que se obtuvieron tras el preprocesamiento de la información obtenida de la Web. Para las palabras clave se tomará una lista formada con la unión de todos los términos extraídos del preprocesamiento, los cuales tendrán un peso asociado según la siguiente expresión:

$$w(PC) = \begin{cases} w_1, & \text{si } PC \in \text{titulo;} \\ w_2, & \text{si } PC \in \text{abstract;} \\ w_3, & \text{en otro caso.} \end{cases} \quad (1)$$

Donde fijaremos $w_1 = 2$, $w_2 = 1$ y $w_3 = 0$ con el objetivo de favorecer a aquellas palabras que aparezcan en el título frente a las que aparezcan en el abstract. Las componentes del vector serán posteriormente usadas para calcular la distancia entre elementos, que permitirá calcular la matriz de distancia, necesaria para el proceso de Clustering. Dicha distancia se calcula según la siguiente expresión:

$$d(f_1, f_2) = \begin{cases} 1, & d_n = 1; \\ (1 - \alpha_1 - \alpha_2) \times d_n + & \\ \alpha_1 \times d_t + \alpha_2 \times d_c, & \text{en otro caso.} \end{cases} \quad (2)$$

Donde α_1 y α_2 son pesos en el intervalo [0-1], establecidos por el usuario en función de los datos de entrada. d_n, d_t y d_c son respectivamente, las componentes de la distancia: la *distancia de nombre*, *distancia de términos* y *distancia de coautores*. Cada una de estas tres componentes se calcula a partir de la información contenida en la *Representación Intermedia*. Cabe destacar que estamos ante una medida de similitud.

La distancia de nombres, d_n, compara los nombres de los autores que se han obtenido como resultado de la búsqueda, es decir, si el apellido consultado era *García*, se compararán aquellos nombres que contengan dicho apellido. Para hacer esta comparación se seguirá el esquema de pesos descrito en [2]. Esto nos devolverá un valor en el intervalo [0-1] indicándonos cual es la similitud entre los nombres considerados.

La distancia de términos, d_t se calcula utilizando los pesos de las palabras clave. Para ello usaremos la similitud del Coseno, con la siguiente expresión:

$$\cos(s,t) = \frac{\sum_{j=1}^{k} s_j t_j}{\sqrt{\sum_{j=1}^{k} s_j^2} \sqrt{\sum_{j=1}^{k} t_j^2}}. \quad (3)$$

Finalmente, la distancia de coautores, d_c se calcula comparando todas las parejas de nombres de coautores de las dos listas y tomando la de mayor similitud. Para hacer las comparaciones de nombres se utiliza el mismo esquema que para la distancia de nombres, y que puede verse en [2].

Definida la medida de distancia, crearemos una matriz simétrica $N \times N$ en la que se definirá la distancia entre dos resultados cualesquiera de la búsqueda. Esta matriz servirá de entrada a un algoritmo de *Clustering jerárquico*, el cual nos dará el agrupamiento en autoridades. Se ha utilizado clustering jerárquico en lugar de otro tipo de enfoque porque no disponemos de ningún tipo de información a priori sobre el número de autoridades que vamos a obtener. Dentro de los algoritmos de clustering jerárquico, vamos a utilizar el algoritmo *Complete Linkage* dado que es apropiado para nuestra medida de distancia, y su forma de calcular la distancia intra-cluster nos ofrece mejores resultados experimentales que otros enfoques clásicos.

Una vez calculado el árbol del cluster jerárquico, debemos definir el punto en el cual realizaremos el corte que separará los clusters resultado de la clasificación. Para esta tarea utilizaremos el enfoque propuesto en [1] el cual se basa en las media de las distancia intra-cluster y entre-clusters. Por tanto, el punto de corte del cluster vendrá dado por el máximo de dicha función definida en cada α-corte del mismo. Descubierto dicho punto, un corte del cluster en esa magnitud nos calculará los grupos correspondientes a las autoridades que estábamos buscando. Por lo tanto, una autoridad estará formada por un grupo de vectores de elementos correspondientes cada uno de ellos a una publicación.

2.4 INDICES PARA EL ANÁLISIS DE LA BONDAD DE LOS RESULTADOS

Para evaluar la corrección de los grupos de autoridades devueltos por el método se van a definir tres índices que valorarán la calidad del resultado obtenido. Dichos índices son los siguientes:

- **Correctos:** Cantidad de elementos que aparecen correctamente agrupados.
- **Incorrectos:** Cantidad de elementos que aparecen en un grupo de forma incorrecta
- **No Agrupados:** Cantidad de elementos que no aparecen dentro un grupo, pero deberían hacerlo.

El primer valor, *Correctos*, mide la cantidad de elementos que se considera que se han agrupado correctamente, es decir, aquellos que constituyen una autoridad. Se entiende que un elemento es correcto cuando aparece dentro de un grupo (cluster) y debe estarlo. *Incorrectos*, hace referencia a aquellos elementos incorrectamente clasificados, es decir, aquellos que se han unido dentro de un grupo y no deberían hacerlo. Finalmente, *No agrupados*, representa aquellos elementos que aparecen aislados del grupo al que les corresponde pertenecer.

3 RESULTADOS

Durante esta sección daremos una serie de resultados obtenidos de la aplicación del método a datos obtenidos de distintas fuentes cada una de ellas con unas características determinadas, tal y como se detallará en la Sección 3.1. De entre todas las fuentes existentes, se han seleccionado tres: *CiteSeerX*, *DBLP* e *INSPEC*, debido a que cada una de ellas ofrece una interfaz única para todos sus contenidos, lo que permite realizar la extracción de datos detallada en la Sección 2.1.

3.1 DESCRIPCIÓN DE LAS FUENTES DE DATOS

CiteSeerX es una *Biblioteca Digital y Motor de Búsqueda de Literatura Científica* que cuenta con una interfaz muy completa a la hora de mostrar resultados, presentando para cada registro: *Título, Abstract, Número y lista de citas y Referencia bibliográfica*. De los datos que nos ofrece esta fuente, nosotros estamos interesados en dos elementos, la *referencia bibliográfica* en formato BibTeX y el *Abstract*, que son los que nos permitirán obtener toda la información necesaria para este proceso.

Otro ejemplo de Biblioteca Digital, y uno de los más populares, es DBLP. Este repositorio tiene la peculiaridad de estar indexado por autor. Dentro de un registro típico de esta Biblioteca, encontramos las publicaciones del autor ordenadas cronológicamente, mostrando primero las más recientes. Para cada publicación, se muestra una breve referencia y una serie de enlaces que permiten buscar el registro en otras fuentes de datos, como CiteSeerX. Del mismo modo, se ofrece una referencia bibliográfica completa del registro tanto en formato BibTex como en XML. Al igual que ocurre con CiteSeerX, pese a la indexación por nombre de autor, esta fuente presenta problemas de ambigüedad. Entre sus registros podemos encontrar entradas pertenecientes a distintos autores agrupados bajo un mismo nombre, así como distintas entradas para un mismo autor usando diferentes representaciones para su nombre. Considerando que esta fuente no nos ofrece el Abstract de la publicación, extraeremos de ella únicamente la *referencia bibliográfica* en formato BibTeX.

INSPEC, es un completo índice de material sobre física, tecnología eléctrica/electrónica, computación, ingeniería de control y tecnología de la información Al realizar una búsqueda en esta fuente, se muestra una lista de los resultados obtenidos, junto con un resumen de la información contenida acerca de la publicación. A la hora de extraer los datos, se nos ofrece la posibilidad de exportar los resultados de la búsqueda en diversos formatos, entre los que no se incluye el BibTeX, que estábamos usando hasta ahora. Sin embargo, permite la exportación en texto plano de los mismos campos que ofrece BibTeX, por lo que utilizamos dicha característica para obtener los datos deseados e introducirlos en nuestro base de datos.

3.2 COMPARACIÓN DE LAS DIVERSAS FUENTES DE DATOS CONSIDERADAS

Como hemos descrito en la sección anterior, existen diversas fuentes de datos presentando cada una de ellas la información con un tipo de representación diferente, así como distintos tipos de problemas relacionados con inconsistencias e incertidumbre que hacen difícil la búsqueda por autor en sus registros. Por lo tanto, partiendo de la base de la existencia de las diferentes fuentes, queremos ver cómo se comporta el método para cada una de ellas y cuales son las diferencias existentes en las autoridades obtenidas.

Para realizar esta comprobación, realizaremos la búsqueda de cuatro apellidos en español y cuatro en inglés, en cada una de estas fuentes: *Delgado, García, Sánchez, Vila, Brown, Davies, Smith y Williams*. Una vez obtenidos los conjuntos de datos, con tamaños que varían desde los pocos cientos de publicaciones hasta varios miles, les aplicamos el método descrito anteriormente y obtenemos los resultados que pueden verse en las Tablas 2 y 3. Considerando que no para todas las fuentes disponemos de *Abstract*, realizaremos esta prueba en dos etapas, primero compararemos los conjuntos de datos sin incluir información obtenida del Abstract, para posteriormente hacer un estudio de los mismos incluyéndolo en los casos en los que sea posible.

Tabla 2: Resultados de la aplicación del método sin incluir el Abstract.

Fuente	Apellido	Corr.	Inc.	N.A.
DBLP	Brown	89.3	9.37	1.33
	Delgado	96.09	2.61	1.3
	García	86.24	11.91	1.85
	Johnson	89.72	9.47	0.81
	Sánchez	83.99	14.19	1.82
	Smith	90.01	8.57	1.42
	Vila	97.69	1.54	0.77
	Williams	89.2	9.87	0.94
	Media	90.28	8.44	1.28
CiteSeerX	Brown	87.39	9.67	2.94
	Delgado	51.85	0	48.15
	García	88.71	6.61	4.68
	Johnson	87.06	7.52	5.43
	Sánchez	84.68	11.26	4.05
	Smith	83.26	10.12	6.61
	Vila	89.81	2.78	7.41
	Williams	86.22	9.39	4.38
	Media	82.37	7.17	10.46
INSPEC	Brown	91.44	0	8.56
	Delgado	89.92	8.06	2.02
	García	90.2	0	9.8
	Johnson	85.05	9.28	5.67
	Sánchez	85.71	11.76	2.52
	Smith	81.96	12.89	5.15
	Vila	75.72	0	24.27
	Williams	83.05	0	16.95
	Media	85.38	5.25	9.37

Observando la Tablas 2 y 3 puede verse que los porcentajes de aciertos son muy similares en las tres fuentes, siendo ligeramente superiores en media para los datos de *DBLP*, pese a no disponer de Abstract. Así mismo, puede verse que los datos de *INSPEC* muestran un porcentaje de errores muy bajo, aunque sí un número alto de no agrupados. Estos resultados son explicables si consideramos la naturaleza de los datos de entrada. Por norma general, los resulta-

Tabla 3: Resultados de la aplicación del método incluyendo el Abstract.

Fuente	Apellido	Corr.	Inc.	N. A.
CiteSeerX	Brown	89.26	9.68	1.05
	Delgado	82.01	6.88	11.11
	García	88.74	5.77	5.49
	Johnson	88.1	7.72	4.18
	Sánchez	84.68	11.26	4.05
	Smith	84.95	8.66	6.39
	Vila	89.81	2.78	7.41
	Williams	87.24	7.11	5.65
	Media	86.86	7.48	5.68
INSPEC	Brown	92.23	0	7.77
	Delgado	92.37	6.83	0.8
	García	90	0	10
	Johnson	89.64	0	10.36
	Sánchez	86.67	11.25	2.08
	Smith	89.29	0	10.71
	Vila	75.72	0	24.28
	Williams	83.05	0	16.95
	Media	87.37	2.26	10.37

dos obtenidos de *DBLP* contienen un nivel menor de ruido que los de *CiteSeerX* e *INSPEC*, existiendo menos problemas de representación que en estas dos últimas fuentes. El alto número de no agrupados que se obtienen en los datos extraídos de *INSPEC* parece ser causado por el hecho de existir más dispersión en estos datos, explicada como un número menor de trabajos por autor, así como trabajos con un número de coautores muy elevado. Sin embargo, pese a estas peculiaridades, vemos que la media de aciertos para cualquiera de los casos supera el 85 %, por lo que consideramos nuestro método como válido para estas tres fuentes de datos.

Con respecto a la inclusión o no de los *Abstract* (Tabla 3), vemos que los resultados que se obtienen son visiblemente mejores, pasando de un 82.37 % a un 86.86 % de aciertos en un caso y de un 85.38 % a un 87.37 % en el otro, a la vez que bajan los porcentajes de errores y no clasificados. Este aumento en el porcentaje de elementos correctos nos lleva a la conclusión de que efectivamente, el uso del abstract cuando está disponible, es beneficioso para el proceso, sobre todo en le caso de datos ruidosos.

4 CONCLUSIONES

En este documento hemos presentado un método que permite aplicar el control de autoridades a datos obtenidos de Bibliotecas Digitales en la Web. Para ello, permitimos el uso de distintas Bibliotecas en la Web, cada una de ellas con unas características y un tipo de información determinados. Partiendo de dicha información hemos realizado una serie de búsquedas obteniendo las autoridades deseadas y hemos comparado los resultados.

Los resultados nos han mostrado que el método que proponemos es eficaz para cualquier fuente de datos, independientemente de su procedencia. Sí bien hemos comprobado que existen pequeñas diferencias entre ellas, quedando el descubrimiento de sus causas como trabajo futuro. En cuanto la inclusión del Abstract como elemento a la hora de comparar palabras clave, se ha visto claro que resulta muy útil siempre y cuando esté disponible.

Referencias

[1] M. Delgado, A. F. Gómez-Skarmeta, and A. Vila. On the use of hierarchical clustering in fuzzy modeling. *International Journal of Approximate Reasoning*, 14(4):237 – 257, 1996.

[2] I. Diaz-Valenzuela, M.J. Martin-Bautista, and M.A. Vila. An automatic data mining authority control system: A first approach. In *Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on*, pages 569 –574, Cairo, Egypt, December 2010.

[3] Hui Han, Lee Giles, Hongyuan Zha, Cheng Li, and Kostas Tsioutsiouliklis. Two supervised learning approaches for name disambiguation in author citations. In *In JCDL '04: Proceedings of the 4th ACM/IEEE joint conference on Digital libraries*, pages 296–305, 2004.

[4] Hui Han, Hongyuan Zha, and C. Lee Giles. Name disambiguation in author citations using a k-way spectral clustering method. In *INTERNATIONAL CONFERENCE ON DIGITAL LIBRARIES*, pages 334–343. ACM, 2005.

[5] Cristina Herrero Pascual. El control de autoridades. *Anales de Documentación*, (2):121–136, 1999.

[6] Yoojin Hong, Byung won On, and Dongwon Lee. System support for name authority control. In *Problem in Digital Libraries: OpenDBLP Approach. In European Conf. on Digital Libraries (ECDL*, pages 134–144, 2004.

[7] James W. Warnner and Elizabeth W. Brown. Automated name authority control. In *JCDL '01: Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries*, pages 21–22, New York, NY, USA, 2001. ACM.

Recuperación de Preguntas Frecuentes a través de SMS

David Pinto[1], Darnes Vilariño[1], Saúl León[2], Esteban Castillo[3], Mireya Tovar[1]

[1] *Benemérita Universidad Autónoma de Puebla*, {dpinto,darnes,mtovar}@cs.buap.mx

[2] *Facultad de Ciencias de la Computación*, saul.ls@live.com

[3] *Maestría en Ciencias de la Computación*, ecjbuap@gmail.com

Resumen

En este trabajo se presenta un análisis del uso de técnicas de traducción automática para sistemas de recuperación de información monolingüe, crosslingüe y multilingüe. Abordamos el problema de búsqueda de respuestas asociadas a preguntas frecuentes (FAQ), considerando la consulta como un mensaje de texto corto llamado SMS. En particular, usamos el modelo de traducción IBM-5 para alinear en un corpus de entrenamiento los SMS y las preguntas frecuentes. El corpus usado en los experimentos contiene textos cortos que han sido escritos en Inglés, Hindú y Malayalam. La baja frecuencia de los términos, el tipo de vocabulario usado en los SMS y los diferentes idiomas da como resultado un problema que es ciertamente un reto importante en el área de recuperación de información. Los resultados obtenidos muestran que es posible proponer una técnica automática para normalizar el vocabulario de los SMS, con la finalidad de recuperar los FAQ asociados al SMS de entrada, obteniendo un resultado aceptable.

Palabras Clave: Recuperación de información, Normalización de consultas, FAQ, SMS.

1 INTRODUCCIÓN

En años recientes ha habido un gran crecimiento en servicios basados en dispositivos móviles, derivado principalmente de la masificación de los mismos. Uno de los mecanismos de comunicación usados en este tipo de dispositivos es el SMS, el cual básicamente es un mensaje textual de no más de 160 caracteres. El bajo costo para el envío de los SMS y el fácil acceso a la compra de dispositivos móviles ha hecho emerger a los mensajes instantáneos como el medio preferido de comunicación después del medio hablado y de los e-mails.

Desde el punto de vista de los sistemas de recuperación de información es importante explotar el potencial existente en el número de usuarios que hacen uso de este tipo de servicios. Por tanto, en este trabajo, se plantea tratar el problema de búsqueda de respuestas asociadas a preguntas frecuentes, usando un SMS como texto de consulta para el sistema de recuperación de información.

Es bien conocido que los millones de usuarios de mensajería instantánea, en particular de mensajes cortos, generan contenido electrónico en un lenguaje que no se adhiere a la gramática convencional, ni a las reglas de puntuación y de pronunciación usuales. Las palabras son intencionalmente comprimidas, usando abreviaciones, acrónimos, transliteraciones fonéticas e incluso neologismos. Adicionalmente, es importante notar que la falta del espacio de pantalla y el tamaño del teclado hace que a menudo los mensajes contengan errores tipográficos involuntarios y de ortografía. Lo anterior deja claro la dificultad existente en la construcción de un sistema de recuperación de información que contemple como texto de consulta (entrada) a un SMS, incluso para preguntas cuyas respuestas están bien documentadas como son las FAQ (Frequently Asked Questions).

En el caso de estudio que atañe a este trabajo de investigación, hemos considerado el uso de un corpus multilingüe que contiene textos escritos en tres diferentes idiomas: Inglés, Hindú y Malayalam. Por esta razón, hemos dividido los experimentos en tres: recuperación de información monolingüe, crosslingüe y multilingüe. El corpus usado fue obtenido del foro de evaluación de recuperación de información FIRE[i] fomentado por IBM Research India, y ha sido usado en esta investigación con el propósito de evaluar un método de normalización de mensajes cortos y su uso en la tarea de búsqueda de FAQ a partir de SMS.

Los avances obtenidos en este trabajo pueden ser de gran utilidad de manera particular en la India, en donde se estima que existen aproximadamente 400 millones de usuarios de telefonía móvil.

2 TRABAJOS PREVIOS

Existen algunos trabajos que han abordado el problema de la recuperación de preguntas frecuentes (FAQs). En particular, Harksoo Kim ha estudiado por varios años la tarea de recuperación de FAQs. En [6], por ejemplo, presenta una manera confiable de recuperar FAQs usando un agrupamiento de las bitácoras de consultas previas. En [2] mejora su trabajo previo, ahora usando semántica latente, y finalmente, en [3] enfatiza el trabajo usando los pesos de los términos latentes.

Por otro lado, en [5] se hace uso de métodos de traducción automática para alinear un corpus de FAQs (pregunta con su correspondiente respuesta), con la finalidad de construir un modelo estadístico de traducción que es utilizado posteriormente para expandir las consultas introducidas en un sistema de recuperación de información. Los experimentos se llevan a cabo considerando los idiomas de Inglés y Chino. Es interesante el hecho de usar métodos de traducción automática para generar un modelo de asociación de terminología entre el vocabulario de preguntas y el de respuestas, el cual puede posteriormente ser usado para estimar la probabilidad de una respuesta, dado un conjunto de términos existentes en la consulta del sistema de recuperación de información.

En [7], se presenta una aproximación a la recuperación de FAQs de dominio específico, usando lo que ellos denominan "aspectos independientes". Básicamente se trata de extraer términos y relaciones empleando WordNet y HowNet, para posteriormente usarlos en un modelo probabilístico basado en mixturas con la finalidad de interpretar la consulta y los pares consulta-respuesta, basándose en los aspectos independientes.

Un punto importante a remarcar, es que en ninguno de los trabajos anteriores se usan consultas tipo SMS, sino consultas escritas en texto normal, lo cual es una diferencia importante con respecto a los experimentos hechos en este trabajo de investigación. En [1], se trata de manera particular el problema de trabajar con textos tipo SMS. Se presenta en ese trabajo un modelo estadístico basado en frases para normalizar mensajes SMS. Se argumenta que los SMS son muy diferentes de los textos escritos normales, debido al estilo particular de escritura de éstos, y la alta frecuencia de ocurrencia de términos no estandarizados, usualmente en versiones cortas, resumidas, truncadas, o fonéticamente transliteradas.

Aun así, algunos trabajos han atacado el problema de usar SMS como mecanismo de consulta en la recuperación de FAQs. En [4] y [8], por ejemplo, se implementa un servicio web que permite consultar una base de datos de FAQs escritas en lenguaje Hindi usando consultas tipo SMS. La propuesta consiste en formular el criterio de similitud en el proceso de búsqueda como un problema combinatorio, en donde el espacio de búsqueda consiste de todas las combinaciones de las variaciones del diccionario de los términos de la consulta y sus N mejores traducciones. Lamentablemente el corpus usado en los experimentos no se encuentra disponible y por tanto, no es posible usarlo con fines de comparación.

3 FORMULACIÓN DEL PROBLEMA

En este trabajo de investigación se propone un modelo de recuperación de FAQs, usando como consultas a textos cortos tipo SMS (mensajes). Este problema tiene relevancia, pues como ya se ha comentado anteriormente, se necesita un proceso para normalizar las consultas. Adicionalmente, se debe abordar la problemática del cambio de idioma, pues se trabaja con tres tareas distintas: monolingüe, crosslingüe y multilingüe.

Para el caso monolingüe, se considera que tanto la consulta como las FAQs se encuentran escritas en el mismo idioma (Inglés, Hindi o Malayalam). Para el caso de la tarea crosslingüe, solamente se ha considerado que la consulta se encuentra escrita en Inglés y las FAQs en Hindi (por disponibilidad de corpus). Finalmente, en el caso multilingüe se considera que las consultas pueden estar escritas en cualquiera de los tres idiomas, al igual que las FAQs. Esta última es sin duda una de las tareas más complicadas.

Formalmente, el problema queda formulado de la siguiente manera: Sea ζ el conjunto de las preguntas en el corpus (FAQ) y $S=s_1s_2...s_n$ un SMS. Tanto el SMS como cada pregunta $q \in \zeta$, son vistos como una secuencia de términos. El objetivo es encontrar la pregunta q^*, del corpus ζ, que tengan el mayor grado de similitud con respecto al SMS *S*.

Para este propósito, hemos usado un modelo de recuperación de información basado en intersección de conjuntos de términos, sin embargo, hemos propuesto un modelo de normalización de consultas con la finalidad de observar el comportamiento de la técnica desarrollada, pero en contextos multilingües. A continuación se describe el modelo usado para normalizar la terminología de los SMS. Este modelo maneja información incompleta, con un cierto grado de incertidumbre, pues lo que ofrece el diccionario estadístico no es completamente exacto.

3.1. NORMALIZACIÓN DE LOS SMS

Los mensajes escritos en SMS contienen un alto grado de términos que no existen en los diccionarios normales. Es común, introducir nueva terminología que regularmente tiene que ver con las contracciones de las palabras originales o con una representación ortográfica de su correspondiente fonética. Es importante notar que dicha terminología cambia de acuerdo a los rangos de edad del usuario, ya que las personas jóvenes suelen tener mayor uso de representaciones fonéticas, lo cual puede estar derivado del hecho que ellos no dominan aún el vocabulario de su lenguaje nativo. Algunos ejemplos de los términos usados

y su correspondiente interpretación se muestran en la Tabla 1.

Tabla 1: Ejemplos de terminología usada en SMS.

Mensaje SMS	Interpretación
10q	Thank you
tna	Temporarily not available
afk	Away from keyboard
L8	Late
LOL	Laugh Out Loud
26Y4U	To sexy for you
@WRK	At work
⋮	⋮

Para la tarea que nos atañe, debemos considerar cada uno de los términos que se usan frecuentemente, pero también aquellos que no son tan frecuentes pero que podrían ocurrir. Tome por ejemplo el caso de una consulta del corpus de prueba: "whr can i find info abt pesticide estb reg and rep", la que podría ser interpretada como "where can i find information about pesticide establishment registration and reporting".

Con la finalidad de determinar el significado de aquellos términos escritos en la consulta tipo SMS que caen fuera del vocabulario común, se propone sustituir cada término de la consulta por la traducción más cercana ofrecida por medio de un diccionario estadístico bilingüe. Este diccionario ha sido creado mediante la herramienta Giza++, usando el modelo IBM-4, y entrenando el modelo de traducción usando los pares de frases, alineando los SMS y las preguntas correspondientes de las FAQs. En total, se crearon 13 diferentes diccionarios estadísticos bilingües, que corresponden a diferentes variantes consideradas en cada una de las tareas contempladas en este trabajo de investigación, de acuerdo a los datos de entrenamiento disponibles. El uso de los idiomas fuente y destino ocupados para cada tarea se muestra en la Tabla 2.

Para encontrar la similitud entre los términos de los mensajes (*SMS*) y cada una de las preguntas de las FAQs (P_{FAQ}), se aplicó la similitud de Jaccard, tal y como se muestra en la Ecuación (1).

$$Similitud(SMS, P_{FAQ}) = \frac{|\ SMS\ \cap\ P_{FAQ}\ |}{|\ SMS\ \cup\ P_{FAQ}\ |} \quad (1)$$

El pseudocódigo asociado al proceso de búsqueda de FAQs se muestra en el Algoritmo 1. Básicamente, se recibe como entrada el conjunto de consultas tipo SMS, las FAQs (ζ) sobre las cuales se desea hacer el proceso de recuperación de información y el diccionario estadístico bilingüe (ϕ) usado en el proceso de normalización de los textos SMS. Cada mensaje de consulta es normalizado de acuerdo al criterio explicado en la sección 1. El mensaje normalizado es entonces comparado con cada una de las preguntas asociadas a cada FAQ por medio del criterio de similitud usado.

Tabla 2: Distribución de los diccionarios estadísticos bilingües.

Tarea	Idioma Fuente (SMS)	Idioma Destino (Pregunta-FAQ)
Monolingüe	Inglés	Inglés
	Hindi	Hindi
	Malayalam	Malayalam
Crosslingüe	Inglés	Hindi
Multilingüe	Inglés	Inglés
	Inglés	Hindi
	Inglés	Malayalam
	Hindi	Hindi
	Hindi	Inglés
	Hindi	Malayalam
	Malayalam	Malayalam
	Malayalam	Inglés
	Malayalam	Hindi

En nuestro caso utilizamos el coeficiente de Jaccard, y todas aquellas preguntas FAQ (P_{FAQ}) que superen el valor de similitud *Umbral*, definido como la menor similitud de los *N* mejores valores de similitud, son agregados al conjunto respuesta.

Algorithm 1: Recuperación de FAQs a partir de SMS

Input: Consultas: $SMS = \{sms_1, sms_2, \ldots, sms_n\}$
Input: FAQs: $\zeta = \{q_1, \cdots, q_n\}$
Input: Diccionario estadístico bilingüe: $\phi = p(t_{SMS}, t_q)$
Output: N Mejores respuestas para cada SMS: Q

```
foreach sms_i ∈ SMS do
    smsN_i = Normaliza(sms_i, φ);
    Q[i] ← {0};
    foreach P_FAQ ∈ ζ do
        if Similitud(smsN_i, P_FAQ) > Umbral then
            Q[i] = Q[i] ∪ {P_FAQ, Similitud(smsN_i, P_FAQ)};
        end
    end
    if |Q[i]| > N then
        Q[i] = NMejoresValores(Q[i]);
    end
end
return Q
```

4 RESULTADOS EXPERIMENTALES

En esta sección se presentan los resultados obtenidos para cada una de las tareas evaluadas. Primeramente, se presenta una descripción de los corpora de entrenamiento y prueba usados en los experimentos. Posteriormente, se discuten los valores de evaluación encontrados.

4.1. CORPORA DE ENTRENAMIENTO

Los corpora de entrenamiento asocian un determinado SMS a uno o más FAQs. Los diferentes corpora de las FAQs se encuentran divididos de acuerdo al idioma. En la Tabla 3 se puede observar la cantidad de FAQs asociadas a cada idioma en los corpora de entrenamiento.

Tabla 3: Distribución de los SMS en los corpora de entrenamiento.

Idioma	# de FAQs
Inglés	7,251
Hindi	1,994
Malayalam	681

Los SMS usados como consulta se encuentran distribuidos de acuerdo al idioma y a la tarea, tal y como se puede ver en la Tabla 4.

Tabla 4: Distribución de los SMS en los corpora de entrenamiento.

Tarea	Idioma	# de SMS
Monolingüe	Inglés	560
	Hindi	144
	Malayalam	96
Crosslingüe	Inglés	232
Multilingüe	Inglés	232
	Hindi	146
	Malayalam	48

4.2. CORPORA DE PRUEBA

En los corpora de prueba, se manejan exactamente el mismo conjunto de FAQs que los presentados en la Tabla 3, sin embargo, el número de SMS varia, considerando un conjunto disjunto con respecto a aquellos considerados en la fase de entrenamiento. Los SMS usados como consulta en la fase de prueba se encuentran distribuidos de acuerdo al idioma y a la tarea, tal y como se puede ver en la Tabla 5.

Tabla 5: Distribución de los SMS en los corpora de prueba.

Tarea	Idioma	# de SMS
Monolingüe	Inglés	131
	Hindi	37
	Malayalam	24
Crosslingüe	Inglés	59
Multilingüe	Inglés	58
	Hindi	37
	Malayalam	12

4.3. EVALUACIÓN

Con el propósito de evaluar el método de normalización en la tarea de búsqueda de FAQs a partir de SMS, hemos considerado la exactitud del sistema de recuperación de información. Se usa una medida básica que determina como positiva, a una respuesta que se encuentre en los primeros *N* documentos (FAQs) regresados por el sistema (*N=5*). De manera alternativa, se considera una medida ponderada calculada como el cociente de la posición que ocupa el documento dentro de los *N* documentos devueltos (ranking), dividida entre el total de documentos correctos esperados.

Sea *RO* el número de respuestas correctas obtenidas por el sistema y *RE* el número de respuestas correctas esperadas para cada consulta (sms_i) a evaluar en el corpus de prueba SMS_t. La exactitud se calcula como se muestra en la Ecuación (2), y la exactitud ponderada como se muestra en la Ecuación (3).

$$\text{Exactitud} = \frac{1}{|SMS_t|} * \sum_{sms_i}^{|SMS_t|} (RO/RE) \quad (2)$$

$$\text{EPonderada} = \frac{1}{|SMS_t|} * \sum_{sms_i}^{|SMS_t|} \left(\frac{Ranking\,(RO)/N}{RE}\right) \quad (3)$$

4.4. RESULTADOS OBTENIDOS

En la Tabla 4 puede observarse la exactitud obtenida por el sistema para las diferentes tareas ejecutadas en los experimentos. Se aprecia una exactitud baja generalizada para el idioma Hindi, lo cual es una evidencia del uso constante de terminología fuera del vocabulario, que probablemente se deba a la gran cantidad de usuarios que parecen no manejar un vocabulario estándar del idioma (clase media y baja).

El caso del idioma Malayalam es interesante, dado que fue el que obtuvo el mayor valor de exactitud. Este valor puede ser un indicativo del bajo uso de términos fuera del vocabulario y una constancia en el uso de vocabulario estándar. De hecho, no es raro, dado que este idioma se

puede considerar una forma separada del Tamil, un idioma usado en la educación y la administración; además de que el Malayalam ha tomado miles de sustantivos, cientos de verbos y algunos indeclinables del sánscrito, un idioma considerado aristocrático y propio de estudiosos, tal como ocurría con el latín en los países europeos.

Tabla 6: Resultados Obtenidos en la Evaluación de las diferentes subtareas.

Sub-Tarea	FAQS	Respuestas Esperadas	Respuestas Obtenidas	Exactitud	Exactitud Ponderada
EN-Monolingüe	7251	131	72	0.550	0.474
HI-Monolingüe	1994	37	13	0.351	0.297
MA-Monolingüe	681	24	20	0.833	0.825
EN-HI (Crosslingüe)	7251	59	21	0.356	0.281
EN-Multilingüe	9926	58	29	0.500	0.500
HI-Multilingüe	9926	37	10	0.270	0.183
MA-Multilingüe	9926	12	4	0.333	0.316

Finalmente, para el idioma Inglés se tiene una exactitud promedio del 50%, observando un uso generalizado de contracciones en los SMS, heredadas de otros países de habla inglesa. Se evaluó también el rendimiento del sistema usando el vocabulario original de los SMS, validando, como ya es conocido en la literatura, que es conveniente eliminar las palabras cerradas para mejorar la exactitud del sistema de recuperación de información.

5 CONCLUCIONES Y TRABAJO A FUTURO

En este trabajo hemos abordado el problema de búsqueda de FAQs a partir de textos tipo SMS. Se evaluaron las tareas monolingüe, crosslingüe y multilingüe para los idiomas Inglés, Hindi y Malayalam. La propuesta de normalización hecha en este trabajo puede, sin embargo, extenderse a cualquier otro idioma si es que se posee un corpus alineado de SMS con sus correspondientes respuestas FAQ.

Los resultados obtenidos muestran que es posible normalizar los textos SMS para obtener un valor de exactitud aceptable. No hemos introducido valores *baseline* para cada una de las tareas, sin embargo, una evaluación adicional realizada, muestra que para el caso monolingüe, el valor de exactitud obtenido cuando no se normalizan los textos es de 0.081 para el idioma Inglés, de 0 para el Hindi, y de 0.241 para el Malayalam. Estos valores refuerzan la discusión presentada anteriormente con respecto al uso correcto del vocabulario en ciertos idiomas.

Como trabajo a futuro, consideramos importante, evaluar el rendimiento con otros modelos de recuperación de información y esquemas de similitud que consideren la frecuencia de ocurrencia de los términos. Adicionalmente, nos gustaría evaluar el comportamiento de la propuesta con otros idiomas y un mayor número de documentos en los corpora de evaluación.

Agradecimientos

Este trabajo ha sido apoyado parcialmente por los proyectos: CONACYT #106625, VIEP #VIAD-ING11-I y #PIAD-ING11-I.

Referencias

[1] AiTi Aw, Min Zhang, Juan Xiao, and Jian Su. A phrase-based statistical model for sms text normalization. In *Proceedings of the COLING/ACL on Main conference poster sessions*, COLING-ACL '06, pages 33–40, Stroudsburg, PA, USA, 2006. Association for Computational Linguistics..

[2] Chung-Hsien Wu, Jui-Feng Yeh, and Ming-Jun Chen. Domain-specific faq retrieval using independent aspects. *ACM Transactions on Asian Language Information Processing (TALIP)*, 4:1–17, March 2005.

[3] Danish Contractor, Govind Kothari, Tanveer A. Faruquie, L. Venkata Subramaniam, and Sumit Negi. Handling noisy queries in cross language faq retrieval. *In Proceedings of the 2010 Conference on EmpiricalMethods in Natural Language Processing*, EMNLP '10, pages 87–96, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[4] Govind Kothari, Sumit Negi, Tanveer A. Faruquie, Venkatesan T. Chakaravarthy, and L. Venkata Subramaniam. SMS based interface for FAQ retrieval. *In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL-IJCNLP '09, pages 852–860, Morristown, NJ, USA, 2009. Association for Computational Linguistics.

[5] Harksoo Kim, Hyunjung Lee, and Jungyun Seo. A reliable faq retrieval system using a query log classification technique based on atent semantic analysis. *Inf. Process. Manage.*, 43:420–430, March 2007.

[6] Harksoo Kim and Jungyun Seo. High-performance faq retrieval using an automatic clustering method of query logs. *Inf. Process. Manage.*, 42:650–661, May 2006.

[7] Harksoo Kim and Jungyun Seo. Cluster-based faq retrieval using latent term weights. *IEEE Intelligent Systems*, 23:58–65, March 2008.

[8] Stefan Riezler, Alexander Vasserman, Ioannis Tsochantaridis, Vibhu Mittal, and Yi Liu. Statistical machine translation for query expansion in answer re-

trieval. *In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 464–471, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

MECANISMO PARA LA EXTRACCIÓN Y ALMACENAMIENTO DE GRADOS BORROSOS A PARTIR DE LAS RELACIONES SEMÁNTICAS DE WORDNET

Francis C. Fernández-Reyes[1], José A. Olivas [2], Exiquio C. Leyva [1], Rogelio Lau [1]

[1] *Instituto Superior Politécnico "José Antonio Echeverría", La Habana, Cuba*, {ffernandez, exiquio, lau}@ceis.cujae.edu.cu

[2] *Universidad de Castilla – La Mancha, España*, JoseAngel.Olivas@uclm.es

Resumen

En este trabajo se presenta un mecanismo automatizado para la extracción de grados borrosos a partir de las principales relaciones semánticas encontradas en la base de datos léxica WordNet. De igual manera, se describen las posibles formas de almacenamiento para su posterior uso. El resultado aquí obtenido servirá para enriquecer los modelos de representación de documentos desde un punto de vista semántico.

Palabras Clave: WordNet, Lógica Borrosa, Recuperación de Información, Modelos de Representación de documentos.

1 INTRODUCCIÓN

La base de datos léxica WordNet [1], creada en la Universidad de Princeton – así como todas las WordNets surgidas a partir de ella [2]– son recursos estándares utilizados en aplicaciones que procesan en alguna medida el lenguaje natural. Esta base de datos léxica se construye sobre la base de las categorías gramaticales de sustantivo, verbo, adjetivo y adverbio y las relaciones semánticas de hiponimia, hiperonimia, meronimia, holonimia, sinonimia, antonimia y términos coordinados. Las relaciones de hiponimia e hiperonimia expresan relaciones "tipo de", es decir, dados dos términos: *árbol* y *pino*; *pino* es un "tipo de" *árbol*, por lo que *pino* es hipónimo de *árbol*, mientras que, *árbol* es hiperónimo de *pino*. Los términos coordinados se basan en la hiperonimia, se dice que dos términos son coordinados siempre que compartan un hiperónimo común, por ejemplo, *pino* y *roble* que comparten árbol como hiperónimo común. Las relaciones de meronimia y holonimia expresan relaciones "parte de", es decir, dados dos términos: *automóvil* y *rueda*; *rueda* es "parte de" *automóvil*, por lo que *rueda* es merónimo de *automóvil*, mientras que, *automóvil* es holónimo de rueda.

En WordNet los conceptos son representados por la lista de palabras que pueden ser usadas para expresarlo, es decir, por todos los términos sinónimos entre sí. Los grados de sinonimia entre dos términos, en algunas propuestas son calculados haciendo uso de sus sentidos obtenidos de un diccionario, la función de pertenencia es calculada mediante el número de sinónimos que comparten entre sí [3] [4].

Otros trabajos tratan la sinonimia como una relación usualmente simétrica y transitiva. Una vía para interpretar los grados de sinonimia es la utilización de la distancia euclidea bajo el supuesto que los dos enlaces de la sinonimia se encuentran en ángulo uno a la derecha de otro sobre un plano. La vía que utilizan para calcular estos grados es la anotación de cada *synset* con algunas medidas estadísticas que reflejan la composición de estos [5].

Las relaciones de hiperonimia y meronimia, así como sus relaciones inversas, aparecen en más de una categoría gramatical y comparten su significado entre varios *synsets*. El fenómeno ocurre cuando el significado de una palabra se incluye en el significado de otra pero esta inclusión no es total lo que incentiva su estudio para determinar en qué grado ocurre el fenómeno dentro de un contexto.

El presente trabajo expone un mecanismo para extraer grados borrosos a partir de determinadas relaciones semánticas que se obtienen de la base de datos léxica WordNet. Siguiendo este propósito se divide en dos secciones, en la primera se expondrán las funciones de pertenencia de cada *synset* asociadas a las diferentes relaciones semánticas utilizadas. Mientras que en la segunda, se expondrán las características de las dos fuentes de almacenamiento propuestas. Finalmente, se arriban a conclusiones y recomendaciones para trabajos futuros.

2 EXTRACCIÓN DE GRADOS BORROSOS A PARTIR DE LAS PRINCIPALES RELACIONES DE WORDNET

La lógica difusa permite tratar información imprecisa, en términos de conjuntos borrosos, debido a la falta de definición del concepto al que se aplica. Es una lógica que utiliza expresiones que no son ni totalmente verdaderas ni completamente falsas, es decir, es una lógica aplicada a conceptos que pueden tomar un valor cualquiera de veracidad dentro de un conjunto de valores que abarcan el rango entre la verdad absoluta y la falsedad total.

Cada uno de los *synsets*, definidos en WordNet, pertenecen en determinada medida a los conjuntos definidos por cada una de las relaciones semánticas que aparecen en esta base de datos léxica. Por ejemplo, entre cada par de *synsets*, las relación de sinonimia puede ser calculada a través de la intersección entre los conjuntos de palabras que son expresadas a través de estos, tal y como aparece esquematizado en la Figura 1. Aquí se pueden apreciar resaltadas las palabras que comparten en común un *synset* asociado a *marvelous* con uno asociado a *grand*.

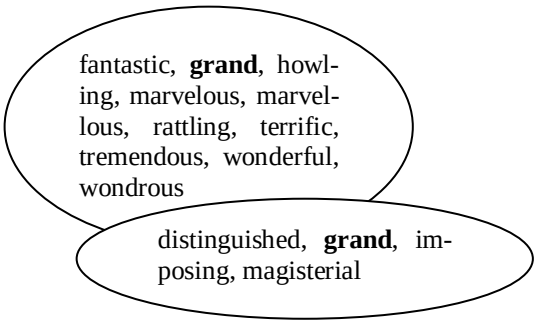


Figura 1: Conjunto de palabras que son expresadas con dos *synsets* distintos.

Existen medidas de distancia semántica en la literatura [6] que no establecen los grados de pertenencia entre dos pares de *synsets* a través de la concepción del conjunto de palabras o de relaciones que comparten entre sí. Sin embargo, esta distancia es calculada mediante la longitud del camino que separa estos *synsets* a través de relaciones de hiperonimia o incluso es calculada mediante el uso del solapamiento de sus glosas.

2.1. RELACIÓN DE SINONIMIA

Los grados de sinonimia se calcularán entre dos *synsets* y no entre dos palabras. Un *synset* es un conjunto de palabras que son sinónimas entre sí para un sentido en particular. Los grados de sinonimia se calcularán de la siguiente forma:

$$SD^1(s_1, s_2) = \frac{|s_1 \cap s_2|}{|s_1 \cup s_2|} \quad (1)$$

La fórmula (1) se calculará como el número de palabras en común entre los dos *synsets* (s_1, s_2) dividido por el total de palabras que comparten entre sí. Siempre y cuando existan entre un par de *synsets* al menos una palabra en común, se puede obtener un grado de pertenencia mayor que 0. Si no comparten palabras en común sería 0 y se concluye que los *synsets* no se comportan como sinónimos entre sí. La sinonimia se cuantifica en ambos sentidos.

Por ejemplo la palabra *unit* tiene 7 *synsets* asociados o 7 sentidos posibles: {*unit_of_measurement, unit*}, {*unit*}, {*unit, social_unit*}, {*unit*}, {*system, unit*}, {*unit, building_block*}, {*whole, whole_thing, unit*};

SD^1 ({*unit_of_measurement, unit*}, {*unit, social_unit*}) = 1 / 3 = 0.33.

SD^1 ({*unit*}, {*unit*}) = 1 / 1 = 1.

Aunque estos dos *synsets* contienen la misma palabra, tienen dos glosas definidas independientes pero relacionadas. Por lo tanto, no basta con definir la sinonimia en cuanto a las palabras que comparten significado dentro de él, sino hay que distinguir los *synsets* de acuerdo a la glosa que expresan y que los hace únicos.

Debido a lo anterior se definió otra forma de calcular los grados de sinonimia entre dos *synsets* utilizando el solapamiento de las glosas que cada uno de ellos representa. De esta forma se aplicó la Ley de Zipf con el objetivo de lograr medidas mayores en caso que el solapamiento se establezca con cadenas de términos en forma consecutiva. Banerjee [7] propuso el cálculo de un incremento cuadrático de acuerdo al número de términos consecutivos que se solapan. Nuestra propuesta está definida de acuerdo a la prueba exacta de Fisher [8]. De esta forma se puede analizar si dos variables dicotómicas están asociadas cuando la muestra a estudiar es demasiado pequeña. Los solapamientos pueden ser a lo sumo el tamaño máximo de una glosa en particular. Sin embargo esto no es totalmente cierto pues no existen *synsets* repetidos en WordNet. La prueba inversa queda definida por la fórmula del cálculo de los grados de sinonimia entre glosas:

$$SD^2(s_1, s_2) = \frac{e^{2|gloss(s_1) \cap gloss(s_2)|} - 1}{e^{2|gloss(s_1) \cap gloss(s_2)|} + 1} \quad (2)$$

El tamaño de la intersección de las glosas se calcula de acuerdo a la suma del incremento cuadrático de cada cadena de términos contenidas en ambas glosas simultáneamente.

Otra forma para el cálculo de los grados de sinonimia, tomando como precedente los estudios del trabajo [3] [4, 9], sería la pertenencia de una palabra particular a un determinado *synset*.

$$SD^3(s_1, w) = \frac{1}{|s_1|} \quad (3)$$

La fórmula (3) calcula el inverso del número de palabras representadas por el *synset* (s_1) donde w pertenece. Esta

fórmula se aplica solo si la palabra en cuestión aparece dentro del *synset*, en caso contrario es 0. Para el ejemplo anterior, si se toma la palabra *unit* entonces los grados de sinonimia para los synsets {*unit_of_measurement*, *unit*} y {*unit*, *social_unit*} sería el mismo ½=0.5. Esta variante tiene como desventaja que no se observa directamente la relación entre los *synsets*.

2.2. RELACIÓN DE HIPERONOMIA

La hiperonimia o cuasisinonimia es una relación de inclusión de significados. Mientras menor sea el número de hipónimos de un hiperónimo, mayor será el número de sentidos específicos de éstos y, por tanto, mayor su precisión de significado.

La hiponimia es una implicación unilateral: un hipónimo supone su hiperónimo, pero no viceversa y los cohipónimos no se suponen entre sí y no son sinónimos los unos de los otros. *Roble* es hipónimo de *Árbol*, lo que evidencia la relación unilateral, pero *Árbol* tiene más de un hipónimo asociado, la hiperonimia es una relación multilateral, por ejemplo: *Árbol* es hiperónimo no sólo de *Roble*, sino también y en la misma medida de *Abedul*, *Acacia*, etc.

Los grados de hiperonimia se calcularán entre dos *synsets* ubicados en la misma rama del árbol de herencia producido por esta relación semántica.

$$HypD(s_1, s_2) = \left[\sum_{s:hiperonimo(s_1)}^{s_2} 1 \right]^{-1} \quad (4)$$

La fórmula (4) calcula el grado de hiperonimia entre dos *synsets* como el inverso de la distancia encontrada entre ambos a través de esta relación. El cálculo se establece en un solo sentido, es decir, $HypD(s_1, s_2) \neq HypD(s_2, s_1)$ pues si el s_1 es hiperónimo de s_2, entonces s_2 es hipónimo de s_1 y puede compartir la hiponimia entre otros *synsets*.

En la Figura 2 aparece el *synset* {*Rose*} con todos sus hiperónimos asociados, hasta terminar en el tope de la jerarquía de hiperonimia. Para este *synset* es calculado su grado de pertenecía a esta relación en particular.

{Rose} @-› {shrub, bush} @-› {woody_plant, ligneous_plant} @-› {vascular_plant, tracheophyte} @-› {plant, flora, plant_life} @-› {organism, being, living_thing} @-› {entity}

Figura 2: Hiperónimos del *synset* {Rose}

A partir de lo representado en la Figura 1 se pueden calcular los grados de hiperonimia haciendo uso de la fórmula (4), quedando de la siguiente forma:

HypD ({*Rose*}, {*vascular_plant*, *tracheophyte*}) = 1 / 3 = 0.33.

HypD ({*Rose*}, {*plant*, *flora*, *plant_life*}) = 1 / 4 = 0.25.

HypD ({*Rose*}, {*organism*, *being*, *living_thing*}) = 1 / 5 = 0.2.

HypD ({*Rose*}, {*entity*}) = 1 / 6 = 0.17.

Los grados de hiponimia se calculan de acuerdo a la cantidad de hipónimos que comparten un hiperónimo en común de manera directa, lo que para la lingüística significa que los términos son co-hipónimos.

$$HipD(s_1, s_2) = \left[\sum_{s:hiperonimo(s_1)}^{s_2} hiponimos(s) \right]^{-1} \quad (5)$$

Por ejemplo el *synset* {*Rose*} tiene como hiperónimo inmediato a {*shrub*, *bush*}, a su vez los hipónimos inmediatos de éste son en total 302.

HipD ({Rose}, {*shrub*, *bush*}) = 1 / 302 = 0.003.

2.3. RELACIÓN DE MERONIMIA

Las relaciones de meronimia y holonimia se calculan de forma similar que las relaciones de hiperonimia y hiponimia. Estas forman un esquema de herencia similar al visto anteriormente, sólo que semánticamente no son equivalentes y la herencia se puede establecer a más de un *synset* simultáneamente. En el caso de este tipo de relación se va a almacenar de acuerdo a la sub-clasificación establecida por WordNet en *Parte* (p), *Sustancia* (s), *Miembro* (m) para cada una de estas relaciones (Mero/Holonimia). En esta base de datos léxica las relaciones de meronimia son etiquetadas con el carácter % y las de holonimia con #. Por ello, para hacer referencia a un merónimo-parte se utiliza la etiqueta %p, y así sucesivamente para las distintas combinaciones de las relaciones con sus clasificaciones.

$$MerD^1(s_1, s_2) = \left[\sum_{s \in \Omega} meronimos(s) \right]^{-1} \quad (6)$$

$$HolD^1(s_1, s_2) = \left[\sum_{s \in \Psi} holonimos(s) \right]^{-1} \quad (7)$$

Donde Ω es el conjunto de todos los *synsets* que trazan un camino (de cualquier longitud) entre s_1 y s_2 a través de relaciones de meronimia y Ψ es el conjunto que representa todos los *synsets* que descubren el camino a través de las relaciones de holonimia.

El *synset {wheeled_vehicle}* es merónimo de los *synsets {axle}, {brake}, {splasher}, {wheel}*. A su vez el *synset {wheel}* es merónimo de los *synsets {felloe, felly}* y *{ rim}*, tal como se aparece en la Figura 2.

{wheeled_vehicle} %p-› {axle}, {brake}, {splasher}, {wheel}

{wheel} %p-› {felloe, felly}, {rim}

Figura 3: Meronimia del *synset* {wheeled_vehicle}

A partir de lo representado en la Figura 3 se pueden calcular los grados de meronimia haciendo uso de la fórmula (6), quedando de la siguiente forma:

MerD^1 ({*wheeled_vehicle* },{*rim*})= 1/6= 0.17

MerD^1 ({*wheeled_vehicle* },{*brake*})= 1/4= 0.25

2.4. RELACIONES COMPLEJAS TRANSITIVAS EN SUSTANTIVOS

La holonimia y la hiponimia se cruzan en ocasiones de forma compleja. Si s_1 es holónimo en cualquiera de sus clasificaciones (#p, #m, #s) de s_n y éste a su vez es hipónimo de s_2 entonces s_1 es holónimo de s_2 y de todos sus hiperónimos. El *synset* s_n solamente es utilizado como conexión intermedia para descubrir la relación de holonima que indirectamente existe entre s_1 y s_2.

Se propone primeramente intentar calcular los grados de holonimia entre un par de *synsets* cualesquiera calculando el inverso de la distancia que tienen en el esquema de herencia de esta relación (utilizando fórmula 7). Si los *synsets* no aparecen relacionados en éste, entonces se propone buscar algún hipónimo intermedio entre ambos *synsets.*

Una buena práctica sería definir este caso particular de Holonimia integrando la información contenida en la fórmula 5, para el cálculo de la pertenencia a la Hiperonimia, con la fórmula 7 referida a la Holonima común. Por tanto, se debe definir la Holonimia indirecta en función de:

$$HolD^2(s_1,s_2) \Rightarrow HolD^1(s_1,s_n) \wedge HipD(s_n,s_2) \quad (8)$$

Suponiendo que s_1 está conectado de forma directa con s_n (solo se calcula $HolD^2$ bajo este supuesto), entonces el conjunto Ψ está compuesto únicamente por s_1, s_n y la sumatoria se reduce a la suma de los holónimos de s_1 con los de s_n. De esta forma se define la fórmula 9. La fórmula 10 se obtiene sustituyendo en (4):

$$HolD^1(s_1,s_n) = \frac{1}{holonimos(s_1) + holonimos(s_n)} \quad (9)$$

$$HipD(s_n,s_2) = \left[\sum_{s:hiperonimo(s_n)}^{s_2} hiponimos(s) \right]^{-1} \quad (10)$$

Finalmente, se define la $HolD^2(s_1,s_2)$ de acuerdo a 8, 9 y 10:

$$HolD^2(s_1,s_2) = \min \left(HolD^1(s_1,s_n), HipD(s_n,s_2) \right) \quad (11)$$

La meronimia se comporta de forma análoga, si s_1 es merónimo en cualquiera de sus clasificaciones (%p, %m, %s) de s_n y éste a su vez es hipónimo de s_2, entonces s_1 es merónimo de s_2. Igualmente se integra la información de la fórmula 6 con la fórmula 4, definiendo la meronimia indirecta en función de:

$$MerD^2(s_1,s_2) = MerD^1(s_1,s_n) \wedge HipD(s_n,s_2) \quad (12)$$

Suponiendo que s_1 está conectado de forma directa con s_n (solo se calcula $MerD^1$ bajo este supuesto), entonces el conjunto Ω está compuesto únicamente por s_1 y s_n, la sumatoria se reduce a la suma de los merónimos de cada *synset*. De esta forma se define la fórmula 13:

$$MerD^1(s_1,s_n) = \frac{1}{meronimos(s_1) + meronimos(s_n)} \quad (13)$$

Finalmente, se define la $MerD^2(s_1,s_2)$ de acuerdo a 10, 12 y 13:

$$MerD^2(s_1,s_2) = \min \left(MerD^1(s_1,s_n), HipD(s_n,s_2) \right) \quad (14)$$

El *synset {wheel}* es holónimo del *synset {wheeled_vehicle},* tal como aparece reflejado en la Figura 4.

{wheel} #p-› {wheeled_vehicle} @-› {vehicle} @-› {conveyance, transport} @-› {instrumentality, instrumentation} @-› {artifact, artefact} @-› {whole, unit} @-› {object, physical_object} @-› {physical_entity} @-› {entity}

Figura 4: Conexión de la meronimia con la hiperonimia.

De acuerdo a las fórmulas planteadas anteriormente en las secciones 2.2 y 2.3, entre *{wheel}* y *{vehicle}* no existía relación aparente, sin embargo a través de *{wheeled_vehicle}* aparece una relación que debe ser considerada. El *synset* {vehicle} tiene 8 hipónimos relacionados.

HolD^2 ({wheel}, {vehicle}) = min (1, 1/8) = 1/8

HolD^2 ({wheel}, {conveyance, transport}) = min (1, 1/22) = 1/22

HolD^2 ({wheel}, {instrumentality, instrumentation}) = min (1, 1/36) = 1/36

Otra relación compleja que puede aparecer con la meronimia es la siguiente: Si s_1 es hipónimo de s_2 y s_1 es merónimo de s_3, entonces s_2 es merónimo de s_3 en la misma medida que lo es s_1.

Un ejemplo de esta relación lo constituye el primer *synset* de *{house}* como se aprecia en la Figura 5. Los cálculos de los grados de pertenencia a la meronimia aparecen a continuación, se decidió establecer la interconexión a través del *synset {library},* pero se puede proceder de igual manera para cada uno de los merónimos asociados.

{house} %p-› {library}, {loft, attic, garret}, {porch}, {study}

{house} ~-› {beach_house}, {boarding_house, boardinghouse}, {bungalow, cottage}, {cabin}, {chalet}, {chapterhouse, fraternity_house, frat_house} …

Figura 5: Relación entre la meronimia y la hiponimia.

MerD^{3} ({*beach_house*}, {*library*}) = **MerD**^{1} ({*house*}, {*library*}) = 1/4 = 0.25

3 ALMACENAMIENTO DE GRADOS BORROSOS OBTENIDOS A PARTIR DE LAS PRINCIPALES RELACIONES DE WORDNET

Se proponen dos formas de almacenamiento independientes una de otra. Se pueden salvar los grados borrosos en ficheros textos, similares a los ficheros de datos que ofrece WordNet. También se pueden almacenar en un script de base de datos que es generado una vez completada la operación solicitada.

Actualmente, la aplicación informática encargada de calcular estos grados de forma automática para cada una de las relaciones semánticas se encuentra en desarrollo, la implementación está siendo realizada en lenguaje Java y el gestor de base de datos utilizado es PostGresSQL.

3.1. FICHEROS TEXTO

Se propone un único fichero donde se representarán por cada *synset* todas las relaciones semánticas acompañadas del correspondiente grado borroso de asociación. Cada línea de este fichero tendrá el siguiente formato:

(i) *synset_offset – PoS*: El *synset_offset* representa la etiqueta de WordNet asociada al sentido y *PoS* representa la categoría gramatical de este sentido.

(ii) *rel_count*: cantidad de relaciones semánticas representadas con su respectivo grado de asociación, aparecerán tantos *ptr_rel* como *rel_count* hayan.

(iii) *ptr_rel*: Cada uno de los *ptr_rel* tiene el formato [*rel grado synset_offset PoS*] donde *rel* especifica la relación semántica de acuerdo a la especificación de WordNet, *grado* especifica el grado borroso de dicha relación entre el *synset_offset* que aparece en (i) y el *synset_offset* que aparece en el *ptr_rel* particular, *PoS* es la categoría gramatical del *synset_offset* que aparece en *ptr_rel*.

Para encontrar un determinado *synset_offset* se implementa una búsqueda binaria que permite agilizar los procesos de búsqueda, debido a que existe un ordenamiento de las etiquetas de los sentidos.

3.2. ESQUEMA DE BASE DE DATOS

En la Figura 6 se aprecia que el modelo de la base de datos tiene dos tablas fundamentales: *tbl_synset* y *tbl_fuzzy_degree*. La primera representa los datos de todos los *synsets* de WordNet y la segunda es la que contiene los grados borrosos de acuerdo a cada una de las relaciones semánticas encontradas. Existe también una tabla *tbl_rel* que funciona como nomenclador para cada una de las relaciones semánticas analizadas. Para que la aplicación sea extensible a todas las versiones de WordNet creadas por Princeton, se incluyó también un nomenclador para manejar las versiones –*tbl-version.*

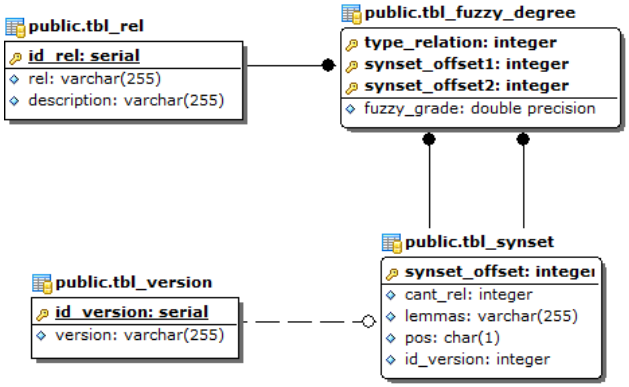


Figura 6: Esquema de base de datos que representa los grados borrosos de los distintos synsets de WordNet.

Después de gestionar cada uno de los procesos para el cálculo de los grados borrosos, estos se almacenan automáticamente en ambos medios (base de datos y fichero). Una vez que la base de datos se haya llenado se genera un script para utilizarla en varios gestores de datos.

4 CONCLUSIONES Y TRABAJOS FUTUROS

El mecanismo para la extracción y almacenamiento de grados borrosos a partir de las principales relaciones léxicas que aparecen en WordNet constituye un primer paso para la incorporación de información semántica a los modelos de representación de documentos. El mecanismo aquí propuesto permite analizar las relaciones que se establecen, tanto directa como indirectamente, entre los diferentes *synsets* a través de diversas relaciones.

Como posibles desarrollos futuros se podrían destacar los siguientes:

Generar un mecanismo para representar las principales relaciones obtenidas de adjetivos, verbos y adverbios. Atender las relaciones de oposición, fundamentalmente la antonimia que puede resultar útil para descubrir el significado de una palabra y es una relación que aparece en más de una categoría gramatical.

Incorporar a las funciones que calculan los grados de pertenencia en cada relación semántica, alguna variable que indique la utilización de los dominios en cada uno de los *synsets*.

Incorporar al cálculo de los grados de pertenencia otras fuentes de información semántica proveniente de ontologías dedicadas al dominio de la lingüística.

Agradecimientos

Esta investigación está parcialmente financiada por los proyectos FIDELIO (TIN2010-20395), MEC-FEDER, SCAIWEB2 (PEIC09-0196-3018) y PLINIO (POII10-0133-3516) de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1] G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross, y K. Miller, "Introduction to WordNet: An on-line lexical database". International Journal of Lexicography.Vol. **3**, No., pp. 235-244. 1990.

[2] P. Vossen, "EuroWordNet: Building a Multilingual WordNet Database with Semantic Relations between Words". Procesamiento del lenguaje natural.Vol. **18**, No., pp. 145-158. 1996.

[3] P. Garcés, J. Olivas, y F. Romero, "Concept-matching IR systems versus Word-matching IR systems: Considering fuzzy interrelations for indexing web pages". Journal of the American Society for Information Science and Technology JASIST.Vol. **4**, No. 57, pp. 564-576. 2006.

[4] J. Olivas, P. Garcés, J. Mata, F. Romero, y J. Serrano-Guerrero, "Conceptual Soft-Computing based Web search: FIS-CRM, FISS Metasearcher and GUMSe Architecture", en "Forging New Frontiers: Fuzzy Pioneers II, Studies in Fuzziness and Soft Computing", M. Nikravesh, J. Kacprzyk, and L. Zadeh, Editors. 2008, Springer Verlag. pp. 107 - 134.

[5] L. Borin y M. Forsberg. "From the People's Synonym Dictionary to fuzzy synsets – first steps". en *Workshop: Semantic Relations. Theory and Applications. Seventh Conference on International Language Resources and Evaluation (LREC'10)*. República de Malta. 2010

[6] T. Pedersen, S. Patwardhan, y J. Michelizzi, *WordNet::Similarity - Measuring the Relatedness of Concepts*, in *Nineteenth National Conference on Artificial Intelligence (AAAI-04)*. 2004: San Jose, CA. p. 1024-1025.

[7] S. Banerjee y T. Pedersen. "An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet". en *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*. 2002

[8] I. StatSoft, *Electronic Statistics Textbook*, Tulsa, Editor. 2011.

[9] M. Nikravesh, J. Kacprzyk, L. Zadeh, J. Olivas, P. Garcés, J. Mata, F. Romero, y J. Serrano-Guerrero, "Conceptual Soft-Computing based Web search: FIS-CRM, FISS Metasearcher and GUMSe Architecture", en "Forging New Frontiers: Fuzzy Pioneers II". 2008, Springer Berlin / Heidelberg. pp. 107 - 134.

Técnicas borrosas de clasificación y clustering de documentos aplicadas a la anotación semántica de contenidos web

Francisco P. Romero[1], **Jose A. Olivas**[1], **Jesús Serrano-Guerrero**[1], **Mateus Ferreira-Satler**[1], **Andres Soto**[2]

[1]*Depto. Tecnologías y Sistemas de la Información, Universidad de Castilla-La Mancha*,
{FranciscoP.Romero,JoseAngel.Olivas, Jesus.Serrano, Mateus.Ferreira}@uclm.es
[2]*Depto. de Ciencias de la Computación, Universidad Autónoma del Carmen, Campeche, México*,
soto_andres@yahoo.com

Resumen

La anotación semántica de contenidos web permite la mejora del rendimiento de los sistemas de recuperación y gestión de la información. La realización manual de ésta resulta ser un proceso totalmente ineficiente tanto en rapidez como en la corrección de los resultados. Es por ello, que se utilizan generalmente métodos automáticos o semiautomáticos de anotación que permiten que el proceso sea mucho más rápido, fácil y correcto. En este trabajo se presentan dos métodos automáticos de anotación semántica de contenidos web basados en algoritmos borrosos de clasificación y clustering de documentos. Los resultados obtenidos dentro de la anotación semántica de contenidos web de un dominio específico demuestran la viabilidad de las propuestas y la aportación de las técnicas borrosas en el manejo de la semántica del lenguaje natural.

Palabras Clave: Anotación Semántica, Clasificación de Documentos, Clustering

1 Introducción

Uno de los principales requisitos para una aplicación o portal Web que ofrezca información recopilada y agregada de Internet mediante técnicas de Web Semántica es la capacidad de anotar semánticamente los contenidos a procesar. La anotación semántica [9] supone la asignación de una o varias etiquetas a cada documento, siendo éstas instancias de clases de una ontología, lo cual permite realizar la correspondencia directa entre el contenido del documento y la semántica de la clase asignada.

La utilización de herramientas como Semantic Word [13] permite realizar la anotación semántica de textos de forma manual, pero suponiendo una gran gran cantidad de trabajo para el administrador de los contenidos de la aplicación. Un enfoque semiautomático permite reducir la carga de trabajo dado que permitirá la identificación y clasificación automática de las entidades localizadas, aunque también exigirá la intervención humana para verificar que las anotaciones se han realizado de forma correcta [7]. En [6] se presenta un asistente para la anotación de textos mediante términos definidos por el usuario apoyándose en OWL con el fin de describir los datos semánticos necesarios.

Entre los diferentes sistemas que realizan una anotación automática de documentos web, cabe destacar el presentado en [8]. Este sistema permite anotar documentos en diferentes idiomas mediante el uso de los conceptos existentes en el tesauro multidioma EUROVOC [1]. Este proceso se realiza mediante una serie de algoritmos estadísticos apoyados en un proceso de entrenamiento sobre textos indexados de forma manual.

En este trabajo se presentan dos alternativas con el fin de anotar documentos de forma semántica mediante la utilización de conceptos extraídos del tesauro EUROVOC. Con el primer método se pretende afrontar aquellas situaciones en las que no se dispone de una colección de documentos previamente clasificados con el fin de entrenar el sistema anotador. La propuesta consiste en la utilización de un algoritmo de clasificación borroso que consigue anotar eficientemente los contenidos web mediante una serie de etiquetas prefijadas sin entrenamiento previo[11]. Por otro lado, la segunda aproximación está basada en la realización de un clustering previo de un conjunto de documentos significativo que permite obtener un conjunto de conceptos de referencia para realizar la anotación.

El resto de secciones del trabajo están organizadas de la siguiente forma: En la sección 2 se explican con detalle los diferentes algoritmos que se proponen para realizar la anotación semántica de documentos web. En la sección 3 se presentarán los experimentos realizados con el fin de evaluar las propuestas realizadas. Finalmente, en la sección 4 se incluirán las conclusiones y las líneas de trabajo futuro.

[1] http://eurovoc.europa.eu/

2 Algoritmos para la Anotación de Documentos basada en EUROVOC

EuroVoc es un tesauro multilingüe y multidisciplinar que es usado por diversas organizaciones gubernamentales de la Unión Europea en sus tareas administrativas. EUROVOC dispone de una organización jerárquica básica de los conceptos mediante dos estructuras, los campos y microtesauros, en las que se organizan tanto los conceptos o "descriptores", que designan sin ambigüedad el nombre de los conceptos, como los "no descriptores" que son las diferentes variaciones lingüísticas o regionales que puede tener un mismo concepto. En la actualidad se cuenta con más de 6700 conceptos entre los que hay desde lugares geográficos, términos económicos y administrativos un largo etcétera.

En las siguientes secciones se presentan las diferentes aproximaciones propuestas para la anotación semántica de contenido web expresado en diferentes idiomas mediante la utilización de descriptores EUROVOC. La primera de las aproximaciones está basada en clasificación sin entrenamiento de textos y la segunda en algoritmos de clustering. Previamente, se explica un procedimiento base dentro de las dos aproximaciones como es la transformación de los contenidos textuales en contenidos conceptuales basados en descriptores EUROVOC.

2.1 Transformación de Textos a Descriptores EUROVOC

Para cualquier proceso de etiquetación en el que intervenga el tesauro EUROVOC es recomendable realizar previamente diferentes transformaciones que permitan representar los textos en base a descriptores y viceversa. La base de toda transformación es el diccionario EUROVOC que en este caso es una estructura con dos componentes principales:

- *Traductor Texto-EUROVOC*: Se dispone una estructura organizada para cada uno de los diferentes idiomas para los cuales se va a almacenar la información. Dentro de la estructura de cada idioma se dispone de una tabla *hash*, siendo sus claves las palabras del idioma y el valor el *id* del descriptor EUROVOC correspondiente.
- *Traductor EUROVOC-Texto*: Se dispone de una tabla de equivalencia unívoca entre cada uno de los descriptores EUROVOC y una palabra o grupo de palabras en inglés que lo define.

Una vez construido este diccionario EUROVOC se puede utilizar para transformar el texto original de una contenido web en un texto basado únicamente en descriptores EUROVOC, lo cual permite una representación basada en conceptos independiente del idioma de trabajo (Algoritmo 1).

Algoritmo 1 Transformación de un texto en palabras a descriptores EUROVOC

$eurovoc \leftarrow LoadDictionary$
$lang \leftarrow languageDetector(doc)$
$descriptors \leftarrow eurovoc.get(lang)$
$eurovocText \leftarrow EmptyString$
for all $desc \in descriptors$ **do**
  $f \leftarrow frecuencia(doc, desc_{def})$
  for $i = 0$ to f **do**
    $eurovocText = eurovocText + desc_{id}$
  end for
end for

2.2 Anotación semántica basada en clasificación sin entrenamiento

La categorización o clasificación de documentos es la tarea de asignar de forma automática un conjunto de documentos a un conjunto de categorías predefinidas, normalmente conceptuales. El método habitual de construcción del clasificador de textos supone la utilización de un conjunto significativo de documentos etiquetados con una o más categorías de entre las predefinidas como conjunto de entrenamiento. Dentro de un entorno vinculado con Internet como puede ser el correo electrónico, o los flujos de noticias, la existencia de este conjunto significativo de documentos etiquetados no siempre es posible (las etiquetas no existen, no están normalizadas o no son fiables), a lo que hay que añadir el crecimiento exponencial del volumen del conjunto de entrenamiento según pretendamos conseguir una mayor precisión.

En este trabajo se propone prescindir del procedimiento de entrenamiento del clasificador, por el contrario, se utilizará como única base para realizar el proceso de clasificación la descripción semántica de los conceptos vinculados a cada una de las categorías predefinidas. Esto supone que a partir del concepto que define cada categoría se obtendrán las relaciones semánticas positivas que tenga con otros conceptos, tal y como se plantea en [11]. Esta definición semántica permitirá implementar un proceso de clasificación multidioma de forma sencilla. Para obtener estas relaciones semánticas positivas se utiliza como base de conocimientos EUROVOC, introduciendo una graduación borrosa de la relación existente entre diferentes descriptores de EUROVOC mediante la utilización de la fórmula presentada en [1].

En este caso, habiéndose elegido como dominio de trabajo "el cambio climático', un panel de expertos ha establecido el siguiente grupo de categorías para realizar la anotación semántica de contenidos web(entre paréntesis el descriptor EUROVOC equivalente): *water* (597), *atmosphere* (3967), *pollution* (2524), *environmental policy* (2470), *natural hazard* (4075), *technology* (4415), *documentation* (494).

Una vez definidas las categorías EUROVOC de trabajo se construye la base de conocimiento, esto es, a partir de cada categoría se obtienen los descriptores que están relacionados con ella dentro de EUROVOC, calculándose el correspondiente grado borroso de relación entre los distintos descriptores. En la Tabla 1 se pueden observar algunos descriptores asociados a la categoría "atmosphere".

Tabla 1: Descriptores relacionados con "atmosphere".

Nombre	Descriptor	Grado de Relación
ozone	2215	0,80
greenhouse effect	5605	0,60
climate	6011	0,40
humid zone	4797	0,28

Estos grados de relación entre las diferentes descriptores se almacenan en fichero de texto o en base de datos con el fin de que estén disponibles para el proceso de clasificación, el cual se presenta en el siguiente algoritmo.

Algoritmo 2 Proceso Completo de Clasificación.

```
Inicializar Definición de Categorías
Inicializar EUROVOC multidioma
for all doc do
  lang ← langDetector(doc) {palabras comunes [3]}
  eurovoc ← transformer(doc) {Alg. 1}
  cats ← classification(eurovoc)
  if cats ≡ ∅ then
    cats ← DEFAULT_CAT
  end if
end for
```

Como ejemplo del proceso de clasificación supongamos que una vez realizados los procesos de inicialización, se introduce el siguiente texto para clasificar procedente de una noticia de la Web.

```
Insights from Oil Spill Air Pollution Study Have Applications
Beyond Gulf. During a special airborne mission to study the
air-quality impacts of the BP Deepwater Horizon oil spill
last June, NOAA researchers discovered an important new
mechanism by which air pollution particles form.
```

En primer lugar se detecta el idioma y se transforma en forma de descriptores EUROVOC, cuyo resultado es el siguiente:

```
2527 2527 3161 3161 2524 2524 3967 3967
```

Posteriormente, se obtienen la lista de las categorías de entre las establecidas que resultan afines a la noticia. Esto se realiza mediante la aplicación del Algoritmo 3. En este algoritmo se procede iterativamente a identificar descriptor a descriptor si están relacionados con alguna de las categorías, si es así, a las categorías relacionadas se les suma el grado de relación del descriptor que aparece en el documento. Al terminar de inspeccionar el texto se dispone de un grado de aparición para cada una de las categorías, el cual servirá para identificar a las categorías ganadoras. El resto de categorías son descartadas.

Algoritmo 3 Clasificador.

```
cats ← CategoryList
resCats ← EmptyCategoryList {Resultados}
for all word ∈ eurovoc do
  if word ∈ KB then
    res ← KB.get(word)
    for all cat ∈ cats do
      cat.ocurr = cat.ocurr + res[cat] {se suma el
      grado de aparición}
      max = (cat.occur > max)?cat.occur; {se obtie-
      ne el máximo}
    end for
  end if
end for
for all cat ∈ cats do
  if cat.occur = max then
    resCats.add(cat)
  end if
end for
```

Esta claro que existen dos categorías "atmosphere" (3967) y "pollution" (2524) que aparecen directamente en el texto y son claras candidatas a ser asignadas. La categoría ganadora en este caso resulta ser "pollution" dado que obtiene un grado de aparición de 5.0, resultante de la suma de la aparición del propio descriptor de la categoría ($2*1=2$) y la aparición de los descriptores 2527 correspondiente a "air pollution" ($2*0.75=1.5$) y 2527 correspondiente a"oil pollution/oil spill" ($2*0.75=1.5$); mientras que "atmosphere" únicamente tiene un grado de 2.0 resultante de la aparición del propio descriptor de la categoría.

2.3 Anotación de documentos mediante Clustering

Las categorías establecidas según el método expuesto anteriormente son impuestas por un experto y además son estáticas, es decir, no evolucionan con el tiempo. Este tipo de categorías son útiles para organizar un sitio web pero no tanto para reflejar realmente el contenido del documento y la evolución de estos contenidos en el tiempo.

En el método que se expone en este apartado, el proceso de etiquetado de contenidos web se basa en una estructura enriquecida de categorías obtenida mediante la aplicación de un algoritmo de clustering a un conjunto de contenidos web significativos. Este método propuesto si que introduce, por lo tanto, una fase de entrenamiento, aunque es automática y no necesita ningún trabajo manual, ofreciendo como contrapartida ventajas como la independencia con respecto al criterio de ningún experto y que la estructura es capaz de evolucionar de forma automática conforme se van procesando nuevos documentos.

Los pasos del proceso de entrenamiento mediante clustering son los siguientes:

1. *Obtener la colección de textos significativos*: Se extraerán contenidos a partir de una serie de sitios web que el experto haya señalado como relevantes y durante un tiempo significativo.

2. *Transformar cada texto a conjuntos de descriptores EUROVOC*: Para cada texto y tras detectar el idioma en el que está elaborado se transforma su contenido en palabras a descriptores EUROVOC tal y como se especifica en la sección 2.1.

3. *Representar vectorialmente cada una de los textos*: Cada documento en su formato EUROVOC tendrá una representación vectorial en la que cada componente corresponderá a un descriptor EUROVOC y el valor de este componente vendrá dado por la utilización de la ponderación clásica $tf-idf$.

4. *Ejecutar el Algoritmo de Clustering*: Con la ejecución del algoritmo de clustering persigue agrupar todos los elementos del espectro de estudio, en este caso documentos, en un número reducido de grupos que estén formados por los más similares en cuanto a un criterio de comparación semántica [14].

5. *Obtener los descriptores más relevantes de cada uno de los grupos obtenidos*: Un descriptor es considerado relevante en una agrupación de contenidos web si aparece más frecuentemente que otro descriptor en el grupo, pero únicamente de forma ocasional en otros grupos. Para obtener el grado de relevancia de cada descriptor dentro del grupo, es necesario considerar el peso de cada descriptor en cada texto de entre los pertenecientes al grupo, la proporción entre los textos en los cuales el descriptor aparece y un coeficiente que permita definir la discriminación de relevancia entre los distintos grupos. De esta forma, el peso ϖ_i^n de un descriptor d_i dentro de un grupo g puede ser calculado de acuerdo con la Ecuación 1.

$$\varpi_i^n = \sum_{n \in g} w_i^n \times \left(1 + \frac{news(i,g)}{|D|}\right) \times Ln\left(\frac{|G|}{G(i)} + 1\right) \quad (1)$$

donde w_i^n representa el grado de relevancia del descriptor i en la documento n (n pertenece al grupo g), $news(i,g)$ se refiere al número de documentos del grupo g en los cuales el descriptor i aparece, $|D|$ es el número total de documentos considerados en el grupo, $|G|$ es el número total de grupos obtenidos y $G(i)$ representa el número de grupos en los cuales el descriptor i posee un grado de pertenencia positivo. Posteriormente, se normalizan los pesos obtenidos en el intervalo $[0,1]$. Solamente los conceptos que tengan peso suficiente ($w > 0.5$) son considerados útiles para la anotación semántica.

6. *Anotación semántica de las contenidos*: Los contenidos que han sido utilizados para obtener las categorías mediante la aplicación del algoritmo de clustering son etiquetados con el descriptor más relevante del grupo o grupos en los que hayan sido incluidos.

Una vez construida la estructura, cada vez que se vaya a procesar una nuevo contenido se procederá a su anotación semántica mediante el procedimiento detallado en el Algoritmo 4 del cual se obtendrá como resultado una serie de categorías/etiquetas con las que anotar el texto, o bien la categoría/etiqueta por defecto.

Algoritmo 4 Anotación Semántica basada en Clustering.

```
resCats ← EmptyCategoryList {Resultados}
lang ← languageDetector(doc))
eurovoc ← transformer(doc) {Alg. 1}
for all g ∈ ESTRUCTURA do
  s ← matching(g,doc)
  max = (cat.occur > max)?cat.occur {el maximo}
end for
for all g ∈ ESTRUCTURA do
  if s > 0,75 * max then
    resCats.add(g)
    Actualización de los pesos de los términos en el
    grupo g
  end if
end for
if resCats ≡ ∅ then
  resCats ← DEFAULT_CAT
end if
```

Especial mención precisa el proceso que permite calcular el grado de compatibilidad entre una nuevo texto n y cada uno de los grupos ($comp(n,g)$. En él se lleva a cabo una comparación entre los descriptores del documento y los descriptores más relevantes del grupo mediante la utilización de la siguiente fórmula (Ec. 2):

$$comp(n,g) = \frac{\left(\sum_{i \in n} w_i^n * w_i^g\right)}{|D|} \quad (2)$$

donde w_i^n es el peso del descriptor i en el la noticia n, w_i^g es el peso del descriptor i en el grupo g y $|D|$ es el número de descriptores diferentes que aparecen en el documento.

3 Experimentos

En esta sección se presentan los experimentos que se han llevado a cabo con el fin de realizar las verificaciones preliminares del funcionamiento de los algoritmos propuestos.

3.1 Evaluación de la aproximación basada en clasificación

Con el fin de evaluar los algoritmos de anotación de contenidos web, en primer lugar se ha construido una colección de entradas recientes en diferentes blogs relacionados con el cambio climático. La colección final está compuesta de 814 documentos divididos en las categorías siguientes: *water, atmosphere, pollution, environmental policy, natural hazard, technology y documentation*. Estos documentos han sido manualmente etiquetados por un experto utilizando la categoría o categorías que mejor definían a cada uno de los *posts*.

Como medidas de rendimiento, se han calculado, siguiendo su definición estándar, los valores de precisión (*P*), exhaustividad (*R*) y métrica F (la media armónica entre precisión y exhaustividad [14]. Para la evaluación del rendimiento medio para todas las categorías se utilizará el método basado en *micro-medias* (*micro-averaging*[15]).

Con el fin de comparar la base de conocimiento basada en EUROVOC y otras bases de conocimiento se han utilizado otros tesauros de uso general como ConceptNet [5], extrayendo relaciones como la Analogía Conceptual y la Proximidad Contextual, y WordNet [2] utilizando una relación basada en la Sinonimia [12]. La línea base del algoritmo es la no utilización de ninguna base de conocimiento y únicamente trabajar con la frecuencia de aparición de las propias categorías. Los resultados obtenidos pueden verse en la Tabla 2.

El procedimiento de clasificación logra un resultado aceptable en cuanto a clasificaciones precisión (78 %). Comparativamente además el uso de EUROVOC consigue unos mejores resultados, lo cual implica que la definición de las categorías es mucho más completa y correcta; esto se debe a que la utilización de descriptores y no descriptores amplia la definición de una categoría pero siempre dentro de un contexto.

Tabla 2: Resultados en cuanto a Precisión

B. Conocimiento	**P**	**R**	**F**
Eurovoc	**78 %**	**74 %**	**75 %**
Analogía Estructural	52 %	29 %	37 %
Proximidad Contextual	66 %	40 %	50 %
WordNet/Sinonimia	74 %	52 %	61 %
Linea Base	42 %	8 %	13 %

3.2 Evaluación de la aproximación basada en clustering

Con el fin de evaluar esta aproximación se ha ejecutado un proceso clustering sobre una base de 900 noticias en idioma inglés que se ha considerado significativa y que cubría suficientemente todo el espectro de temáticas relacionadas con el cambio climático. Entre las técnicas de clustering elegidas están un algoritmo de clustering jerárquico acumulativo clásico, el algoritmo de soft-clustering descrito en [4] y un algoritmo mixto de clustering borroso que combina parte jerárquica y parte plana no particionante [10]. De esta manera, se han utilizado tres algoritmos de clustering distintos para obtener tres estructuras distintas. Con cada una de ellas se ha procedido a anotar dos conjuntos de noticias uno de 772 noticias en idioma inglés y otro grupo de 469 noticias en castellano. Con este proceso se persigue comparar el resultado de las diferentes estructuras obtenidas y el funcionamiento de la propuesta sobre noticias multiidioma. Los indicadores utilizados para realizar la comparación han sido los porcentajes de anotaciones correctas/incorrectas y no anotadas, esto es así, dado que las categorías no están prefijadas, siendo posible que cada una de las estructuras ofrezca categorías distintas no siendo por ello ninguna de ellas incorrecta. El resultado del proceso de anotación utilizando idioma inglés puede observarse en la Tabla 3:

Tabla 3: Resultados del proceso de anotación en inglés

B. Conocimiento	**Correcto**	**Incorrecto**	**No**
Algoritmo Jerárquico	86 %	9 %	**5 %**
Soft Clustering	77 %	6 %	17 %
Algoritmo Mixto	**89 %**	**5 %**	6 %

Dado que esta aproximación utiliza entrenamiento para su elaboración (las categorías y los criterios son extraídos de una colección de noticias previa) los resultados de clasificación son mejores que en la aproximación anterior. El algoritmo de clustering borroso mixto es el que ofrece mejores resultados en general. También cabe reseñar que los resultados son mejores en cuánto el número de categorías posibles es mayor, lo cual puede ser contraproducente ya que este número puede ser demasiado alto y no ser significativo para la organización de las noticias con respecto al usuario final. A pesar de esto, un número mayor de categorías también ha supuesto que el número de noticias no clasificadas sea mínimo.

Los resultado en castellano (Tabla 4) son peores, principalmente debido a la naturaleza de las propias noticias que no proceden de fuentes tan adecuadas al dominio del cambio climático, sino de fuentes más generalistas, por lo que es más difícil de etiquetarlas con unas categorías definidas a partir de un conjunto de noticias que sí se encontraban dentro de un dominio específico.

Tabla 4: Resultados del proceso de anotación en castellano

B. Conocimiento	**Correcto**	**Incorrecto**	**No**
Algoritmo Jerárquico	65 %	22 %	13 %
Soft Clustering	50 %	26 %	24 %
Algoritmo Mixto	65 %	19 %	16 %

4 Conclusiones y Trabajos Futuros

En este trabajo se han presentado dos aproximaciones para la anotación semántica de contenidos web basadas en EUROVOC y técnicas borrosas de clasificación y clustering. La primera de ellas, basada en un algoritmo de clasificación sin entrenamiento, permite la anotación semántica de los contenidos web a partir de una serie de categorías fijas y su definición semántica extraída de EUROVOC. La segunda de las aproximaciones se basa en los resultados de la estructuración mediante clustering de un conjunto significativo de contenidos web. La estructura de categorías obtenida servirá como referencia para anotar semánticamente los nuevos contenidos que se vayan procesando por la aplicación.

Como principales líneas de trabajo futuro se afrontará la realización de nuevos experimentos con colecciones estándares de prueba con el fin de comparar estos algoritmos con aproximaciones similares, así como la utilización de varias ontologías combinadas como referencia con el fin de mejorar los resultados tanto en eficiencia como en corrección.

Agradecimientos

La investigación realizada está financiada por el proyecto FIDELIO (TIN2010-20395) del Ministerio de Educación y Ciencia y por los proyectos de excelencia SCAIWEB2 (PEIC09-0196-3018) y proyecto de investigación orientada PLINIO (POII10-0133-3516), ambos de la Junta de Comunidades de Castilla-La Mancha, así como el proyecto HECAOT (TC20112330, UCLM).

Referencias

[1] M. Batet, A. Valls, K. Gibert: A Distance Function to assess the Similarity of Words using Ontologies, En: Proceedings of the XV Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF 2010), Huelva (España), pp. 561–566, 2010.

[2] C. Fellbaum: WordNet: An Electronic Lexical Database. MIT Press. 1998.

[3] G. Grefenstette: Comparing two language identification schemes. En: Proceedings of JADT-95, 3rd International Conference on the Statistical Analysis of Textual Data, Roma (Italia), pp. 1–6, 1995.

[4] K. Lin, R. Kondadadi: A similarity-based soft clustering algorithm for documents. En DASFAA '01: Proceedings of the 7th International Conference on Database Systems for Advanced Applications, Washington, DC, USA, pp. 40–47. 2001.

[5] H. Liu, P. Singh: Commonsense reasoning in and over natural language. En Proceedings of the 8th Int. Conf. on Knowledge-Based Intelligent Information and Engineering Systems, Wellingtong (New Zealand), pp. 293–306, 2004.

[6] H. Liu, H.C. Chen, J.L. Jain, J.Y. Chen: Semiautomatic annotation system for OWL-based semantic search. En: Proceedings of the Int. Conf.on Complex, Intelligent and Software Intensive Systems, Seul (Corea del Sur), Vol.1(3), pp. 243–247, 2009.

[7] A. Maedche, S. Staab: Ontology learning for the Semantic Web. IEEE Intelligent Systems, Vol. 16(2), pp.72–79, 2001.

[8] B. Pouliquen, R. Steinberger, C. Ignat: Automatic annotation of multilingual text collections with a conceptual thesaurus. En: Proceedings of the Workshop "Ontologies and Information Extraction" at EUROLAN'2003, Bucharest (Rumanía), pp 9–28, 2006.

[9] L. Reeve, H. Han: Survey of semantic annotation platforms. En: Proceedings of the 2005 ACM symposium on Applied computing (SAC 2005), Santa Fe (USA), Vol. 2, p. 1634–1638, 2001.

[10] F. P. Romero, J.A. Olivas, P.J. Garcés: A soft approach to hybrid models for document clustering. En: Proceedings of the Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris (Francia), Vol 1, pp: 1040–1045, 2006.

[11] F. P. Romero, P. Julian-Iranzo, A. Soto, M. Ferreira-Satler, J. Gallardo-Casero: A Fuzzy Declarative Approach for Classifying Unlabeled Short Texts Using Thesauri. En: Proceedings of the 9th International Workshop on Fuzzy Logic and Applications (WILF 2011), Trani (Italia), p. 180–187. 2011.

[12] A. Soto, J.A. Olivas, M. Prieto: Fuzzy approach of synonymy and polysemy for information retrieval. Enn: R. Bello et al. (Eds.), Granular Computing: At the Junction of Rough Sets and Fuzzy Sets, Vol. 224 of Studies in Fuzziness and Soft Computing, Springer Berlin / Heidelberg (Alemania), pp. 179–198, 2008.

[13] M. Tallis: Semantic Word Processing for Content Authors. En: Proceedings of the 2nd International Conference on Knowledge Capture (KCAP 2003), Sanibel (USA), 2003.

[14] C. Van Rijsbergen: Information Retrieval. Butterworth, London, 1979.

[15] Y. Yang, X. Liu: A re-examination of text categorization methods. En: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, New York (USA), pp 42–49, 1999.

CLASIFICACIÓN DE DOCUMENTOS A PARTIR DE LA DEFINICIÓN CONCEPTUAL DE LAS CATEGORÍAS

Andrés Soto Villaverde[1], Francisco P. Romero[2], Jesús Serrano-Guerrero[2], José A. Olivas[2]

[1] *Universidad Autónoma del Carmen, Ciudad del Carmen, México,* soto_andres@yahoo.com

[2] *Universidad de Castilla La Mancha, Ciudad Real, España,* {FranciscoP.Romero, jesus.serrano, joseangel.olivas}@e-mail1.com

Resumen

Ante el crecimiento vertiginoso de la WWW, los algoritmos de aprendizaje automático para clasificación de documentos enfrentan dificultades, pues es más costoso o difícil disponer de conjuntos de documentos correctamente etiquetados para entrenar los algoritmos. En el presente trabajo se expone una metodología para clasificar documentos de texto en general a partir de la definición conceptual de las categorías. El enfoque se basa en construir alrededor de dichas categorías, una base de conocimientos empleando conjuntos y relaciones borrosas, que permita asociarles el máximo contenido semántico posible. Para construir la base de conocimientos se utilizaron diccionarios y tesauros como WordNet y EUROVOC. Los resultados preliminares obtenidos resultan satisfactorios y corroboran las posibilidades del enfoque.

Palabras Clave: Clasificación de documentos, Reuters, conjuntos y relaciones borrosas

1 INTRODUCCION

A través de la Historia, la clasificación o categorización ha pasado por diferentes etapas. Platón, filósofo griego, alumno de Sócrates y maestro de Aristóteles, es el primero en introducir el enfoque de agrupar los objetos basados en sus propiedades semejantes. Este enfoque fue desarrollado posteriormente por Aristóteles, el cual analiza las diferencias entre clases y objetos y aplica este enfoque a la clasificación de seres vivos [19]. De acuerdo con Cohen and Lefebvre [5], categorizar es un proceso cognitivo básico, fundamental para los humanos, el cual consiste en organizar y agrupar objetos en categorías con un propósito determinado. F. Sebastiani [17], por su parte, considera que la categorización (o clasificación) de textos es la actividad de etiquetar textos en lenguaje natural mediante categorías temáticas tomadas de un conjunto predefinido.

La categorización automática de textos en base a categorías o tópicos se remonta a los años 60. En este sentido, Borko y Bernick [4] plantean que la clasificación automática de documentos se refiere a la determinación de los temas a que hace referencia el documento. Para que un documento pueda ser clasificado en una categoría determinada debe asegurarse que el tema de que trata dicho documento se refiere a la categoría.

Posteriormente, hasta finales de los años 80, el enfoque más efectivo consistió en la construcción manual de programas clasificadores mediante la definición de conjuntos de reglas que describían el conocimiento de los expertos sobre cómo clasificar los documentos. Estos métodos han brindado muy buenos resultados en cuanto a eficiencia y efectividad.

A partir de los años 90, ante el crecimiento vertiginoso de documentos en línea, el enfoque basado en el paradigma de aprendizaje automático (machine learning) sustituyó al enfoque basado en la ingeniería del conocimiento, teniendo en cuenta la efectividad obtenida y los ahorros considerables en términos de tiempo empleado por los expertos.

El aprendizaje automático, en general, se refiere a lograr que una computadora sea capaz de cambiar su estructura, programa o datos (en base a información externa que reciben) de manera que su desenvolvimiento futuro sea mejor de acuerdo con algún criterio o métrica [23][14]. El aprendizaje automático para clasificación de documentos se considera un proceso inductivo que construye un programa clasificador de manera automática, a partir de un conjunto de documentos previamente clasificados, "aprendiendo" de ellos las características de las categorías de interés [17].

Este método ha mostrado muy buena efectividad, con independencia del dominio de clasificación y representa un ahorro sustancial de tiempo de los expertos que se requerían previamente para la construcción de programas clasificadores basados en reglas. Bajo este enfoque, las categorías se consideran como meras etiquetas simbólicas, sin que se admita la utilización de ningún conocimiento extra sobre el significado de las mismas para la construcción del clasificador.

Sin embargo, existen diferentes situaciones donde los documentos etiquetados son escasos o resulta costoso obtenerlos [13]. Esta situación limita las posibilidades de aplicar el aprendizaje automático a nuevos dominios. Los algoritmos de aprendizaje semi supervisados son de in-

terés en estos casos pues requieren menor cantidad de objetos etiquetados, a la vez que brindan la posibilidad de entender mejor los mecanismos de categorización humanos [24]. A pesar de ello, las aplicaciones de algoritmos semisupervisados a situaciones reales son aún escasas [6].
Otros enfoques plantean el uso de modelos basados en ontologías para mejorar la eficiencia y la efectividad de la clasificación de documentos [2], mientras otros plantean el uso de reglas del tipo Ripple-Down como método de adquisición incremental del conocimiento [15].
A continuación plantearemos una metodología para clasificar documentos de texto en general, aunque en este trabajo solo se aplica a la colección Reuters 21578. Teniendo en cuenta que las categorías son preestablecidas y conocidas, nuestro enfoque se basa en construir alrededor de las mismas una base de conocimientos que permita asociarles el máximo contenido semántico posible. Teniendo en cuenta que las categorías son elementos estables, menos proclives a cambiar que los documentos a clasificar, consideramos que la inversión requerida en tiempo de trabajo y de esfuerzos por parte de expertos se justifica. Por otra parte, para construir la base de conocimientos utilizaremos diccionarios y tesauros como WordNet [12] y EUROVOC[1]. Posteriormente incorporaremos diferentes ontologías tales como Concept Net [9][7], YAGO [21][8], DBpedia [3], FreeBase[2], entre otras. Finalmente, la base de conocimientos se complementará con meta heurísticas apropiadas.
En este trabajo, utilizaremos el término 'ontología' como concepto que engloba los conceptos de diccionario y tesauro, por cuanto la **ontología** se ocupa del estudio de las categorías de cosas que existen o pueden existir en un cierto dominio D [20]. Como resultado de ese estudio se obtiene una ontología particular consistente en un catalogo de cosas que existen en dicho dominio, aspecto este que forma parte del enfoque del presente trabajo. La ontología define al ser y establece las categorías fundamentales de las cosas a partir del estudio de sus propiedades, sistemas y estructuras. Dicha ontología define el vocabulario del dominio mediante un conjunto de términos básicos y relaciones entre dichos términos, así como las reglas que los combinan y relaciones que amplían las definiciones dadas en el vocabulario, lo cual es precisamente lo que pretendemos lograr. Los tipos en la ontología representan los conceptos y relaciones utilizados para expresarse alrededor de los tópicos del dominio D mediante un lenguaje L dado. Una ontología formal se especifica mediante una colección de nombres utilizados para identificar conceptos y relaciones tipo, organizados en un ordenamiento parcial mediante la relación tipo-subtipo, en la cual nos basaremos con frecuencia.
El resto del artículo está organizado como sigue. En la sección 2 se presenta una versión más detallada de nuestro enfoque para la clasificación de documentos. En la sección 3 se explica la estructura general del sistema. En la sección 4 se presentan los resultados obtenidos hasta el momento. En la sección 5 se exponen las conclusiones y los trabajos futuros.

[1] EuroVoc Thesaurus © European Union, 2010, http://eurovoc.europa.eu/
[2] Freebase http://www.freebase.com/

2 CLASIFICACIÓN DE DOCUMENTOS BASADA EN EL SIGNIFICADO INTRÍNSECO DE LAS CATEGORÍAS

Para la clasificación de documentos se parte de una descripción de un documento, $d \in D$, donde $D = \{d_j, j=1,\ldots,n\}$ es el espacio de documentos, así como de un conjunto fijo de clases o categorías $C = \{c_i, i=1,\ldots,m\}$, las cuales son definidas por humanos [10].
Habitualmente se emplea un algoritmo de aprendizaje para determinar una función de clasificación que asocie documentos a las clases con la tarea de determinar una asignación de un valor del conjunto {0, 1} a cada par <d,c> [17]. Se tiene además un conjunto de entrenamiento E de documentos etiquetados. El algoritmo de aprendizaje devuelve la función de clasificación aprendida, a partir de las características significativas detectadas en el conjunto de entrenamiento. Los algoritmos tratan de identificar las características en común que tienen los documentos pre clasificados en una categoría dada, sin tener en cuenta el significado semántico de los términos que definen las categorías en sí, las cuales se consideran meras etiquetas simbólicas, sin que se admita la utilización de ningún conocimiento extra sobre el significado de las mismas para la construcción del clasificador. El objetivo es que, a partir del conjunto de entrenamiento, el algoritmo aprenda por sí solo a reconocer cuáles documentos (nuevos) caen en una u otra categoría.
Un problema que se presenta actualmente con este enfoque es la gran cantidad y variedad de documentos (noticias, correos electrónicos, mensajes desde celulares, blogs, bases de datos corporativas, registros de pacientes médicos, bibliotecas digitales, etc) que se generan a diario y que se pretende clasificar, por lo que los documentos a clasificar varían mucho en sus características y muchas veces no se cuenta con conjuntos de entrenamiento apropiados o resulta muy costoso obtenerlos [13]. De esta manera, se requiere una inversión significativa en términos de tiempo empleado por los expertos para crear dichos conjuntos de entrenamiento, por lo que muchos trabajos de investigación se enfocan a reducir el tamaño del conjunto de entrenamiento, mediante algoritmos semi supervisados.
Otro aspecto a tener en cuenta es que las categorías, para un sistema o aplicación dado, se mantienen estables (no cambian) por períodos prolongados de tiempo, a diferencia del conjunto de documentos a clasificar que varía constantemente, así como su contenido, temática, estilo de redacción, etc.
El enfoque que se plantea en este trabajo no requiere del uso de conjunto de documentos de entrenamiento, por lo

cual permite soslayar las limitaciones y dificultades que estos imponen. El enfoque consiste en explicitar el significado de las categorías, mediante el establecimiento de relaciones semánticas con otros términos. De esta manera se construye una base de conocimientos formada por términos y relaciones entre los mismos que describe, en un cierto grado, el significado de las categorías. Esta base se complementa con un conjunto de reglas y procedimientos para aplicar las relaciones.

Teniendo en cuenta lo antes mencionado, el conjunto de categorías C se extiende para formar un sistema de clasificación S = (C, T, R, P) formado por:

- Conjunto de categorías C predefinidas por los humanos.
- Conjunto T de términos de un vocabulario V, asociados a las categorías predefinidas mediante relaciones de diferentes tipos.
- Conjunto de relaciones R borrosas, las cuales pueden ser de diferentes tipos y aridad, expresando los vínculos existentes entre los elementos, tales como equivalencia, sinonimia, meronimia, asociatividad, etc. Las relaciones asocian un grado de pertenencia (i. e. en qué medida o con qué fuerza se cumple la relación) entre los elementos.
- Conjunto de reglas o procedimientos para aplicar las relaciones.

El vocabulario de términos que se ha tomado como base es el de WordNet. No se requiere, en principio, incorporarlo todo, solo aquellos términos vinculados con las categorías. Por otro lado, determinados términos que no aparezcan en dicho diccionario se tomarán de otro. Al incorporar dichos términos al vocabulario V, se incorporan también las relaciones entre los mismos que se dispongan.

Hasta el momento, se han analizado los siguientes tipos de categorías:

1. Categorías definidas mediante términos simples: 'wheat', 'corn', 'barley', 'cocoa', 'grain'
2. Categorías definidas mediante abreviaturas simples: 'acq', 'cpi', 'gnp', 'dlr'
3. Categorías definidas mediante términos compuestos con '-' intermedio: 'money-supply', 'meal-feed', 'red-bean', 'copra-cake'
4. Categorías definidas mediante términos compuestos concatenados: 'cornglutenfeed', 'citrus-pulp', 'castorseed'
5. Categorías definidas mediante términos simples con sinónimos compuestos: 'linseed', 'soybean', 'oilseed'
6. Categorías definidas mediante abreviaturas y términos compuestos con '-' intermedio: 'veg-oil', 'lin-oil' , 'soy-oil'

Por ejemplo, entre los tópicos o categorías de la colección Reuters aparecen muchos términos abreviados como 'acq', 'cpi', 'gnp', 'dlr'. En estos casos, una relación denominada ***abreviatura*** entre C y T permite explicitar este vínculo: abreviatura(acq, [acquisition], grado1), abreviatura(dlr, [dollar], grado2). El grado permite indicar en qué medida se puede tomar un término dado como sustituto de una abreviatura dada, inclusive en dependencia del contexto donde se esté utilizando. Por ejemplo, la abreviatura 'dlr' tiene 39 significados posibles en AF AcronymFinder.com[3], apareciendo 'dollar' solamente en el contexto de Bussiness & Finance.

En otros casos, las categorías Reuters, como 'veg-oil', 'soy-oil' y 'money-fx', están formadas por dos palabras y algunas de ellas pueden estar abreviadas. La relación ***abreviatura*** se puede ampliar de forma que tenga en cuenta estos casos: abreviatura('veg-oil', [vegetable, oil], grado).

Por otra parte, hay muchas categorías que están relacionadas entre sí, como por ejemplo, 'dlr', 'austdlr', 'peseta', 'cruzado', 'yen', que corresponden a la unidad monetaria de diversos países. Se puede definir una relación entre dichas monedas y los respectivos países, así como una relación entre todas ellas como unidad monetaria.

Con estas relaciones se conforma un grafo a manera de red semántica o grafo conceptual [20] [18] entre los diferentes conceptos relacionados. Los nodos (términos, conceptos) que estén más interconectados tendrán un vínculo más estrecho, aplicando un criterio similar al del Page Rank de Google [1][22][16], lo cual nos define un grado de proximidad entre dichos términos. Las relaciones se importan de las diferentes ontologías existentes, tratando de no aumentar excesivamente el número de términos del vocabulario. En un principio se pretende incluir solo aquellos términos que estén vinculados directamente con las categorías a través de una u otra relación como la sinonimia, hiperonimia, meronimia, etc, así como otro tipo de relaciones existentes en las bases de conocimientos con las que se trabaja. Dichas relaciones pre definidas pueden tener ya asociados grados o niveles en su ontología de origen que puedan ser utilizados para medir el grado de la relación entre términos del vocabulario y las categorías.

De esta manera tendremos una red de conceptos asociados e interconectados entre sí, que irradian de un punto central o eje, formado por cada una de las categorías, con valores y pesos asociados, formando una suerte de nube conceptual a su alrededor. Dichas redes de conceptos se interceptan y entre mezclan entre sí como mencionábamos anteriormente formando la base de conocimientos asociada.

3 ESTRUCTURA GENERAL DEL SISTEMA

El sistema está conformado por tres componentes principales (ver figura 1). El primero que corresponde a la Base de Conocimientos asociada a las Categorías, denominado BCC, donde se intenta darle una connotación semántica a

[3] AF Acronym Finder.com http://www.acronymfinder.com/DLR.html

las mismas y su significado. El segundo formado por las diferentes ontologías, diccionarios y tesauros que se utilizan y denominado BCO. Y el tercero, más tradicional, corresponde al reconocimiento de los términos que aparecen en los documentos, la construcción de la tabla de índices y la determinación de la frecuencia con que aparece cada término en los diferentes documentos, denominado STI (Sistema de la Tabla de Índices) [11].

Figura 1 Esquema general del sistema

Con respecto al primer componente, ya se ha explicado cómo se ha ido construyendo. Hasta el momento, solo se están incluyendo las relaciones directas entre categorías y otros términos que aparecen en la BCO, para lo cual se buscan todas las relaciones que contienen como argumento, alguna de las categorías de Reuters y se incorporan a la BCC.

En el segundo componente cabe mencionar las herramientas de búsqueda y recuperación de relaciones y términos que incluyen las diferentes ontologías para su uso. Dichas herramientas se complementan con una interface para facilitar su uso desde el sistema.

El tercer componente comprende los analizadores léxico y sintáctico para recuperar y reconocer los diferentes términos existentes en los documentos, así como herramientas para reconocer los tokens, las raíces de las palabras, etiquetado gramatical (Part of Speech Tagger, POST), etc. También se incluyen las herramientas para insertar y buscar los términos en la tabla de índices, calcular su frecuencia, etc.

4 RESULTADOS OBTENIDOS

Hasta el momento, los experimentos realizados se han aplicado a los 21578 documentos de la colección Reuters considerando solo el primer tipo de categorías, las definidas mediante términos simples como: 'wheat', 'corn', 'barley', 'cocoa', 'grain'. En un experimento inicial E0, sin utilizar sinónimos ni otras relaciones, se obtuvieron resultados satisfactorios en muchos casos, pero otros decididamente malos (ver tabla 1).

Como no había ninguna explicación aparente para esta situación y eran muchos casos a analizar, se decidió concentrarse en los tópicos más frecuentes: 'earn', 'acq', 'money-fx', 'crude', 'grain', 'trade', 'interest', 'wheat', 'ship', 'corn'. Se eliminaron las abreviaturas (‘acq’, 'money-fx'), quedando solamente 8 categorías. Con dichas categorías, se realizó un experimento E1 consistente en comprobar en cuántos documentos aparecía referenciada la categoría sin utilizar sinonimia u otra relación y otro experimento donde ya se incluyeron los sinónimos o algún otro término adecuado (i.e. earn – earnings).

Tabla 1 Resultados experimento inicial E0

Topico	Precision	Recall	F1
'cocoa'	0.566	0.961	0.712
'grain'	0.684	0.441	0.536
'wheat'	0.654	0.974	0.782
'corn'	0.601	0.727	0.658
'barley'	0.568	1.000	0.725
'ship'	0.477	0.505	0.490
'trade'	0.218	0.933	0.353
'cotton'	0.562	1.000	0.720
'crude'	0.686	0.468	0.557

Como se muestra en la tabla 2 los resultados de recall mejoraron abiertamente (un 23% como promedio) entre un experimento y el otro, lo cual resulta lógico al tener en cuenta más términos a buscar (i.e. los sinónimos). En algunos casos, como 'earn', 'crude' y 'corn', el recall aumentó significativamente, más de un 50%. En cuanto a la precisión, se observa una disminución general, menos de un 10% como promedio, pero solo en los casos de 'crude' (39%) y 'ship' (24%) es importante. Con respecto al F1, podemos observar una mejoría grande en el caso de 'earn' (47%) y una mejoría leve para ‘corn’ (12%) y disminuciones leves para 'crude' y 'ship', lo cual es lógico teniendo en cuenta la afectación de la precisión antes señalada. Es conveniente resaltar que no se tomaron medidas especiales a favor de la precisión.

Teniendo en cuenta que ha habido un avance significativo general en los valores del recall, que indica que gran parte de los documentos relevantes han sido recuperados (i.e. ‘corn’ 100%, ‘wheat’ y ‘crude’ 87%, ‘trade’ 93% y ‘earn’ 89%), en próximos experimentos nos concentraremos en mejorar la precisión.

Tabla 2 Resultados experimentos 1 y 2

	Precision		Recall		F1	
Tópico	E1	E2	E1	E2	E1	E2
'earn'	0,37	0,37	0,03	0,89	0,05	0,52
'crude'	0,69	0,29	0,47	0,97	0,56	0,45
'grain'	0,68	0,67	0,44	0,53	0,54	0,59
'trade'	0,22	0,17	0,93	0,93	0,35	0,28
'interest'	0,10	0,07	0,39	0,39	0,15	0,12
'wheat'	0,65	0,64	0,97	0,97	0,78	0,78
'ship'	0,48	0,24	0,51	0,58	0,49	0,34
'corn'	0,60	0,64	0,73	1,00	0,66	0,78

5 CONCLUSIONES

Los algoritmos de aprendizaje automático para clasificación de documentos han sido los más eficientes desde la década del 90. Sin embargo, ante el crecimiento vertiginoso de la WWW, enfrentan dificultades ya que los documentos etiquetados, necesarios para entrenar los algoritmos, son más escasos o resulta costoso obtenerlos. Ello ha conllevado a buscar enfoques alternativos.

En este trabajo se presenta un enfoque diferente, ya que no parte de un conjunto de documentos previamente clasificados para ejercitar el algoritmo de reconocimiento. Por el contrario, se pretende crear una base de conocimientos, a partir de ontologías ya existentes, que caracterice de alguna forma el significado semántico de los tópicos o categorías empleados para la clasificación.

El método aquí planteado pudiera considerarse como una versión renovada del enfoque basado en reglas, el más efectivo hasta finales de los años 80. La diferencia fundamental radica en que, en nuestro caso, las reglas se importan de ontologías pre existentes, creadas por especialistas con objetivos generales, por lo cual no hay que invertir tanto tiempo y esfuerzo en crear el sistema.

Hasta el momento se ha trabajado en crear una base de conocimientos primaria en torno a las categorías o tópicos considerados en la colección Reuters. En dicha base se ha incluido un conjunto de relaciones borrosas para representar la proximidad entre las categorías Reuters y diferentes términos considerados sinónimos de las mismas en WordNet y otros diccionarios y tesauros.

Los resultados que se reportan en este trabajo corresponden a una investigación en curso que consideramos de interés por su enfoque. Los experimentos se han realizado con los 21578 documentos de la colección Reuters, teniendo solo en cuenta las categorías definidas mediante términos simples (sin abreviaturas ni términos compuestos). Para facilitar el análisis se consideraron solamente las categorías más frecuentes con estas características, es decir 8 categorías. Los resultados muestran que el recall obtenido en los experimentos utilizando la base de conocimientos (i.e. sinónimos) ha mejorado como promedio un 23% en los casos analizados, y en algunos casos más de un 50%, lo cual se corresponde con lo que se esperaba, dado el enfoque. La precisión, por otro lado, disminuyó en menos de un 10% como promedio, aunque en un caso, la disminución llego hasta un 39%. Es conveniente resaltar que no se tomaron medidas especiales a favor de la precisión, lo cual nos proponemos hacer en experimentos futuros. Con respecto al F1, la disminución fue del orden de un 4% con una afectación máxima de 47%. Consideramos que, como resultados primarios, son satisfactorios.

En el futuro, nos proponemos afinar los criterios utilizados para determinar si un documento se clasifica o no en una categoría dada, a la vez que ampliar nuestro estudio al resto de las categorías consideradas en Reuters.

Agradecimientos

Esta investigación está parcialmente financiada por los proyectos FIDELIO (TIN2010-20395), MEC-FEDER, SCAIWEB2 (PEIC09-0196-3018) y PLINIO (POII10-0133-3516) de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1] Altman, A. and Tennenholtz, M. 2005. Ranking systems. (2005), 1-8.

[2] An Ontology Based Approach for Chinese Web Texts Classification. Information Technology Journal, pp 796-801, 2008.

[3] Bizer, C. et al. 2009. DBpedia - A crystallization point for the Web of Data. *Web Semantics: Science, Services and Agents on the World Wide Web*. 7, (Sep. 2009), 154-165.

[4] Borko, H. and Bernick, M. 1963. Automatic Document Classification. *Journal of the ACM*. 10, (Apr. 1963), 151-162.

[5] Cohen, H. and Lefebvre, C. 2005. *Handbook of categorization in cognitive science*. Elsevier.

[6] Druck, G. et al. 2008. Learning from labeled features using generalized expectation criteria. (2008), 595.

[7] H Liu and P Singh 2004. ConceptNet - a practical commonsense reasoning tool-kit. *BT Technology Journal*. 22, 4 (2004), 211-226.

[8] Kasneci, G. et al. 2009. The YAGO-NAGA approach to knowledge discovery. *ACM SIGMOD Record*. 37, (Mar. 2009), 41.

[9] Liu, H. and Singh, P. 2004. Commonsense reasoning in and over natural language. *Proceedings of the 8th International Conference On Knowledge-Based Intelligent Information And Engineering Systems (KES-2004*. 3215, (2004).

[10] Manning, C.D. et al. 2008. *Introduction to Information Retrieval*. Cambridge University Press.

[11] Manning, C.D. et al. 2008. *Introduction to Information Retrieval*. Cambridge University Press.

[12] Miller, G.A. 1995. WordNet: a lexical database for English. *Communications of the ACM*. 38, (Nov. 1995), 39-41.

[13] Nigam, K. et al. 1999. Text Classification from Labeled and Unlabeled Documents using EM. *Machine Learning*. 39, (1999), 103--134.

[14] Nils J. Nilsson 1998. *Introduction To Machine Learning*.

[15] Park, G. and Kim, S. 2006. Web Document Classification Using Changing Training Data Set. *Computational Science and Its Applications - ICCSA 2006*. M.L. Gavrilova et al., eds. Springer Berlin Heidelberg. 565-574.

[16] Richardson, M. and Domingos, P. 2002. The Intelligent Surfer: Probabilistic Combination of Link

and Content Information in PageRank. *Proceedings of Advances in Neural Information Processing Systems. 14*. (2002).

[17] Sebastiani, F. 2002. Machine learning in automated text categorization. *ACM Computing Surveys*. 34, (Mar. 2002), 1-47.

[18] Shapiro, S.C. 1992. *Encyclopedia of Artificial Intelligence*. John Wiley & Sons.

[19] Sowa, J.F. 1999. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks / Cole.

[20] Sowa, J.F. 1999. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks / Cole.

[21] Suchanek, F. et al. 2007. Yago: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. *Proceedings of the 16th international conference on World Wide Web* (Banff, Alberta, Canada, 2007), 697-706.

[22] Taher, P.P. et al. 2003. An Analytical Comparison of Approaches to Personalizing PageRank. *Stanford University Technical Report.* (2003).

[23] Tom M. Mitchell 2006. The Discipline of Machine Learning. Carnegie Mellon University.

[24] Zhu, X. and Goldberg, A.B. 2009. Introduction to Semi-Supervised Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*. 3, (Jan. 2009), 1-130.

AN IMAGE RETRIEVAL APPROACH BASED ON FUZZY DOMINANT COLOR DESCRIPTORS

J. Chamorro-Martínez[1], **J.M. Soto-Hidalgo**[2], **D. Sánchez**[3], **P. Martínez-Jiménez**[1]

[1]*Dept. of Computer Science and A.I., University of Granada, Spain*, {jesus,pedromartinez}@decsai.ugr.es
[2]*Dept. of Comp. Architecture, Electronics, and Electronic Tech., University of Córdoba, Spain*, jmsoto@uco.es
[3]*European Centre for Soft Computing, Mieres, Asturias, Spain*, daniel.sanchezf@softcomputing.es

Abstract

In this paper, an image retrieval approach based on Dominant Color Descriptors is proposed. These descriptors are defined as a fuzzy set over a finite universe of fuzzy colors, in which membership degrees represent the dominance of each color. The fuzzy colors employed are fuzzy sets defined on an ordinary color space, filling the semantic gap between the color representation in computers and its human perception. The dominance of each fuzzy color is calculated on the basis of a fuzzy quantifier representing the notion of dominance, and a fuzzy histogram representing as a fuzzy quantity the percentage of pixels that match each fuzzy color. The obtained descriptors are applyed for describing dominant colors and retrieving images on the basis of these descriptions.

Keywords: Fuzzy set, Dominant color descriptor, Image retrieval.

1 INTRODUCTION

Traditional and most common methods for describing visual information are based on attaching keywords to the images. Although this approach has proved to be useful, it is time-consuming, laborious, expensive, and subjective due to a person is required for making the description of the large amount of visual information available in digital format. In addition, it is almost impossible for a person to label an image with every possible keyword that may be relevant. To overcome this limitation, the use of visual descriptors has been suggested [1, 11]. They describe elementary characteristics such as color, texture and shape, which are automatically extracted from images [3].

Among these characteristics, color plays an important role because of its robustness and independence from image size and orientation [10]. In this sense, several proposals are formulated for describing color properties in images. For example, the well-known MPEG-7 standard [14] proposes color descriptors such as the Color Structure Descriptor, that tries to identify color distributions in an image, and the Scalable Color Descriptor, that deals with a scalable representation of the color.

In this paper we focus on the dominant color descriptor, one of the most important descriptors in MPEG-7. A dominant color descriptor must provide an effective, compact, and intuitive representation of the most representative colors presented in an image. In this context, many approaches to dominant color extraction have been proposed in the literature [7, 15]. Most of them perform the extraction process based on histogram analysis [16] or clustering techniques [6] in color domain. Nevertheless, these approaches consider a crisp notion of dominance, when in fact for human's perception there are degrees of dominance, that is, colors can be clearly dominant, clearly not dominant, or can be dominant to a certain degree. In addition, most of the times they do not consider subsets of crisp colors, as represented in computers, that fully match human perception as expressed by linguistic color terms.

In this paper, we provide a solution to the previous problems by introducing a fuzzy descriptor for dominant colors as a level-two fuzzy sets on fuzzy colors. To the best of our knowledge, this is a completely new approach for the definition of dominant color descriptor. As a basis for the descriptor we rely on the definition of fuzzy colors as fuzzy subsets of a crisp color space (RGB for instance), introduced in [9]. Concretely, a color will be modeled by means of a fuzzy set (a fuzzy color) and a fuzzy partition will be defined in the color feature domain (a fuzzy color space). In this paper we propose to define this partition on the basis of the ISCC-NBS color naming system [4, 5], that is based on the human perception of color. In order to obtain the descriptor we use a histogram of fuzzy colors, as the dominance is related with the frequency of the colors in the image. A fuzzy quantifier will be employed in

order to represent the semantics of *dominant* on the basis of the amount of pixels having a certain color, in a natural way. The counting will be performed by using the scalar sigma-count (sum of membership degrees), and the evaluation of the dominance will be finally the compatibility between the quantifier defining the semantics of dominance, and the sigma-count. We use this fast calculation since it is well known to be completely equivalent to other quantification models for evaluating type I sentences with non-decreasing quantifiers, like in this case. On the basis of this dominance, a new descriptor "*Fuzzy Dominant Color Descriptor*" is proposed in this paper. This descriptor will be defined as a fuzzy set over the finite universe of color fuzzy sets (with membership degrees according to its dominance degrees).

The rest of the paper is organized as follows. As a basis for our descriptor, in section 2 we recall fuzzy colors as defined and calculated in [9]. The dominance-based color fuzzy descriptor which is our main contribution here, is defined in section 3. In section 4 we show some results of calculating color descriptors for real images; we also show the result of some image retrieval experiments for illustrative purposes. Finally, conclusions and future work are summarized in section 5.

2 FUZZY MODELLING OF COLORS

In this section, the notions of fuzzy color (section 2.1) and fuzzy color space (section 2.2) we presented in a previous work [9] are summarized. Based on it, a fuzzy partition will be defined in the color feature domain (our fuzzy color space) according to the ISCC-NBS color naming system [5].

2.1 Fuzzy color

For representing colors, several color spaces can be used. In essence, a color space is a specification of a coordinate system and a subspace within that system where each color is represented by a single point. The most commonly used color space in practice is RGB because is the one employed in hardware devices (like monitors and digital cameras). It is based on a cartesian coordinate system, where each color consists of three components corresponding to the primary colors red, green, and blue. Other color spaces are also popular in the image processing field: linear combination of RGB (like CMY, YCbCr, or YUV), color spaces based on human color terms like hue or saturation (HSI, HSV or HSL), or perceptually uniform color spaces (like CIELa*b*, CIELuv, etc.).

In order to manage the imprecision in color description, we introduce the following definition of fuzzy color:

Definition 2.1 *A fuzzy color $\widetilde{C}$ is a normalized fuzzy subset of colors.*

As previously explained, colors can be represented as a triplet of real numbers corresponding to coordinates in a color space. Hence, a fuzzy color can be defined as a normalized fuzzy subset of points of a color space. From now on, we shall note XYZ a generic color space with components X, Y and Z[1], and we shall assume that a color space XYZ, with domains D_X, D_Y and D_Z of the corresponding color components is employed. This leads to the following more specific definition:

Definition 2.2 *A fuzzy color $\widetilde{C}$ is a linguistic label whose semantic is represented in a color space XYZ by a normalized fuzzy subset of $D_X \times D_Y \times D_Z$.*

Notice that the above definition implies that for each fuzzy color $\widetilde{C}$ there is at least one crisp color $\mathbf{r}$ such that $\widetilde{C}(\mathbf{r}) = 1$.

In this paper, and following [9], we will define the membership function of $\widetilde{C}$ as

$$\widetilde{C}(\mathbf{c};\mathbf{r},S,\Omega) = f\left(\left|\overrightarrow{\mathbf{rc}}\right|;t_1^c,\ldots,t_n^c\right) \quad (1)$$

depending on three parameters: $S = \{S_1,\ldots,S_n\}$ a set of bounded surfaces in XYZ verifying $S_i \cap S_j = \emptyset\ \forall i,j$ (i.e., pairwise disjoint) and such that $Volume(S_i) \subset Volume(S_{i+1})$; $\Omega = \{\alpha_1,\ldots,\alpha_n\} \subseteq (0,1]$, with $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_n = 0$, the membership degrees associated to S verifying $\widetilde{C}(\mathbf{s};\mathbf{r},S,\Omega) = \alpha_i\ \forall \mathbf{s} \in S_i$; and $\mathbf{r}$ a point inside $Volume(S_1)$ that is assumed to be a crisp color representative of $\widetilde{C}$.

In Eq.1, $f : \mathbb{R} \to [0,1]$ is a piecewise function with knots $\{t_1^c,\ldots,t_n^c\}$ verifying $f(t_i^c) = \alpha_i \in \Omega$, where these knots are calculated from the parameters $\mathbf{r}$, S and Ω as follows: $t_i^c = \left|\overrightarrow{\mathbf{rp}_i}\right|$ with $\mathbf{p}_i = S_i \cap \overline{\mathbf{rc}}$ being the intersection between the line $\overline{\mathbf{rc}}$ (straight line containing the points $\mathbf{r}$ and $\mathbf{c}$) and the surface S_i, and $\left|\overrightarrow{\mathbf{rp}_i}\right|$ the length of the vector $\overrightarrow{\mathbf{rp}_i}$.

2.2 Fuzzy color space

For extending the concept of color space to the case of fuzzy colors, and assuming a fixed color space XYZ, with D_X, D_Y and D_Z being the domains of the corresponding color components, the following definition is introduced:

Definition 2.3 *A fuzzy color space $\widetilde{XYZ}$ is a set of fuzzy colors that define a partition of $D_X \times D_Y \times D_Z$.*

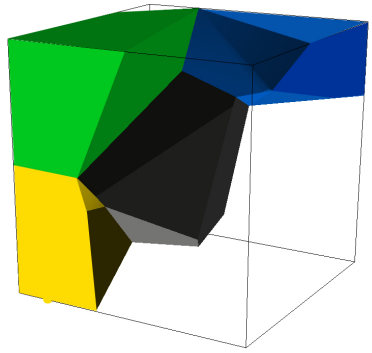
[1] Although we are assuming a three dimensional color space, the proposal can be easily extended to color spaces with more components.

As we introduced in the previous section (see Eq.1), each fuzzy color $\widetilde{C}_i \in \widetilde{XYZ}$ will have associated a representative crisp color $\mathbf{r}_i$. Therefore, for defining our fuzzy color space, a set of representative crisp colors $R = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ is needed. In this paper we propose to use the color names (and color points) provided by the ISCC-NBS system [4, 5], witch is based on human tests about color perception. ISCC-NBS system defines a set of valid colors terms. In this paper, the basic set formed by 13 color names (10 hues -pink, red, orange, yellow, brown, olive, green, blue, violet, purple- and 3 achromatic colors -white, grey and black-) will be used (i.e, $R = \{\mathbf{r}_1, \ldots, \mathbf{r}_{13}\}$ with $\mathbf{r}_i$ a RGB color).

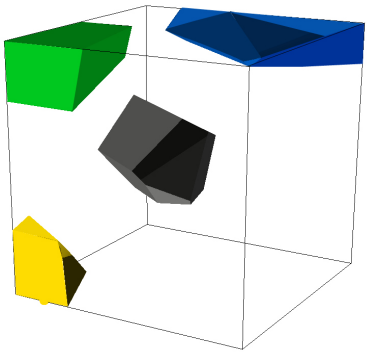
For defining each fuzzy color $\widetilde{C}_i \in \widetilde{XYZ}$, we also need to fix the set of surfaces S_i and the associated memberships degrees Ω_i (see Eq.1). In this paper, we have focused on the case of convex surfaces defined as a polyhedra (i.e, a set of faces). Concretely, three surfaces $S_i = \{S_1^i, S_2^i, S_3^i\}$ have been used for each fuzzy color $\widetilde{C}_i$ with $\Omega_i = \{1, 0.5, 0\} \forall i$.

To obtain $S_2^i \in S_i \; \forall i$, a Voronoi diagram has been calculated [8] with R as centroid points . As results, a crisp partition of the color domain given by convex volumes is obtained (each volume will define a Voronoi cell). The surfaces of the Voronoi cells will define the surfaces $S_2^i \in S_i \; \forall i$. Once S_2^i is obtained, the surface S_1^i (resp. S_3^i) is calculated as a scaled surface of S_2^i with scale factor of 0.5 (resp. 1.5). For more details about the parameter values which define each polyhedra, see [9].

Figure 1 shows the surfaces S_2^{yellow}, S_2^{blue}, S_2^{green} and S_2^{grey} (on the top), and S_1^{yellow}, S_1^{blue}, S_1^{green} and S_1^{grey} (on the bottom). The representative crisp values are $\mathbf{r}_{yellow} = [254, 220, 1]$, $\mathbf{r}_{blue} = [1, 90, 200]$, $\mathbf{r}_{green} = [1, 220, 30]$ and $\mathbf{r}_{gray} = [132, 132, 132]$.



(a)



(b)

Figure 1: Some surfaces associated to the fuzzy color space. (a) S_2^i (b) S_1^i for $i = \{yellow, blue, green, grey\}$

3 DOMINANCE-BASED COLOR FUZZY DESCRIPTOR

For describing semantically an image, the dominant colors will be used. In this section, Fuzzy Descriptors for dominant colors are proposed (section 3.2) on the basis of the dominance degree of a given color (section 3.1).

3.1 Dominant Fuzzy Colors

Intuitively, a color is dominant to the extend it appears frequently in a given image. As it is well known in the computer vision field, the histogram is a powerful tool for measuring the frequency in which a property appears in an image. The histogram is a function $h(x) = n_x$ where x is a pixel property (grey level, color, texture value, etc.) and n_x is the number of pixels in the image having the property x. It is common to normalize a histogram by dividing each of its values by the total number of pixels, obtaining an estimate of the probability of occurrence of x.

Working with fuzzy properties suggests to extend the notion of histogram to "fuzzy histogram". In this sense, a fuzzy histogram will give us information about the frequency of each fuzzy color. In this paper, the counting will be performed by using the scalar sigma-count (i.e., the sum of membership degrees). Thus, for any fuzzy set F with membership function $F : X \to [0, 1]$, the fuzzy histogram is defined as

$$h(F) = \frac{1}{NP} \sum_{x \in X} F(x) \tag{2}$$

with NP being the number of pixels.

Using the information given by the histogram, we will measure the "dominance" of a color fuzzy set. Dominance is an imprecise concept, i.e., it is possible in general to find colors that are clearly dominant, colors that are clearly not dominant, and colors that are dominant to a certain degree, that depends on the percentage of pixels where the color appears.

It seems natural to model the idea of dominance by means of a fuzzy set over the percentages given by $h(F)$, i.e.,

a fuzzy quantifier defined by a non-decreasing subset of the real interval $[0,1]$. Hence, we define the fuzzy subset "Dominant", noted as *Dom*, as follows:

$$Dom(F) = \begin{cases} 0 & h(F) \leq u_1 \\ \frac{h(F)-u_1}{u_2-u_1} & u_1 \leq h(F) \leq u_2 \\ 1 & h(F) \geq u_2 \end{cases} \quad (3)$$

where u_1 and u_2 are two parameters such that $0 \leq u_1 < u_2 \leq 1$, and $h(F)$ is calculated by means of Eq. 2.

3.2 Dominance-based fuzzy descriptors

On the basis of the dominance of colors, a new image descriptor is proposed in this section. For a generic image property, we introduce the following definition of fuzzy descriptor:

Definition 3.1 *Let $\mathscr{P}$ be a finite universe of fuzzy sets associated with a given image property*[2]*. A fuzzy descriptor is defined as a fuzzy set over the reference universe $\mathscr{P}$*

In the case of the "Color" property, we define the following fuzzy descriptor for dominant colors:

Definition 3.2 *Let $\mathscr{C}$ be a finite reference universe of color fuzzy sets. We define the* Fuzzy Dominant Color Descriptor *as the fuzzy set*

$$FDCD = \sum_{C \in \mathscr{C}} Dom(C)/C \quad (4)$$

with $Dom(C)$ being the dominance degree of C given by Eq. 3.

4 RESULTS

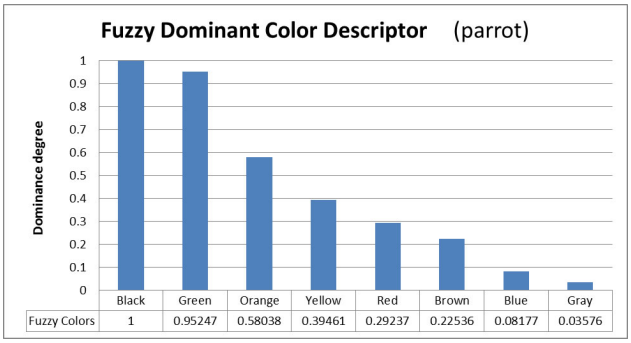
In this section, we will use the proposed Fuzzy Dominant Color Descriptor (FDCD) in order to obtain the dominant colors present in an image. In addition, the FDCD will be used to define conditions in image retrieval queries.

In this paper, we have used the fuzzy color space defined in section 2 which is designed on the basis of a collection of crisp colors and color names. The color names (and color points) provided by the ISCC-NBS system [4, 5], based on human tests about color perception, have been used. Specifically, we have used the set formed by 13 color names (10 hues -pink, red, orange, yellow, brown, olive, green, blue, violet, purple- and 3 achromatic colors -white, grey and black-).

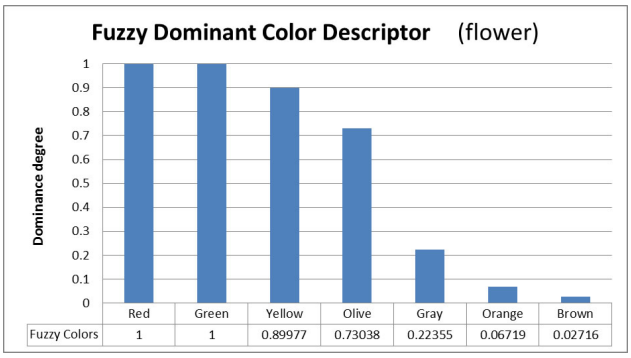
[2]For example, the property "Color" with $\mathscr{P}$ being a set of fuzzy colors.

(a)

(b)



(c)



(d)

Figure 2: **(a)** Parrot image. **(b)** Flower image. **(c)(d)** A graphical representation of the *FDCD* obtained from (a) and (b) images respectively.

In our experiments we have empirically fixed the parameters u_1 and u_2 in equation 3 to 0.05 and 0.25 respectively.

4.1 Fuzzy Dominant Color Descriptor examples

In order to show the suitability of the Fuzzy Dominant Color Descriptor (FDCD) for describing the dominant colors in an image, we have calculated it in real images with a large variety of colors. This way, we get a description of the dominant colors present in an image in terms of fuzzy colors and the corresponding dominance degrees.

Figure 2 shows a graphical representation of the FDCD obtained from two color images. For the *parrot* image (Figure 2(a)), the Fuzzy Dominant Color Descriptor (Figure 2(c)) corresponding to fuzzy colors and its

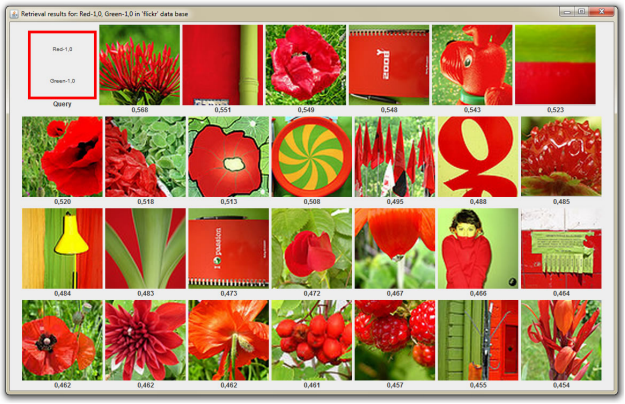


Figure 3: Retrieval results on Flickr "interestingness" collection for the color inclusion query with the labels *Red* and *Green*

(a)

Figure 4: Retrieval results on Flickr "interestingness" collection using the previously colorful image (Flower) as query.

dominance degree (ordered in decreasing order of dominance degree) is $FDCD^{parrot} = 1/Black + 0.95/Green + 0.58/Orange + 0.39/Yellow + 0.29/Red + 0.22/Brown + 0.08/Blue + 0.03/Gray$.

Notice that the FDCD provides a full description of the perceived dominant colors present in Figure 2(a). As it is shown in this image, the background color (*Black*) is clearly dominant, which match with the $FDCD^{parrot}$ ($1/Black$). In addition, the colors present in the wings of the parrot (*Green*, *Orange*, *Yellow* and *Red*) are less dominant than the background, as well as the colors of the trunk (*Brown*) and the head (*Blue*) are less dominant than the colors of the wings. It can be seen that our fuzzy descriptor captures the dominance degree of the wings colors ($0.95/Green$, $0.58/Orange$, $0.39/Yellow$, $0.29/Red$) and the trunk and the head colors ($0.22/Brown$ and $0.08/Blue$), respectively.

Similarly, in Figure 2(d) we can see the Fuzzy Dominant Color Descriptor obtained from the *flower* image (Figure 2(b)). In this case, $FDCD^{flower} = 1/Red + 1/Green + 0.89/Yellow + 0.73/Olive + 0.22/Gray + 0.06/Orange + 0.02/Brown$. Thus, the color of the petals and the leaves are clearly dominant ($1/Red$ and $1/Green$ respectively) and the colors of the stamen and pistil area dominate with less degree ($0.89/Yellow$). Finally, we can see that the color of the shadow in the petals is dominant to an insignificant degree ($0.02/Brown$).

Unlike existing techniques that consider *dominant* as a crisp concept, the FDCD describes dominance degrees to the fuzzy colors considered.

4.2 Retrieval examples

In this section, the FDCD will be used to define conditions in image retrieval queries. For this purpose, in order to match descriptors, we propose to use fuzzy resemblance operators between dominant color descriptors. Given two Fuzzy Dominant Color Descriptors, $FDCD^i$ and $FDCD^j$, we propose to calculate the resemblance degree between $FDCD^i$ and $FDCD^j$ by means of the *Generalized Resemblance between Fuzzy Sets* proposed in [2], which is based on the concept of double inclusion.

In this illustrative application, the *Flickr* online image database has been used. Flickr [12] is an image hosting which has increasingly been adopted by many web users to share and embed personal photographs. Nowadays, there are more than 5 billion images hosted in *Flickr*. We have used a collection of 10000 images provided by *Flickr* API [13] as the most interesting photos for a certain month.

Two retrieval examples will be shown in order to illustrate the performance of the FDCD for retrieving images. In the first example, linguistic labels will be used as query, whereas in the other example the query is specified as an image. Notice that our approach to dominant color descriptor allows us to perform both kind of queries with a single representation.

Figure 3 shows the retrieval results for the query *images where red and green are dominant*, using the color-based linguistic labels *Red* and *Green* as dominant colors. This query can be solved by considering the query descriptor will be $FDCD^{query} = 1/Red + 1/Green$, and determining its degree of inclusion in the dominant color description of each image. Retrieved images are shown in decreasing order of resemblance degree, that is shown below each image. Notice that images containing similar dominant colors can be found in these results. Furthermore, the most representative colors in these images match human perception of the color-based linguistic labels *Red* and *Green*.

Also, we may be interested in getting images with similar

dominant colors to a sample image. This condition can be defined by using the fuzzy resemblance operator to compute the resemblance degrees between the fuzzy descriptor $FDCD$ and the corresponding $FDCD$ of each image in the collection.

We have used the $FDCD$ calculated for the previous image (Figure 2(b)) for retrieving images with similar dominant colors to this. In this case, the retrieval results are shown in Figure 4. These results are ordered by its resemblance degree. It can be shown that images with similar FDCD (similar dominant colors and dominance degree) can be found in these results.

5 CONCLUSIONS

In this paper, a new Fuzzy Dominant Color Descriptor has been proposed. This descriptor has been defined as a fuzzy set over a finite universe of fuzzy colors, in which membership degrees represent the dominance of each color. The color fuzzy sets have been defined taking into account the relationship between the color representation in computers and its human perception. We have illustrated the usefulness of our proposals with an application in image retrieval.

Several future work related to this will be to apply the descriptor in the linguistic description of images and other interesting applications. We also plan to define other color descriptors on the basis of the fuzzy color spaces we have introduced in previous works.

Acknowledgements

This work has been partially supported by the Spanish Government under the TIN2009-08296 and Consolider Ingenio 2010 MIPRCV (CSD2007-00018) projects.

References

[1] E. Akbas and F.T. Yarman-Vural. Automatic image annotation by ensemble of visual descriptors. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[2] J. Chamorro-Martínez, J.M. Medina, C. Barranco, E. Galán-Perales, and J.M. Soto-Hidalgo. Retrieving images in fuzzy object-relational databases using dominant color descriptors. *Fuzzy Sets and Systems*, 158(3):312–324, February 2007.

[3] S.F. Chang, T. Sikora, and A. Puri. Overview of the mpeg-7 standard. *IEEE transaction on Circuits and Systems for Video Technology*, 11:688–695, 2001.

[4] K.L. Kelly and D.B. Judd. The iscc-nbs method of designating colors and a dictionary of color names. *National Bureau of Standards (USA)*.

[5] K.L. Kelly and D.B. Judd. *Color universal color language and dictionary of names*. Number 440. 1976.

[6] S. Kiranyaz, S. Uhlmann, and M. Gabbouj. Dominant color extraction based on dynamic clustering by multi-dimensional particle swarm optimization. In *Content-Based Multimedia Indexing, 2009. CBMI '09. Seventh International Workshop on*, pages 181 –188, 2009.

[7] A. Li and X. Bao. Extracting image dominant color features based on region growing. In *Web Information Systems and Mining (WISM), 2010 International Conference on*, volume 2, pages 120 –123, 2010.

[8] F.P. Preparata and M.I. Shamos. *Computational geometry : algorithms and applications*. Springer-Verlag, New York, 2nd edition edition, 1988.

[9] J.M. Soto-Hidalgo, J. Chamorro-Martinez, and D. Sanchez. A new approach for defining a fuzzy color space. In *IEEE World Congress on Computational Intelligence (WCCI 2010)*, pages 292–297, July 2010.

[10] K.E.A. van de Sande, T. Gevers, and C.G.M. Snoek. Evaluating color descriptors for object and scene recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(9):1582 –1596, 2010.

[11] J. Wu and J.M. Rehg. Centrist: A visual descriptor for scene categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3:4, 2010.

[12] Yahoo! Flickr. an image hosting and video hosting website @ONLINE, 2011.

[13] Yahoo! Flickr api. a programmers place to create applications @ONLINE, 2011.

[14] A. Yamada, M. Pickering, S. Jeannin, and L.C. Jens. Mpeg-7: Visual part of experimentation model version 9.0. *ISO/IEC JTC1/SC29/WG11/N3914*, 2001.

[15] N. Yang, W. Chang, C. Kuo, and T. Li. A fast mpeg-7 dominant color extraction with new similarity measure for image retrieval. *Journal of Visual Communication and Image Representation*, 19(2):92 – 105, 2008.

[16] S. Yu, S. Huang, Y. Pan, and H. Wu. An easy dominant color extraction and edge valley histogram for image retrieval. In *Computer Symposium (ICS), 2010 International*, pages 159 –164, 2010.

Umbralización de imágenes utilizando funciones de agrupamiento

Aranzazu Jurio[1], **Miguel Pagola**[1], **Luis Gonzalez-Jaime**[2], **Humberto Bustince**[1], **Ana Pradera**[3]

[1]*Dpto. Automática y Computación, Universidad Publica de Navarra, España,*
{aranzazu.jurio,miguel.pagola,bustince}@unavarra.es
[2]*Applied Mathematics and Computer Science, Ghent University, Bélgica,* luis.gonzalez@UGent.be
[3]*Departamento de Ciencias de la Computación, Universidad Rey Juan Carlos, España,* ana.pradera@urjc.es

Resumen

En este trabajo presentamos un algoritmo de umbralización de imágenes en escala de grises. Nuestra propuesta utiliza funciones de agrupamiento para encontrar el mejor umbral. Estas funciones son capaces de representar cuánto pertenece una intensidad de gris al fondo o al objeto de la imagen, por lo que el mejor umbral viene dado por el mayor agrupamiento

Palabras Clave: Umbralización de imágenes, funciones de agrupamiento.

1 Introducción

Una de las técnicas más utilizadas en la segmentación de imágenes es la umbralización o segmentación por niveles de gris [7],[8],[6]. En umbralización, los objetos presentes en las imágenes están caracterizados únicamente por el nivel de intensidad de gris de cada pixel. El proceso consiste en encontrar un umbral *t* tal que los píxeles con intensidades menores o iguales que dicho umbral pertenecen al fondo y los píxeles cuyas intensidades sean mayores que *t* pertenecen al objeto, o viceversa [3]. La ventaja de este tipo de métodos frente a otros algoritmos de segmentación es su sencillez y su bajo coste computacional. Por esta razón, los métodos de umbralización pueden utilizarse como un paso previo para otros procesos de segmentación más costosos.

En este trabajo presentamos un nuevo algoritmo de umbralización que es una generalización de otros métodos de umbralización difusos [3], [5]. Nuestra propuesta se basa en la construcción, para cada nivel de intensidad, de dos conjuntos difusos (Q_{B_t} y Q_{O_t}) que representan la pertenencia de cada nivel de intensidad de gris al fondo ($\mu_{Q_{B_t}}(q)$) y al objeto ($\mu_{Q_{O_t}}(q)$) respectivamente. El objetivo es encontrar un umbral a través del cual la pertenencia de cada nivel de intensidad al fondo o al objeto sea máxima ($\mu_{Q_{B_t}} = 1$ o $\mu_{Q_{O_t}}$), y por tanto estemos seguros de que esos píxeles pertenecen al fondo o al objeto de la imagen.

Para resolver el objetivo del trabajo, proponemos la definición de función de agrupamiento como una función de agregación bivalorada que obtiene el valor máximo si y sólo si alguno de sus argumentos vale 1. En este trabajo estudiamos las axiomatización de estas funciones y estudiamos algunos métodos para su construcción, además de establecer una relación entre este tipo de funciones y las funciones de solapamiento [4].

El resto del trabajo está organizado de la siguiente manera. Comenzamos recordando algunos conceptos preliminares en la Sección 2. En la Sección 3 estudiamos las funciones de agrupamiento, así como su relación con las funciones de solapamiento y algunos métodos de construcción. En la Sección 4 presentamos nuestro algoritmo de umbralización de imágenes, y en la Sección 5 mostramos un ejemplo ilustrativo. Finalizamos con algunas conclusiones en la Sección 6.

2 Preliminares

Una negación estricta [9] es una función contínua y estrictamente decreciente $N : [0,1]^2 \to [0,1]$ tal que $N(0) = 1$ y $N(1) = 0$. Una negación fuerte es una negación estricta que además es involutiva, es decir, $N(N(x)) = x$ para todo $x \in [0,1]$.

Una norma triangular (t-norma) es una función de agregación bivalorada, simétrica y asociativa $T : [0,1]^2 \to [0,1]$ tal que $T(x,1) = x$ para todo $x \in [0,1]$. Algunos ejemplos de t-normas son la función mínimo $T_M(x,y) = min(x,y)$ o la función producto $T_P(x,y) = x \cdot y$. Una conorma triangular (t-conorma) es una función de agregación bivalorada, simétrica y asociativa $S : [0,1]^2 \to [0,1]$ tal que $S(x,0) = x$ para todo $x \in [0,1]$. Algunos ejemplos de t-conormas son la función máximo $S_M(x,y) = max(x,y)$ o la suma probabilística $S_P(x,y) = x + y - x \cdot y$ [1].

En este trabajo utilizamos las funciones de equivalencia restringida (REF) para construir los conjuntos difusos asoci-

ados a una imagen [2],[3].

Definición 1 *Una función* $REF : [0,1]^2 \to [0,1]$ *se dice función de equivalencia restringida si satisface las siguientes condiciones:*

1. $REF(x,y) = REF(y,x)$ *para todo* $x,y \in [0,1]$*;*
2. $REF(x,y) = 1$ *si y sólo si* $x = y$*;*
3. $REF(x,y) = 0$ *si y sólo si* $x = 1$ *e* $y = 0$ *o* $x = 0$ *e* $y = 1$*;*
4. $REF(x,y) = REF(N(x),N(y))$ *para todo* $x,y \in [0,1]$*, siendo N una negación fuerte;*
5. *si* $x \leq y \leq z$ *entonces* $REF(x,y) \geq REF(x,z)$ *y* $REF(y,z) \geq REF(x,z)$*, para todo* $x,y,z \in [0,1]$*.*

3 Funciones de agrupamiento

Definición 2 *Una función* $G_G : [0,1]^2 \to [0,1]$ *es una función de agrupamiento si satisface las siguientes condiciones:*

(G_G1*)* $G_G(x,y) = G_G(y,x)$ *para todo* $x,y \in [0,1]$*;*

(G_G2*)* $G_G(x,y) = 0$ *si y sólo si* $x = y = 0$*;*

(G_G3*)* $G_G(x,y) = 1$ *si y sólo si* $x = 1$ *o* $y = 1$*;*

(G_G4*)* G_G *es no decreciente;*

(G_G5*)* G_G *es continua.*

Obsérvese que una función de agrupamiento es un caso particular de función de agregación binaria.

Podemos encontrar una relación entre las funciones de agrupamiento y las funciones de solapamiento definidas en [4]. Utilizamos este hecho para presentar diferentes métodos de construcción de funciones de agrupamiento.

3.1 Funciones de solapamiento

Definición 3 *Una función* $G_O : [0,1]^2 \to [0,1]$ *es una función de solapamiento si satisface las siguientes condiciones:*

(G_O1*)* G_O *es simétrica.*

(G_O2*)* $G_O(x,y) = 0$ *si y sólo si* $xy = 0$*.*

(G_O3*)* $G_O(x,y) = 1$ *si y sólo si* $xy = 1$*.*

(G_O4*)* G_O *es no decreciente.*

(G_O5*)* G_O *es contínua.*

Teorema 1 *Sea* G_O *una función de solapamiento y sea N una negación estricta. Entonces*

$$G_G(x,y) = N(G_O(N(x),N(y))) \tag{1}$$

es una función de agrupamiento. Recíprocamente, tenemos que

$$G_O(x,y) = N(G_G(N(x),N(y))) \tag{2}$$

es una función de solapamiento.

Demostración 1 *(*G_G1*),* *(*G_G4*)* *y* *(*G_G5*)* *son directas.* *(*G_G2*)* $G_G(x,y) = 0 = N(G_O(N(x),N(y)))$ *si y solo si* $G_O(N(x),N(y)) = 1$ *si y solo si* $N(x) = N(y) = 1$ *si y solo si* $x = y = 0$*.* *(*G_G3*),* $G_G(x,y) = 1 = N(G_O(N(x);N(y)))$ *si y solo si* $G_O(N(x),N(y)) = 0$ *si y solo si* $N(x) = 0$ *o* $N(y) = 0$ *si y solo si* $x = 1$ *o* $y = 1$*.*

Basándonos en la relación entre funciones de solapamiento y t-normas, en este trabajo mostramos que las funciones de agrupamiento que cumplen la asociatividad son a su vez t-conormas. Sin embargo, el recíproco de este teorema no es cierto, pues basta considerar t-conormas no contínuas.

Teorema 2 *Sea* G_G *una función de agrupamiento asociativa. Entonces* G_G *es una t-conorma.*

Demostración 2 *Necesitamos probar solamente que* 0 *es el elemento neutro de* G_G*. Dada la continuidad de* G_G *y el hecho de que* $G_G(0,1) = y$ $G_G(0,0) = 0$*, podemos decir que para cualquier* $x \in]0,1[$ *existe un* $y \in]0,1[$ *tal que* $x = G_G(y,0)$*. Entonces* $G_G(x,0) = G_G(G_G(y,0),0) = G_G(y,G_G(0,0)) = G_G(y,0) = x$ *y de una forma similar* $G_G(0,x) = x$*.*

Ejemplo 1 *Una función de agrupamiento asociativa, y que por tanto es una t-conorma, es la función máximo.*

$$G_G(x,y) = max(x,y) \tag{3}$$

Teorema 3 *Sean* $G_G1,\ldots,G_Gm$ *m funciones de agrupamiento y sean* $w_1,\ldots,w_m$ *m pesos no negativos tales que* $\sum_{i=1}^{m} w_i = 1$*. Entonces la suma convexa* $G = \sum_{i=1}^{m} w_i G_i$ *es una función de agrupamiento.*

Demostración 3 *Directo.*

3.2 Construcción de funciones de agrupamiento

En el siguiente teorema estudiamos un método de construcción de funciones de agrupamiento a partir de dos funciones f y h que satisfacen determinadas propiedades.

Teorema 4 *La función* $G_G : [0,1]^2 \to [0,1]$ *es una función de agrupamiento si y sólo si*

$$G_G(x,y) = \frac{f(x,y)}{f(x,y) + h(x,y)} \tag{4}$$

para $f,h : [0,1]^2 \to [0,1]$ *tales que*

1. *f y h son simétricas;*

2. *f es no decreciente y h es no creciente;*

3. $f(x,y) = 0$ *si y sólo si* $x = y = 0$*;*

4. $h(x,y) = 0$ *si y sólo si* $x = 1$ *o* $y = 1$*;*

5. *f y h son funciones contínuas.*

Demostración 4 *Hay que tener en cuenta que* $f(x,y) + h(x,y) \neq 0$ *para todo* $(x,y) \in [0,1]^2$*. Entonces la necesidad es inmediata tomando* $f(x,y) = G_G(x,y)$ *y* $h(x,y) = 1 - G_G(x,y)$*. (Suficiencia)* $(G_G1),(G_G2),(G_G3)$ *y* (G_G5) *son directos.* (G_G4) *Si* $x_1 \leq x_2$ *entonces* $f(x_1,y) \leq f(x_2,y)$ *y* $h(x_2,y) \leq h(x_1,y)$*. Por tanto, tenemos que* $f(x_1,y)h(x_2,y) \leq f(x_2,y)h(x_1,y)$*. Multiplicamos a ambos lados de la igualdad y obtenemos* $f(x_1,y)f(x_2,y) + f(x_1,y)h(x_2,y) \leq f(x_1,y)f(x_2,y) + f(x_2,y)h(x_1,y)$*. De aquí podemos despejar que* $G_G(x_1,y) = \frac{f(x_1,y)}{f(x_1,y)+h(x_1,y)} \leq \frac{f(x_2,y)}{f(x_2,y)+h(x_2,y)} = G_G(x_2,y)$.

Ejemplo 2 *Si tomamos* $f(x,y) = max(x,y)$ *y* $h(x,y) = \sqrt{(1-x)(1-y)}$ *tenemos*

$$G_G(x,y) = \frac{max(x,y)}{max(x,y) + \sqrt{(1-x)(1-y)}} \tag{5}$$

Ejemplo 3 *Si tomamos* $f(x,y) = max(x,y)$ *y* $h(x,y) = (1-x)(1-y)$ *tenemos*

$$G_G(x,y) = \frac{max(x,y)}{max(x,y) + (1-x)(1-y)} \tag{6}$$

Ejemplo 4 *Si tomamos* $f(x,y) = 1 - \sqrt{(1-x)(1-y)}$ *y* $h(x,y) = min((1-x),(1-y))$ *tenemos*

$$G_G(x,y) = \frac{1 - \sqrt{(1-x)(1-y)}}{1 - \sqrt{(1-x)(1-y)} + min((1-x),(1-y))} \tag{7}$$

4 Algoritmo de umbralización basado en funciones de agrupamiento

En este trabajo proponemos la utilización de funciones de agrupamiento como métrica para el cálculo del umbral óptimo de una imagen. Para ello construimos, para cada nivel de intensidad, un conjunto difuso asociado al fondo y otro conjunto difuso asociado al objeto de la imagen. Aplicando una combinación convexa de diferentes funciones de agrupamiento a estos conjuntos, elegimos el umbral adecuado para cada imagen. El esquema del algoritmo que proponemos se puede ver en Algoritmo 1.

Algoritmo 1 Algoritmo de umbralización

1: **for** $t = \{0,1,\ldots,L-1\}$ (Para cada nivel de intensidad) **do**
2: Construir un conjunto difuso asociado con el fondo de la imagen (Q_{B_t}).
3: Construir un conjunto difuso asociado con el objeto de la imagen (Q_{O_t}).
4: **for** $q = \{0,1,\ldots,L-1\}$ (Para cada nivel de intensidad) **do**
5: Calcular varias funciones de agrupamiento entre $Q_{B_t}(q)$ y $Q_{O_t}(q)$.
6: Calcular la combinación convexa de las funciones anteriores obteniendo una nueva función de agrupamiento $G_{G_{comb}}(q)$.
7: **end for**
8: Calcular la suma ponderada de los agrupamientos anteriores

$$\sum_{q=0}^{L-1} G_{G_{comb}}(\mu_{B_t}(q), \mu_{O_t}(q)) \cdot h(q)$$

donde $h(q)$ es la cantidad de píxeles cuya intensidad es q.
9: **end for**
10: Tomar como mejor umbral $t*$ el asociado a la mayor suma de agrupamientos:

$$t* = \arg\max_t \sum_{q=0}^{L-1} G_{G_{comb}}(\mu_{B_t}(q), \mu_{O_t}(q)) \cdot h(q)$$

4.1 Construcción de conjuntos difusos asociados a la imagen

Cuando umbralizamos una imagen mediante un único umbral, partimos de la idea de que la imagen está dividida en dos zonas, y por tanto podemos diferenciar un objeto y el fondo de la imagen. Siguiendo el estudio realizado en [3], en este trabajo construimos dos conjuntos difusos (Q_{B_t} asociado al fondo y Q_{O_t} asociado al objeto) a partir de funciones de equivalencia restringida, teniendo en cuenta el siguiente razonamiento: cuanto más parecido sea un nivel de intensidad (q) a la media de intensidades del fondo (análogamente con el objeto) mayor será el valor de pertenencia de esa intensidad al conjunto difuso que representa al fondo (objeto).

Fijado un nivel de intensidad t, comenzamos calculando la media de intensidades del fondo ($m_B(t)$) y del objeto ($m_O(t)$) mediante las siguientes expresiones:

$$m_B(t) = \frac{\sum_{q=0}^{t} q \cdot h(q)}{\sum_{q=0}^{t} h(q)} \quad m_O(t) = \frac{\sum_{q=t+1}^{L-1} q \cdot h(q)}{\sum_{q=t+1}^{L-1} h(q)}$$

Sea REF una función de equivalencia restringida, construimos los conjuntos difusos Q_{B_t} y Q_{O_t} con las siguientes

funciones de pertenencia, para cada nivel de intensidad $q = 0, 1, \ldots, L-1$:

$$\mu_{Q_{B_t}}(q) = REF(\frac{q}{L-1}, \frac{m_B(t)}{L-1}) \quad (8)$$

$$\mu_{Q_{O_t}}(q) = REF(\frac{q}{L-1}, \frac{m_O(t)}{L-1}) \quad (9)$$

Podemos comprobar que, con este método de construcción y debido a la propiedad (2) de la Definición 1, un nivel de intensidad sólo va a tener pertenencia máxima al conjunto del fondo (objeto) si su intensidad coincide con la media del fondo (objeto).

- $\mu_{Q_{B_t}}(q) = 1$ si y solo si $q = m_B(t)$.
- $\mu_{Q_{O_t}}(q) = 1$ si y solo si $q = m_O(t)$.

4.2 Cálculo del agrupamiento

Para calcular un valor de agrupamiento asociado a cada posible umbral t, comenzamos utilizando n funciones de agrupamiento distintas ($G_G1, G_G2, \ldots, G_Gn$). Una función de agrupamiento toma dos valores y calcula el nivel de agrupamiento que existe entre ellos. En este caso, calculamos el agrupamiento que existe para cada nivel de intensidad entre la pertenencia al fondo y la pertenencia al objeto.

Utilizando el resultado obtenido en el Teorema 3, combinamos las n funciones de agrupamiento previamente calculadas. De esta forma obtenemos una nueva función de agrupamiento que, experimentalmente, mejora siempre al resultado obtenido con la peor de las funciones elegidas. Este paso nos ayuda a solventar el posible inconveniente de haber elegido una función de agrupamiento que no se adecue bien al problema, y que por tanto no vaya a obtener buenos resultados.

Una vez que tenemos un valor único para el agrupamiento en cada nivel de intensidad, obtenemos su suma. Este es el valor de agrupamiento asociado al umbral t.

4.3 Elección del agrupamiento máximo

Cada uno de los posibles umbrales $t = \{0, 1, \ldots, L-1\}$ tiene asociado un valor de agrupamiento, calculado como la suma de una función de agrupamiento en diferentes puntos. Para obtener el mejor umbral, nos quedamos con aquel asociado al mayor valor de agrupamiento. Elegimos el mayor valor teniendo en cuenta las propiedades de las funciones de agrupamiento. De esta forma, el valor de la suma será máximo si $G_G(\mu_{B_t}(q), \mu_{O_t}(q)) = 1$ para todo $q = \{0, 1, \ldots, L-1\}$. Por la propiedad ($G_G2$) de la Definición 2 este hecho se puede dar en dos casos:

- $\mu_{B_t}(q) = 1$, con lo que $q = m_B(t)$. En este caso estamos completamente seguros de que los píxeles cuya intensidad es q pertenecen al fondo, ya que coinciden exactamente con su media.
- $\mu_{O_t}(q) = 1$, con lo que $q = m_O(t)$. En este caso estamos completamente seguros de que los píxeles cuya intensidad es q pertenecen al objeto, ya que coinciden exactamente con su media.

En este sentido, eligiendo el mayor valor de agrupamiento estamos eligiendo el umbral con el que todos los píxeles cuya intensidad es menor o igual que el umbral están más cercanos a la intensidad media del fondo (objeto) y los píxeles cuya intensidad es mayor que el umbral están más cercanos a la media del objeto (fondo).

5 Ejemplo ilustrativo

En esta sección mostramos el comportamiento del algoritmo que proponemos evaluándolo sobre 10 imágenes en escala de grises (ver Figura 1). Para comprobar la calidad de las segmentaciones obtenidas, las compararemos con una segmentación ideal, realizada manualmente (ver Figura 2).

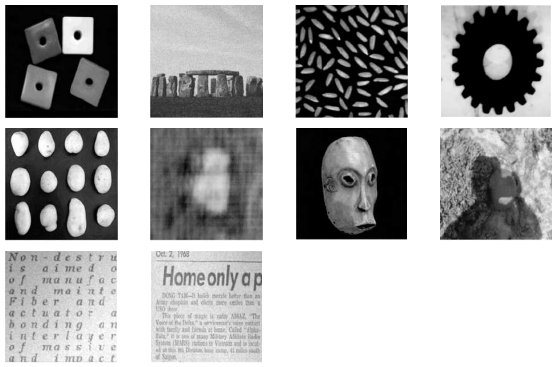


Figura 1: Imágenes originales.

Figura 2: Segmentaciones ideales realizadas manualmente.

En este ejemplo vamos a utilizar 4 funciones de agrupamiento, que se corresponden con las ecuaciones (3), (5), (6) y (7).

- $G_G1(x,y) = max(x,y)$
- $G_G2(x,y) = \frac{max(x,y)}{max(x,y)+\sqrt{(1-x)(1-y)}}$
- $G_G3(x,y) = \frac{max(x,y)}{max(x,y)+(1-x)(1-y)}$
- $G_G4(x,y) = \frac{1-\sqrt{(1-x)(1-y)}}{1-\sqrt{(1-x)(1-y)}+min((1-x),(1-y))}$

Para calcular la combinación convexa de agrupamientos utilizamos pesos iguales. En este caso, cada uno de los cuatro pesos tendrá un valor de 0.25. Si estuviésemos trabajando con un tipo de imágenes específico, por ejemplo imágenes de células, de resonancia magnética, etc., es probable que unas funciones de agrupamiento se adaptasen mejor al problema que otras. En estos casos es recomendable utilizar diferentes valores, aplicando un mayor peso a las funciones que den mejor resultado (por ejemplo, utilizando operadores OWA [10]).

Figura 3: Segmentaciones obtenidas por nuestro método.

En la Figura 3 podemos ver la segmentación obtenida por nuestro método para cada una de las imágenes.

A continuación realizamos un estudio de las diferentes segmentaciones obtenidas con cada una de las 4 funciones de agrupamiento propuestas, en relación con el resultado final obtenido con la combinación convexa de todas ellas. Para ello, en las Figuras 4 y 5 vemos los diferentes resultados para las dos primeras imagenes. En la primera fila podemos ver la segmentación obtenida por cada una de las cuatro funciones de agrupamiento individualmente. En la segunda fila se muestran en blanco los píxeles que han sido bien clasificados por la función de agrupamiento individual, pero que al realizar la combinación convexa de los agrupamientos, se han clasificado incorrectamente. Por contra, en la tercera fila se muestran en blanco los píxeles que en la segmentación realizada por una sola función de agrupamiento se habían clasificado incorrectamente, pero que en la segmentación conjunta se han clasificado bien.

Como podemos ver, dependiendo de la imagen que queramos segmentar, algunas funciones de agrupamiento se

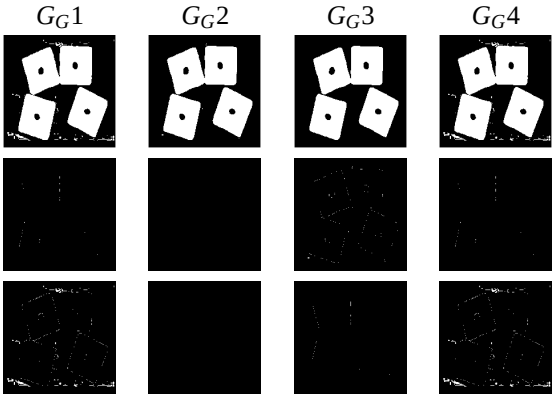


Figura 4: Segmentaciones obtenidas por las cuatro funciones de agrupamiento (primera fila). En la segunda fila están los píxeles bien clasificados por cada función de agrupamiento individual que se clasifican incorrectamente en la combinación convexa. En la tercera fila se muestran los que se clasifican incorrectamente en las segmentaciones individuales pero correctamente en la combinación convexa.

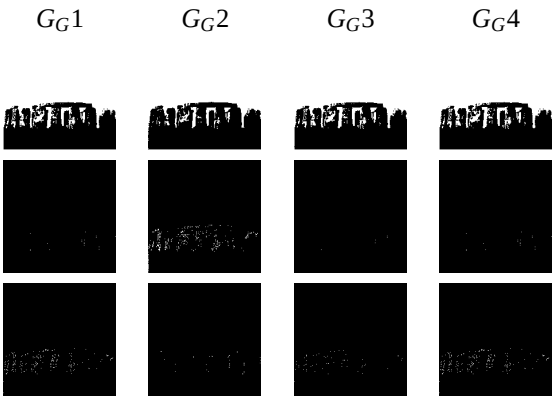


Figura 5: Segmentaciones obtenidas por las cuatro funciones de agrupamiento (primera fila). En la segunda fila están los píxeles bien clasificados por cada función de agrupamiento individual que se clasifican incorrectamente en la combinación convexa. En la tercera fila se muestran los que se clasifican incorrectamente en las segmentaciones individuales pero correctamente en la combinación convexa.

comportan mejor que otras. Este hecho queda demostrado en la cantidad de píxeles bien y mal clasificados con respecto a la combinación convexa. Completamos este estudio viendo la cantidad de píxeles bien clasificados que obtenemos en cada caso. En la Tabla 1 vemos los diferentes umbrales que obtenemos con los cinco agrupamientos (los cuatro propuestos y la combinación de todos ellos). En la Tabla 2 se muestran los porcentajes de píxeles bien clasificados que se obtienen con cada uno de estos umbrales.

Vemos experimentalmente que la combinación convexa de agrupamientos obtiene siempre un umbral intermedio entre los umbrales obtenidos por las funciones de agrupamiento

Tabla 1: Umbrales obtenidos con las diferentes funciones de agrupamiento para cada imagen.

Imagen	G_G1	G_G2	G_G3	G_G4	$G_{G_{comb}}$
1	7	13	18	7	13
4	131	144	134	131	135
2	38	11	45	39	27
3	57	7	71	69	44
5	165	181	158	163	175
6	142	99	114	139	127
7	17	3	49	26	9
8	121	121	127	121	121
9	206	210	202	206	207
10	205	212	202	207	207

Tabla 2: Porcentaje de píxeles bien clasificados con las diferentes funciones de agrupamiento para cada imagen.

Imagen	G_G1	G_G2	G_G3	G_G4	$G_{G_{comb}}$
1	95.31	96.67	96.90	95.31	96.67
4	95.23	96.67	95.59	95.23	95.73
2	92.68	90.89	92.77	92.70	92.24
3	98.35	97.51	98.40	98.37	98.26
5	94.59	92.80	95.19	94.83	93.70
6	96.89	68.41	88.29	96.56	94.62
7	96.83	92.25	96.72	96.96	96.58
8	89.57	89.57	87.93	89.57	89.57
9	56.55	52.08	61.32	56.55	55.54
10	70.07	62.79	72.66	68.26	68.26

por separado, y que el porcentaje de píxeles bien clasificados de la combinación también es siempre intermedio. De esta forma, obtenemos que con nuestro algoritmo no obtenemos el mejor umbral que se puede alcanzar mediante funciones de agrupamiento. Sin embargo, siempre obtenemos un umbral mejor que el obtenido por la peor función de agrupamiento utilizada. Con este hecho podemos solventar el problema de haber elegido una función de agrupamiento que no se adecúe correctamente a la imagen tratada, y aun así obtener una buena solución.

6 Conclusiones

En este trabajo hemos presentado un algoritmo de umbralización de imágenes en escala de grises que se basa en las funciones de agrupamiento. Estas funciones, aplicadas a nuestro problema, miden cuánto pertenece un nivel de intensidad al fondo o al objeto presente en la imagen. De esta forma, escogemos el umbral que obtenga un mayor agrupamiento para segmentar la imagen. Una de las ventajas de este método es que evitamos la compleja tarea de elegir una función de agrupamiento que funcione bien para cada imagen, utilizando una combinación convexa de varias de ellas.

Agradecimientos

Este trabajo ha sido parcialmente apoyado por el servicio de Investigación de la Universidad Pública de Navarra, por la beca TIN2010-15055 y por la Comision Europea bajo el acuerdo de subvencion No. 238819 (MIBISOC Marie Curie ITN).

Referencias

[1] G. Beliakov, A. Pradera, T. Calvo: Aggregation Functions: A Guide for Practitioners. *Studies in Fuzziness and Soft Computing* 221, 2007.

[2] H. Bustince, E. Barrenechea, M. Pagola: Restricted equivalence functions. *Fuzzy Sets and Systems*. 157, pp. 2333–2346, 2006.

[3] H. Bustince, E. Barrenechea, M. Pagola: Image thresholding using restricted equivalence functions and maximizing the measure of similarity. *Fuzzy Sets and Systems* 158, pp. 496–516, 2007.

[4] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna: Overlap functions. *Nonlinear Analysis - Theory, Methods & Applications*. 78, pp. 1488–1499, 2010.

[5] L.K. Huang, M.J.J. Wang: Image thresholding by minimizing the measure of fuzziness. *Pattern Recognition* 28, pp. 41–51, 1995.

[6] N. Otsu: A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*. 9, pp. 62–66, 1979.

[7] P.K. Sahoo, S. Soltani, A.K.C. Wong, Y.C. Chen: A survey of thresholding techniques. *Computer Vision Graphics and Image Processing* 41, pp. 233–260, 1988.

[8] M. Sezgin, B. Sankur: Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging* 13, pp. 146–168, 2004.

[9] E. Trillas: Sobre funciones de negación en la teoría de conjuntos difusos. *Stochastica III* 1, pp. 47–59, 1979.

[10] R.R. Yager: On ordered weighted averaging aggregation operators in multicriteria decision-making. *IEEE Transantions on Systems, Man and Cybernetics* 18, pp. 183-190, 1988.

Clasificación no supervisada y jerárquica basada en coloración en grafos: una aplicación a imágenes astronómicas

Daniel Gómez[1], **Edwin Zarrazola**[2], **Javier Montero**[2], **Javier Yáñez**[2]

[1]*Escuela de Estadística. Universidad Complutense de Madrid. Av. Puerta de Hierro s/n 28040 Madrid,*
dagomez@estad.ucm.es

[2]*Facultad de Matemáticas. Universidad Complutense de Madrid. Plaza de Ciencias s/n 28040 Madrid,*
edwzar@mat.ucm.es, monty@mat.ucm.es, jayage@mat.ucm.es

Resumen

En este trabajo se propone un algoritmo de clasificación no supervisada y jerárquica sobre un conjunto de objetos modelizados mediante un grafo. La salida de este algoritmo muestra la evolución de cómo los grupos se van rompiendo en la red de manera divisiva. Así pues, el procedimiento aquí estudiado muestra de manera evolutiva la sucesión de cortes en la red desde la situación inicial en la que todos los nodos de la red están conectados (en el mismo grupo) hasta la última iteración en la que cada nodo representa un grupo por sí mismo. Esta salida puede ser vista como un agrupamiento difuso en el que para cada alfa corte se tiene una partición nítida de la red. Aplicaremos esta metodología a un problema de segmentación en imágenes astronómicas, en la que los nodos del grafo son los píxeles y las aristas unen píxeles adyacentes, obteniendo buenos resultados desde el punto de vista computacional y desde la calidad de la segmentación realizada.

Palabras Clave: Clasificación no supervisada, Clustering borroso, Segmentación en Redes, Imágenes Astronómicas.

1 Introducción

El problema de la segmentación o clasificación no supervisada en redes ha tenido una importancia creciente durante los últimos años (ver por ejemplo [1, 9]). Al igual que pasa con cualquier problema de clasificación, los métodos que permiten segmentar un conjunto de objetos que están relacionados entre si, pueden dividirse entre jerarquizados y no jerarquizados. En los métodos no jerarquizados el objetivo final de la segmentación es obtener una partición final del conjunto de datos (que este caso serán pixels, ver por ejemplo [1, 10] para más información sobre diferentes técnicas y posibles aplicaciones), mientras que en los métodos jerarquizados (cuya visualización suele darse mediante un dendograma) el objetivo final es mostrar la evolución de como los grupos van formándose (en los métodos aglomerativos) o dividiéndose (en los métodos divisivos) desde la primera iteración hasta la última.

En este trabajo, nos centraremos en los métodos jerárquicos divisivos para aquellas situaciones en las que los objetos que van a agruparse se relacionan mediante un grafo. El algoritmo aquí desarrollado comparte la misma filosofía que el algoritmo presentado en [5], extendido al contexto borroso en [4]. No obstante en este trabajo se reduce de manera significativa el tiempo computacional de previas aproximaciones sin reducir la calidad de las segmentaciones obtenidas.

El artículo está organizado de la siguiente manera. En primer lugar, en la sección 2, se presenta al algoritmo básico de agrupamiento o segmentación jerarquizado en redes. En la sección 3, se presenta el núcleo de la metodología (subsección 3.1), que consiste en un proceso iterativo de coloración binario basado en la secuencia de una serie de valores fijados. La selección apropiada de dichos valores (subsección 3.2) define de manera univoca el algoritmo de agrupamiento jerarquizado (subsección 3.3). El análisis de la complejidad computacional es analizado en la subsección 3.5. Finalmente se aplica el algoritmo aquí desarrollado al problema de la segmentación de imágenes astronómicas dando una nueva perspectiva al problema de segmentación de imágenes en la sección 4. Este trabajo acaba con algunos comentarios y observaciones finales.

2 Preliminares

Sea $V = \{1,2,\ldots,n\}$ un conjunto finito de elementos que deber ser agrupados. Sea $E = \{\{i,j\} \mid i,j \in V\}$ el conjunto de pares no ordenados que representan las relaciones que existen entre el conjunto de nodos V: si dos elementos $i,j \in V$ están relacionados, entonces existe una arista $e = \{i,j\} \in E$; en otro caso, $\{i,j\} \notin E$. De esta manera, un grafo $G = (V,E)$ representa las relaciones existentes entre el conjunto de nodos u objetos que queremos agrupar de

manera unívoca. Para este trabajo se asume que el grafo G es conexo. En otro caso, se podrían obtener resultados similares analizando en paralelo cada una de las componentes conexas del grafo.

Dado $e = \{i, j\} \in E$, denotamos por $d_e \geq 0$ el grado de disimilaridad entre los nodos i y j. Obviamente, a mayor valor del parámetro d_e, más disimilares serán los nodos adyacentes i y j. Aunque no es el principal objetivo de este trabajo cómo obtener esta medida, d_e se debe obtener teniendo en cuenta las especificaciones del problema y las características asociadas a cada uno de los objetos.

En el caso particular del tratamiento de imágenes, las relaciones entre los píxeles serán modelizadas mediante el modelo 4-vecinos (ver por ejemplo sección 3.3) . En resumen, la información disponible para abordar el problema del agrupamiento o segmentación jerarquizado en redes es la siguiente:

$$N = \big(G = (V,E);\ \{d_e\ ,\ e \in E\}\big). \tag{1}$$

3 El algoritmo divisivo

3.1 Proceso de coloreado binario

Partiendo de un nodo $i \in V$ en (1) coloreado de forma binaria $\big(\mathrm{col}(i) \in \{0,1\}\big)$, se puede obtener una coloración binaria de toda la red coloreando los nodos adjacentes j (y así sucesivamente) en función de la medida d_e (con $e = \{i, j\} \in E$) en relación con un valor fijo α de la siguiente manera:

$$\mathrm{col}(j) = \begin{cases} \mathrm{col}(i) & \text{if } d_e < \alpha \\ 1 - \mathrm{col}(i) & \text{si } d_e \geq \alpha \end{cases} \tag{2}$$

para todo $e \in E$.

Aunque este proceso en principio puede parecer consistente, es fácil ver que no lo es, ya que un nodo k puede ser coloreado según (2) de manera diferente en función de la cadena (secuencia de enlaces que unen el nodo coloreado inicialmente) escogida para su coloración. Esta coloración inconsistente podría aparecer cuando existen ciclos en el grafo G, ya que un nodo podría ser coleado desde diferentes cadenas con resultados diversos. En este trabajo se llamarán a estos ciclos, *ciclos inconsistentes*. Para abordar el problema de la inconsistencia un bosque soporte fue definido en [4, 5]. Obviamente, si solamente se tienen en cuenta para el proceso de coloración los arcos pertenecientes a un un bosque soporte la coloración siempre es consistente. El número de ciclos inconsistentes depende del valor elegido α para una iteración particular.

En [4, 5], se diseñó un mecanismo para abordar el problema de los ciclos inconsistentes, pero como se estudió posteriormente en [11], la búsqueda del bosque soporte que produce menos ciclos inconsistentes para un valor fijo de α tiene un coste demasiado elevado en el proceso de segmentación. Para eliminar este problema, en este trabajo se propone un método alternativo en la búsqueda de este bosque teniendo en cuenta la particular topología que tienen la redes que modelizan las imágenes digitales o vía satélite.

Dada la red N mostrada en (1) y una familia de valores α

$$\{\alpha_0 > \alpha_1 > \cdots > \alpha_{K-1} > \alpha_K\} \tag{3}$$

se define la siguiente familia de grafos parciales:

$$\big\{G^i = (V,E^i)\ \mid\ i \in \{0,1,2,\ldots,K\}\big\} \tag{4}$$

donde

$$E^i = \{e \in E\ \mid\ d_e < \alpha_i\}.$$

A partir de un α_0 suficientemente grande, $G^0 = G$ en (4) se construye la familia de grafos parciales $\{F^i = (V,T^i) \mid i \in \{1,\ldots,K\}\}$ siguiendo el siguiente procedimiento en dos fases:

1. De forma similar al algoritmo de Kruskal , se construye el grafo parcial $F^i = (V,T^i)$ de $G^{i-1} \setminus G^i = (V, E^{i-1} - E^i)$ dado por la secuencia decreciente de arcos con valores en el intervalo $[\alpha_i, \alpha_{i-1})$. Estos arcos se incluyen en el árbol soporte mientras no generen un ciclo.

2. Si el grafo parcial F^i es un bosque soporte de $G^{i-1} \setminus G^i = (V, E^{i-1} - E^i)$, paramos; en otro caso, los arcos de E^i (arcos con pesos menores que α_i) son ordenados en orden creciente y son añadidos iterativamente a T^i mientras no formen un ciclo. El proceso acaba cuando F^i es un árbol soporte.

Nótese que las componentes de F^i no se corresponden con árboles soporte de máximo o mínimo peso.

3.2 La elección de los α

Uno de los parámetros básicos en el procedimiento de coloración binaria es la elección de la secuencia de valores de α en cada iteración. Obviamente, diferentes valores de α_i producen diferentes F^i, y consecuentemente diferente coloración y partición de la red. Teniendo en cuenta que en las primeras iteraciones se busca romper aquellas relaciones con disimilaridad alta, la secuencia de estos valores se hace de manera decreciente empezando por aquellas situaciones en los que los objetos son muy disimilares para acabar rompiendo aquellas parejas con valores de disimilaridad bajo. Fijado K, el número de particiones que se desea realizar, se tiene:

$$\alpha_0 > \overline{\alpha} = \alpha_1 > \alpha_2 > \cdots > \alpha_{K-1} \geq \underline{\alpha} > \alpha_K$$

donde $\overline{\alpha} \equiv \max_{e \in E}\{d_e\}$ y $\underline{\alpha} \equiv \min_{e \in E}\{d_e\}$. Obsérvese que F^{i+1} y F^i será coloreado de manera diferente si y solo si

existe un arco $e \in E$ tal que $\alpha_{i+1} \leq d_e < \alpha_i$. De esta manera, el número de diferentes posibles valores de α esta acotado por m, siendo m el número de aristas de la red en estudio, por tanto $K \leq m$. En situaciones reales, si el tamaño del grafo es elevado se verá que es suficiente con asignar a K un valor sensiblemente inferior a m.

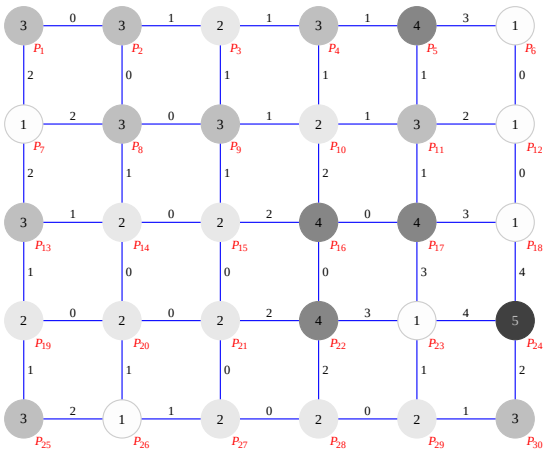
3.3 Agrupamiento jerárquico de la red.

En esta sección se analiza la información obtenida por la aplicación sucesiva de la secuencia de procedimientos de coloración binaria de la red N dada en (1). El procedimiento sucesivo de coloración binario introducida en la anterior subsección da lugar a un proceso de partición jerárquica: la primera partición se obtiene de manera natural por aquellos grupos conectados de igual color. Los nodos separados en una iteración no pueden volver a estar en el mismo grupo en iteraciones posteriores de manera que el proceso define un agrupamiento jerarquizado.

Nótese que las particiones que se obtienen en cada iteración son independientes de la coloración 0 o 1, escogida por el nodo seminal para iniciar el proceso de coloración binaria. Obviamente, la partición empezará a ser no informativa cuando K sea muy grande por lo que si el objetivo final es obtener una partición (y no un agrupamiento jerárquico) el parámetro K no debe ser muy grande. Este parámetro debe ser escogido considerando los diferentes valores que toman d_e, el tamaño de la red, y el tiempo de computación disponible. En cualquier caso, después de que el parámetro K sea elegido, la secuencia de valores α debe satisfacer que: $\alpha_0 > \cdots > \alpha_K$.

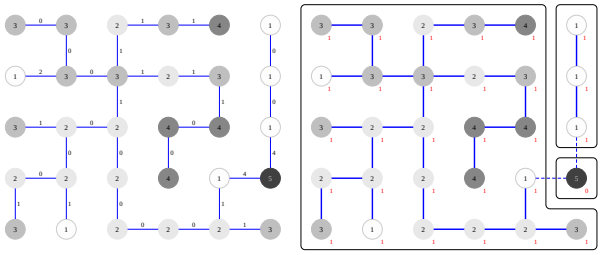
Ejemplo:

Consideremos un ejemplo de 30 pixels como muestra la figura de abajo, donde cada pixel tiene asociado un nivel de intensidad en la escala de grises (1 gris claro, 5 gris oscuro).

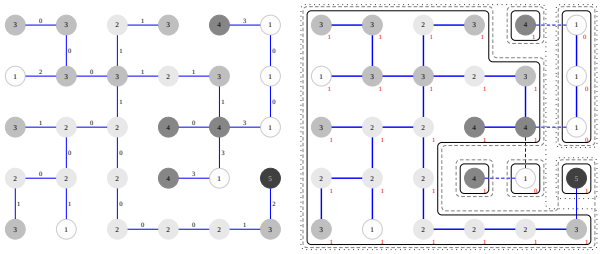


Si tomamos la distancia euclídea entre tonalidades de píxeles adyacentes, la segmentación jerárquica obtenida si $K = 4$ y $\{ \alpha_0 = 5 > \alpha_1 = 4 > \cdots > 1 = \alpha_4 \}$ es:

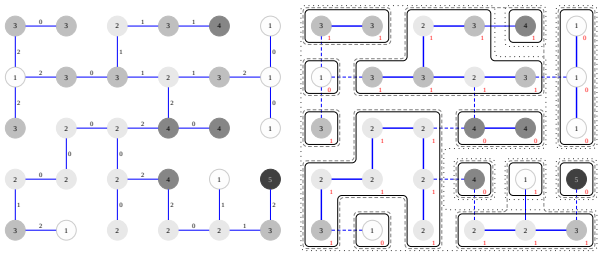
Paso 1: $\alpha = 4$.



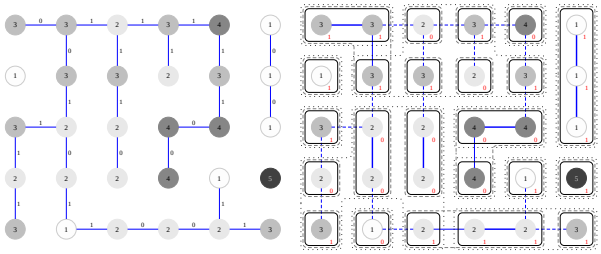
Paso 2: $\alpha = 3$.



Paso 3: $\alpha = 2$.



Paso 4: $\alpha = 1$.



3.4 La clases borrosas borde y no borde y el grafo borroso consistente

De las seccciones anteriores, puede observarse que el algoritmo aquí presentado obtiene, para cada α una partición/segmentación de la red. De manera que las aristas pueden dividirse entre aquéllas que conectan nodos en un mismo grupo y aquéllas que conectan diferentes grupos homogéneos. En base a esto es posible construir las siguientes funciones de pertenencia μ_B, $\mu_{NB} : E \longrightarrow [0,1]$ que corresponderían a las clases *borde* y *no borde* en el proceso de

segmentación:

$$\mu_B(e) = Max\{\alpha \in [0,1] \mid \mathrm{col}_\alpha(i) \neq \mathrm{col}_\alpha(j)\},$$

donde $e = \{i, j\} \in E$. Además,

$$\mu_{NB}(e) = n(\mu_B(e)),$$

siendo *n* una función de negación.

Obsérvese también que desde la clase borrosa $\{\mu_B\}$ es posible construir un grafo borroso consistente (es decir un grafo sin inconsistencias en el proceso de coloración binario). Sea $\widetilde{G} = (V, \mu_B)$ el grafo borroso dado por la función de pertenencia de la clase *borde*. Teniendo en cuenta cómo se ha construido esta clase, se puede garantizar que $\forall \alpha \in [0,1]$ no es posible encontrar una cadena $\{i_1, \ldots, i_k\}$ en $G = (V, E)$ tal que $d_{i_l, i_{l+1}} \leq \alpha$ y $d_{i_1, i_k} \geq \alpha$. Por ejemplo, sea $G = (V = \{1,2,3,4\}, E = \{d_{1,2} = 0.8, d_{1,3} = 0.6, d_{2,4} = 0.7, d_{3,4} = 0.1\})$ un grafo valorado con inconsistencias (por ejemplo para $\alpha = 0.75$). Después del proceso de segmentación jerárquico el grafo borroso consistente dado por la clase borde es: $\widetilde{G} = (V = \{1,2,3,4\}, E = \{\mu_B(1,2) = 0.8, \mu_B(1,3) = 0.6, \mu_B(2,4) = 0.8, \mu_B(3,4) = 0.1\})$ que no tiene ciclos inconsistentes para ningún valor de α.

3.5 Complejidad algorítmica

Teniendo en cuenta que el número de iteraciones *K* esta acotado por *m*, y que el proceso de coloración binaria tiene complejidad polinomial, podemos concluir que el algoritmo obtiene una secuencia de particiones jerárquicas en tiempo polinomial. Además, esta complejidad se asemeja a la del algoritmo de Kruskal para la construcción de un bosque soporte.

4 Aplicación a Imágenes Astronómicas

Las imágenes astronómicas de objetos extendidos proporcionan información crucial sobre las propiedades del medio interestelar. Sin embargo, estas imágenes tienen un fondo débil muy difícil de identificar o tratar. En este marco, las estructuras típicas a estudiar se encuentran en una relación señal-ruido de menos de 3. Estas estructuras suelen ser identificadas después de la inspección visual.

Los algoritmos analizados en este trabajo permiten obtener una buena solución para identificar las estructuras débiles, de una manera no supervisada (necesario en bases de datos de imágenes astronómicas con Terabytes de información). En este sentido, cabe también destacar que el algoritmo aquí presentado podría ser adecuado para identificar estructuras de esa medida sobre varias imágenes sin necesidad de usar procedimientos de normalización, necesarios en el tratamiento de este tipo de imágenes.

El algoritmo aquí desarrollado, tiene en cuenta la borrosidad asociada a las características del fondo, así como la dificultad para evaluar si los fotones detectados en niveles de señal bajos son causadas por el ruido (ruido fotónico, ruido electrónico) o representan alguna señal real. A modo de ejemplo se analiza una imagen astronómica donde $\underline{\alpha} \equiv 0$ y $\overline{\alpha} \equiv 4.7$. Si elegimos, por ejemplo, la siguiente secuencia de umbrales $\alpha \in \{0.6,\ 0.8,\ 1,\ 1.3,\ 1.6\}$; la segmentación jerárquica se muestra en la Figura 1.

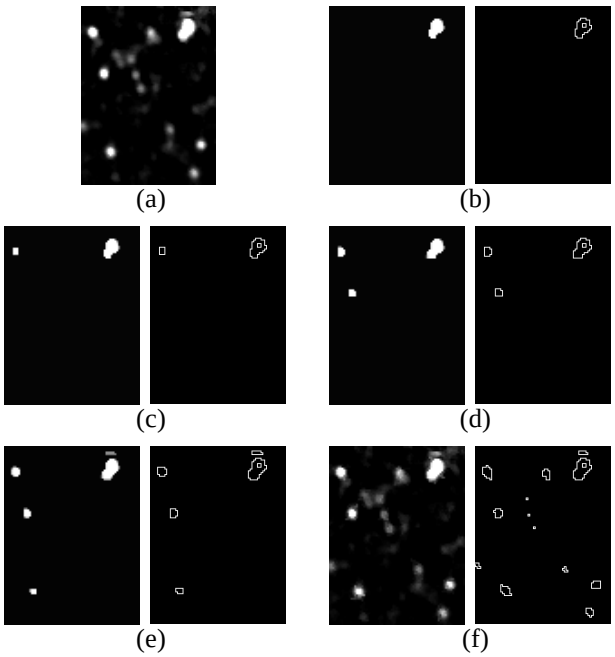


Figura 1: (a): Imagen Original; (b,c,d,e,f): segmentación jerárquica para $\alpha = [1.6,\ 1.3,\ 1,\ 0.8,\ 0.6]$ respectivamente.

Veamos ahora el análisis de la siguiente imagen astronómica dada por la Figura 2, donde $\underline{\alpha} \equiv 0$ and $\overline{\alpha} \equiv 202374840$.

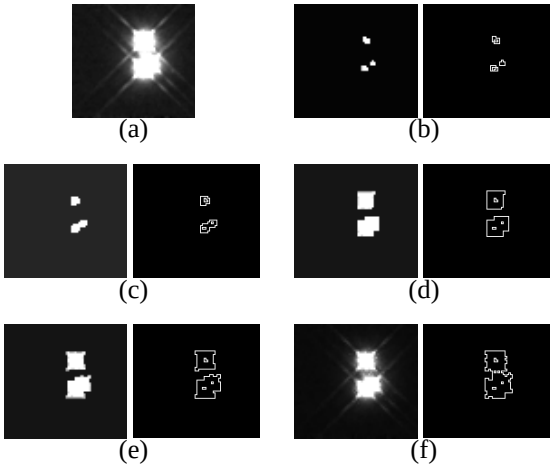


Figura 2: (a): Imagen original; (b,c,d,e,f): segmentación jerárquica con α=$[80,\ 50,\ 10,\ 1,\ 0.5] \times 10^6$ respectivamente.

Ambas imágenes son subconjuntos de imágenes mayores

que han sido descargadas del *Science Archive of the Galaxy Evolution Explorer* (GALEX) ver [7] para más detalles. GALEX es un telescopio espacial que observa el rango ultravioleta del espectro electromagnético. La imagen original cubre un campo circular de visión de 0.6 grados de radio con centro en Ascensión Recta 4h34m8.s0 y Declinación 23d38'33"en la Nube Estelar de Tauro, una región cercana a la formación de estrellas. Se buscan filamentos tenues con el algoritmo. Además, los detectores de GALEX son detectores de conteo de fotones. Por este motivo, no hay lectura de ruido electrónico; el ruido en las imágenes puede ser simplemente modelado por las estadísticas de los fotones.

Las imágenes segmentadas muestran claramente cómo los objetos reales son claramente identificados. Además, la clasificación jerárquica aquí desarrollada permite obtener una graduación de borde y de los objetos, lo que da una visualización más realista de la situación.

5 Comentarios Finales

La principal contribución de este trabajo es la introducción de un algoritmo de agrupamiento jerárquico para redes, donde los elementos están relacionados por medio de un grafo valorado.

El núcleo del algoritmo propuesto es un proceso iterativo de coloración binario, basado en el procedimiento previamente introducido por algunos de los autores en [4, 5]. Este procedimiento se introdujo con el fin de obtener una segmentación de una imagen digital. En el algoritmo que aquí se propone, sin embargo, no hay ninguna restricción sobre la estructura de la red, permitiendo de esta manera la agrupación de cualquier grafo valorado.

Otro logro importante del trabajo aquí presentado es que el algoritmo tiene una complejidad polinomial. Este hecho, junto con la propuesta de una representación automática por medio de un dendograma, abre la posibilidad de un análisis visual de las redes de gran tamaño (véase, por ejemplo, [6]).

También nos gustaría recalcar (como se hizo en [8]) en lo importante que es en la práctica desarrollar instrumentos adecuados para la representación de información con el fin de evitar cajas negras en la toma de decisiones.

Finalmente, nos gustaría recordar (ver por ejemplo [3, 12]) la absoluta necesidad de encontrar metodologías alternativas que permitan determinar la calidad de una clasificación (supervisada o no supervisada) y jerárquica o no jerárquica cuando los objetos que son analizados poseen una estructura subyacente. En el marco de la clasificación no supervisada (y fundamentalmente no jerárquica) se pueden encontrar algunas definiciones para medir la calidad de una partición (ver por ejemplo [1] un resumen general sobre el tema). En este marco aparecen medidas como la modularidad [2] entre otras, que permiten medir la calidad de una partición en una red. No obstante, poco se ha desarrollado cuando el agrupamiento es jerárquico.

Agradecimientos

Esta investigación ha sido realizada en el marco del proyecto TIN2009-07901 Plan Nacional de Investigación.

Referencias

[1] S. Fortunato. "Community detection in graphs". *Physics Reports* **486**, 75–174 (2010).

[2] M. Girvan and M.E.J. Newman. "Community structure in social and biological networks". *Proceedings of the National Academy of Sciences of the USA* **99**, 7821–7826 (2002).

[3] D. Gómez, J. Montero and G. Biging. "Accuracy statistics for judging soft classification". *International Journal of Remote Sensing* **29**(3), 693–709 (2008).

[4] D. Gómez, J. Montero and J. Yánez. "A coloring algorithm for image classification". *Information Sciences* **176**, 3645–3657 (2006).

[5] D. Gómez, J. Montero, J. Yáñez and C. Poidomani. "A graph coloring algorithm approach for image segmentation". *Omega* **35**, 173–183 (2007).

[6] W. Groissboeck, E. Lughofer and S. Thumfart. "Associating visual textures with human perceptions using genetic algorithms". *Information Sciences* **180**(11), 2065–2084 (2010).

[7] C. Martin, T. Barlow, W. Barnhart, L. et al. "The Galaxy Evolution Explorer". *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 336–350 February 2003.

[8] J. Montero, V. López and D. Gómez. "The role of fuzziness in decision making". *Studies in Fuzziness and Soft Computing* **215**, 337–349 (2007).

[9] M.E.J. Newman. "Analysis of weighted networks". *Physical Review E* **70**, 056131 (2004).

[10] F. Tavares-Pereira, J.R. Figueira, V. Mousseau and B. Roy. "Multiple criteria districting problems – The public transportation network pricing system of the Paris region". *Annals of Operations Research* **154**, 69–92 (2007).

[11] J. Yáñez, S. Muñoz and J. Montero. "Graph coloring inconsistencies in image segmentation". *Computer Engineering and Information Sciences* **1**, 435–440 (2008).

[12] S. Yue, J.S. Wang, T. Wu and H. Wang. “A new separation measure for improving the effectiveness of validity indices. *Information Sciences* **180**(5), 748–764 (2010).

EVALUACIÓN DE ALGORITMOS DE CLUSTERING DIFUSO ESPACIALES PARA EL USO DE VECINDADES 3D EN EL ANÁLISIS DE ESTUDIOS TACs DE ALTA RESOLUCIÓN

Alberto Rey, **Alfonso Castro**, **Bernardino Arcay**

Facultad de Informática, Campus de Elviña, Universidade da Coruña,
{alberto.rey,alfonso.castro,bernardino.arcay}@udc.es

Resumen

La detección de nódulos pulmonares ha sido y es uno de los tópicos más estudiados dentro del análisis de imágenes médicas, inicialmente mediante el uso de imágenes radiológicas y en la actualidad a partir de estudios adquiridos mediante tomografía axial computarizada (TAC).

Con el objetivo de asistir al radiólogo en el examen de los estudios, se están desarrollando múltiples sistemas de ayuda al diagnóstico (CAD). Una de sus principales fases es la detección de regiones de interés (ROIs). Para ello es común aplicar diversos algoritmos de clustering difuso entre los que destacan por los resultados obtenidos los kernelized espaciales. Estos algoritmos fueron desarrollados para trabajar con vecindades 2D. Sin embargo, los estudios TAC permiten el análisis de la estructura representada en 3D. En este sentido, este trabajo presenta los resultados obtenidos modificando dos algoritmos de clustering difuso espaciales para usar vecindades 3D, comparando su uso con vecindades 2D y con el Fuzzy C-Means que actualmente es considerado el algoritmo de referencia.

Palabras Clave: Clustering Difuso, Análisis de Imagen Médica, CAD.

1 INTRODUCCIÓN

La gran complejidad de la estructura pulmonar unido a la alta incidencia del carcinoma pulmonar en la sociedad moderna y sus altos índices de mortalidad, ha fomentado la investigación en la detección de nódulos pulmonares. La obtención de un diagnóstico precoz se convierte en un objetivo primordial dado que permite elevar en ciertos casos la tasa de supervivencia en más del 50 % [1].

La aparición de técnicas de adquisición de imágenes como la Tomografía Axial Computarizada (TAC) ha propiciado un gran avance en la detección y diagnóstico de nódulos pulmonares en su estado primario (pueden ser menores de 3mm), obteniendo de una forma no invasiva imágenes de alta definición de estas estructuras. Sin embargo, cada estudio adquirido mediante esta técnica puede contener más de 500 imágenes, con lo que su análisis supone un gran esfuerzo en tiempo para los radiólogos.

Para tratar de superar estas dificultades se están desarrollando numerosos sistemas CAD [10], para la detección de nódulos pulmonares, con el fin de proporcionar una ayuda al diagnóstico efectuado por los radiólogos y mejorar la eficiencia del proceso de análisis de las imágenes.

Una de las principales fases de estos sistemas consiste en la detección de regiones de interés que pueden ser tratadas como posibles nódulos. En este sentido, hemos realizado diferentes trabajos [4] [5], en los que se analizan diferentes algoritmos de clustering difuso para la detección de ROIs en cortes de TACs pulmonares. En ellos, se observa que aquellos algoritmos que incorporan información espacial con funciones kernel son los que ofrecen mejores resultados, permitiendo segmentar de forma más óptima el espacio de entrada.

Por otra parte, los algoritmos presentes en la literatura que se han utilizado en el análisis de imagen médica hacen uso de vecindades 2D [7, 8, 11]. Sin embargo, la estructura anatómica puede ser representada y reconstruida en 3D a partir del TAC, permitiendo realizar el análisis a nivel de estudio en vez de corte y de esta forma disponer de más información durante la clasificación y una mejor aproximación a la distribución real de los píxeles del nódulo.

El objetivo de este trabajo consiste en la modificación de estos algoritmos incorporando vecindades 3D de diferentes tamaños, con el fin de analizar y comparar su comportamiento respecto a vecindades 2D. Se busca así cuantificar si se produce una mejora en la localización de nódulos pulmonares trabajando con el estudio completo.

2 MATERIALES Y MÉTODOS

En la realización de este trabajo hemos seleccionado un conjunto de prueba consistente en nueve estudios de TACs pulmonares de alta resolución obtenidos de la base de datos pública Lung Image Database Consortium (LIDC) [2]. El desarrollo de este proyecto ha sido promovido por el National Institute of Cancer de los Estados Unidos, además de cinco universidades norteamericanas. Su principal objetivo es proporcionar un repositorio público de referencia de imagen médica pulmonar con el fin de desarrollar, entrenar y probar sistemas de ayuda al diagnóstico para la detección de nódulos pulmonares.

Esta base de datos contiene numerosos estudios en los que se detallan las características de los posibles nódulos apreciados a través de un fichero XML, creado mediante un proceso de anotación entre diversos radiólogos. Además, se especifican los puntos que conforman el contorno del nódulo apreciado. Las imágenes están almacenadas de acuerdo al estándar DICOM con un tamaño de 512x512, con un tamaño de píxel variando entre 0.5 a 0.8 mm y una escala de grises de 16 bits en Hounsfiled Units (HU).

A continuación se describen los algoritmos utilizados en este trabajo.

2.1 Fuzzy C-Means (FCM)

El Fuzzy C-Means fue introducido por Bezdek [3] y permite dividir en conjuntos difusos un colección finita de elementos $X = \{x_1, ..., x_n\}$. Los principales pasos del algoritmo son:

1. Cálculo de la pertenencia de cada elemento a cada cluster:

$$u_k(i,j) = \left(\sum \frac{\|y(i,j) - vk\|^{\frac{2}{m-1}}}{\|y(i,j) - vj\|} \right)^{-1} \quad (1)$$

2. Cálculo de los nuevos centroides de la imagen:

$$v_k = \frac{\sum_{i,j} u_k(i,j)^m y(i,j)}{\sum_{i,j} u_k(i,j)^m}, k = \{1, \ldots, c\} \quad (2)$$

3. Si el error permanece por debajo de un determinado umbral, parar. En otro caso, volver al paso 1.

2.2 Spatial Kernelized Fuzzy C-Means (SKFCM)

Este algoritmo, propuesto en [11], propone una modificación del algoritmo KFCM, presentado en el mismo trabajo, en la que se incorpora un término de penalización con información espacial referente a la vecindad. Este factor permite controlar la sensibilidad al ruido y aproximar la salida hacia una solución más homogénea.

Por otro parte, el uso de una función de kernel permite transformar el espacio de entrada de baja dimensionalidad en uno de mayor dimensionalidad, en el que los problemas no lineales complejos pueden ser tratados con una mayor eficiencia [6]. En este trabajo se ha optado por representar los resultados eligiendo la función de kernel Gaussian Radial Basis(GRBF).

$$K(x,y) = exp\left(-\frac{\|x-y\|^2}{\sigma^2} \right) \quad (3)$$

Derivando su función objetivo se obtienen dos condiciones que minimizan dicha función y permiten crear un algoritmo iterativo para este propósito. En la inicialización del algoritmo se deben establecer los parámetros: c (numero de clusters), m y α, los centroides y pertenencias iniciales y el umbral ε usado como condición de parada.

En el primer paso son calculados los factores de pertenencias a partir de la siguiente ecuación:

$$u_{ik} = \frac{((1 - K(x_k, v_i)) + \frac{\alpha}{N_R} \sum_{r \varepsilon N_k} (1 - u_{ir})^m)^{\frac{-1}{(m-1)}}}{\sum_{j=1}^{c} ((1 - K(x_k, v_j)) + \frac{\alpha}{N_R} \sum_{r \varepsilon N_k} (1 - u_{jr})^m)^{\frac{-1}{(m-1)}}} \quad (4)$$

donde N_k representa la vecindad cuadrada o cúbica, según se trabaje con imágenes 2D o 3D respectivamente, y que se encuentra centrada en el píxel x_k, N_R es la cardinalidad de N_k, y α ($0<\alpha<1$) es el parámetro que controla la importancia del factor de penalización.

Posteriormente, se actualizan los centroides:

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^m K(x_k, v_i) x_k}{\sum_{k=1}^{n} u_{ik}^m K(x_k, v_i)} \quad (5)$$

Estos pasos son repetidos hasta que se satisfaga la condición $max_{i,k} \|u_{ik}^t - u_{ik}^{t-1}\| \leq \varepsilon$.

2.3 Modified Spatial Kernelized Fuzzy C-Means (MSKFCM)

Este algoritmo, es una variante del propuesto en [5] y consiste en una combinación de los algoritmos SFCM [8] y KFCM [7]. La principal diferencia radica en la utilización de una implementación diferente del KFCM, de esta manera, se consigue superar el gran coste computacional que supone la implementación original, permitiendo paralelizar el código y reducir notablemente el tiempo de ejecución, indispensable cuando se trata de ejecuciones de estudios completos. Al igual que en el algoritmo anterior la función kernel utilizada es la Gaussian Radial Basis(GRBF).

El principal objetivo de esta implementación consiste en combinar las principales fortalezas de cada uno de los dos algoritmos utilizados, esto es, homogeneidad e insensibilidad al ruido. Combinando los principales pasos de estos

algoritmos se consigue un método iterativo que consta de las siguientes fases:

1. Selección de los parámetros iniciales como el número de clusters, centroides iniciales y valores de p, q y m.
2. Cálculo de los factores de pertenencia:

$$u_{ik} = \frac{(1-K(x_k,v_i))^{-1/(m-1)}}{\sum_{j=1}^{c}(1-K(x_k,v_j))^{-1/(m-1)}} \quad (6)$$

3. Actualización de los factores de pertenencia:

$$u'_{ik} = \frac{u_{ik}^p h_{ik}^q}{\sum_{j=1}^{c} u_{jk}^p h_{jk}^q} \quad (7)$$

con $h_{ik} = \sum_{z \in NB(x_k)} u_{iz}$, donde NB representa la vecindad cuadrada o cúbica (imágenes 2D o 3D), y que se encuentra centrada en el píxel x_k.

4. Cálculo de los nuevos centroides:

$$v'_i = \frac{\sum_{k=1}^{n} {u'_{ik}}^m K(x_k,v_i) x_k}{\sum_{k=1}^{n} {u'_{ik}}^m K(x_k,v_i)} \quad (8)$$

5. Si el error es menor que un umbral $\|V^t - V^{t-1}\| \leq \varepsilon$, parar. En otro caso, volver al paso 2.

3 MÉTRICAS DE EVALUACIÓN

En este análisis se han seleccionado nueve estudios de TACs pulmonares de alta resolución de la base de datos del LIDC y adquiridos por dispositivos de diferentes fabricantes. Cada uno de ellos pertenece a un sujeto diferente y contiene un rango de imágenes de entre 124-282 que representan un total de 1562. Todas los estudios presentan en uno o más cortes nódulos pulmonares de diferente naturaleza: difusos, juxtapleurales, sólidos, etc; y presentes en diferentes partes del tórax.

Los resultados se analizan en base a las métricas de verdaderos positivos (VP) y falsos positivos (FP). Principalmente se han escogido estas medidas debido a que las máscaras proporcionadas por el LIDC sólo muestran la ubicación y forma de los nódulos presentes obviando otras estructuras del tórax presentes en el TAC.

4 RESULTADOS

Los resultados presentados en esta sección se han obtenido tras aplicar los tres algoritmos a cada uno de los nueve estudios variando sus principales parámetros de entrada. Se busca cuantificar su comportamiento y ajustarlos para el conjunto global, intentando evitar la sobreadaptación a un estudio particular, ya que el objetivo final es obtener un buen ajuste que pueda ser incorporado al sistema CAD. Los resultados presentados utilizando vecindades 2D se obtienen para cada corte individual; por el contrario, para obtener los resultados con vecindades 3D se han utilizado todas las imágenes que conforman un estudio.

Todos los algoritmos probados presentan mejores resultados segmentando los cortes en tres clusters que representen: el fondo de la imagen, el tejido pulmonar y las regiones detectadas como nódulos. En el caso de seleccionar cuatro clusters se producen problemas a la hora de clasificar los posibles nódulos ya que en ocasiones se asignan a diferentes clusters lo que dificulta su uso en un sistema CAD.

Los parámetros que se han modificado para el MSKFCM son: la σ entre [150, 700], p y q variando entre [1, 2] y el tamaño de las vecindades a radio 3, 5 y 7. En el caso del SKFCM se ha variado la σ entre [150, 700], el factor α de 0.1 a 0.7 y el tamaño de las vecindades a radio 3, 5 y 7. Por tanto, para cada estudio y cada algoritmo se han realizado 90 pruebas seleccionando diferentes parámetros de entrada.

En la Figura 1, se muestra el índice de acierto en comparación con los falsos positivos para todos los cortes de los nueve estudios aplicando el algoritmo MSKFCM, con la selección de parámetros con los que se obtienen los mejores resultados. En él, se observa que la mayoría de resultados se sitúa en torno al 100 % de VP y con una tasa de FP menor que 18 %.

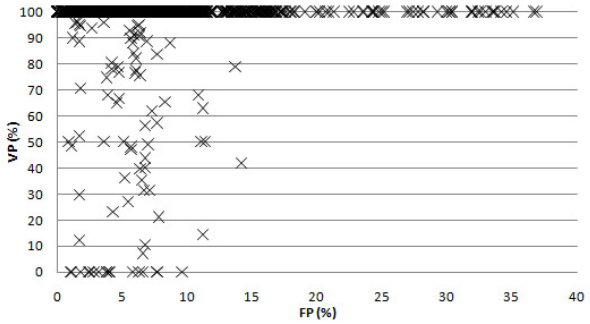


Figura 1: Resultados del MSKFCM para los 9 estudios. Parámetros: $\sigma = 250,\ m = 2,\ p = 2,\ q = 1$.

Es reseñable que los puntos agrupados por debajo del 50 % de acierto se corresponden con nódulos juxtapleurales en los que en la fase de segmentación se pierde cierto borde de incapaz de recuperarse en esta etapa; del mismo modo, aquellos con porcentaje del 0 % son nódulos con un único píxel e identificados por los radiólogos como de naturaleza indefinida, por lo que aún estando representados en los resultados no poseen un elevado interés en nuestro problema. En el caso de los cortes con un índice de FP mayor del 20 % se corresponderían principalmente con aquellos primeros cortes tanto superiores como inferiores en los que el

numero de píxeles es pequeño y las intensidades son muy variables siendo difícil su clasificación.

En la Tabla 1 se presenta una comparativa con los resultados obtenidos con la mejor selección de parámetros, para un corte representativo de cada uno de los estudios y seleccionados de zonas diferentes, variando el tamaño de vecindad tanto en 2D como en 3D. Se observa que cuanto más aumentamos el tamaño de la vecindad más se reducen los VP así como los FP, llegando incluso a no ser capaz de clasificar ningún píxel del nódulo presente, como el caso del corte del estudio cinco con vecindad 2D y tamaño 7x7.

Tabla 1: Resultados del MSKFCM utilizando diferentes vecindades. Parámetros: $\sigma = 250,\ m = 2,\ p = 2,\ q = 1$.

		MSKFCM					
		3x3		5x5		7x7	
		2D	3D	2D	3D	2D	3D
1	VP	61,8	61,8	43,6	60,0	41,8	50,9
	FP	6,9	7,3	3,8	6,3	3,6	5,9
2	VP	30,9	48,1	27,2	45,7	23,5	42,0
	FP	3,3	5,7	3,1	5,0	2,9	4,4
2	VP	100,0	95,7	65,2	30,4	47,8	21,7
	FP	4,8	3,6	3,8	1,9	1,8	1,7
4	VP	100,0	95,0	95,0	95,0	90,0	95,0
	FP	9,4	6,3	6,0	5,5	5,0	5,3
5	VP	78,9	78,9	78,9	78,9	0,0	76,3
	FP	13,6	13,7	9,5	13,9	3,0	14,5
6	VP	96,9	89,7	93,8	88,7	85,6	67,0
	FP	17,4	5,7	15,8	4,3	15,7	3,7
7	VP	43,8	43,8	42,2	42,2	42,2	32,8
	FP	7,0	6,8	6,2	5,5	5,3	4,7
8	VP	98,8	97,6	97,6	84,5	95,2	84,5
	FP	4,6	3,8	3,6	2,6	3,1	2,4
9	VP	97,8	95,1	92,2	92,1	91,6	91,6
	FP	5,5	1,7	1,1	1,1	0,9	0,9

Los mejores resultados, en cuestión de estabilidad y porcentaje, se obtienen con ventanas de 3x3. Por otro lado, si bien los resultados para 2D y 3D son similares, se observa en los cortes representativos del estudio 2 y 6 una variación del 17 % y 11 % en VP y FP, respectivamente, a favor de la vecindad 3D. En general, esta situación se repite para otros cortes de los estudios por lo que se obtienen mejores resultados utilizando vecindades 3D.

Cabe señalar, que los cortes seleccionados del estudio 2 y 7 obtienen bajos porcentajes de VP porque parte del nódulo juxtapleural ha sido eliminado en la fase de preprocesado.

En la Figura 2 se presentan los resultados obtenidos por el algoritmo SKFCM de la misma forma que en el caso del MSKFCM. En este caso, se observa un aumento en el ratio de VP situándose en torno al 100 % para la mayor parte del conjunto de datos. Sin embargo, este resultado se ve penalizado por un destacado aumento de la tasa de FP en comparación con el anterior algoritmo, situándose más del 85 % de los puntos en torno al rango 5 %-60 %.

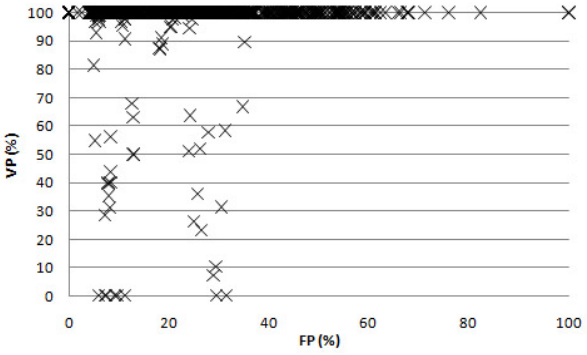


Figura 2: Resultados del SKFCM para los 9 estudios. Parámetros: $\sigma = 550,\ \alpha = 0.5,\ m = 2$

Cabe señalar que estos resultados son obtenidos al seleccionar un factor de penalización $\alpha = 0.5$, con lo que estaríamos considerando un índice medio respecto a la influencia de la vecindad. Si disminuimos esta variable a un valor pequeño se obtienen mejores resultados en torno a los FP disminuyendo en algunos casos más de un 20 %. Estos resultados serían muy similares a los conseguidos por los algoritmos kernelized sin usar la espacialidad.

La Tabla 2 muestra en detalle los problemas que se presentan en algunos cortes. Por ejemplo, en el corte seleccionado del cuarto estudio se observa una tasa de FP que varía desde un 23 % hasta un 41 % cuando en el caso del MSKFCM es de 5 % hasta un 9 %, y únicamente con una reducción del 5 % para los VP. Esta situación se repite en otros cortes indicados como el quinto y el sexto.

Otra característica que penaliza el comportamiento de este algoritmo es la diferencia que se observa en los resultados obtenidos seleccionando vecindades 2D o 3D. Es el caso de los dos primeros estudios, en los que la variación en la tasa de FP puede llegar a más de un 28 % no siendo en ningún caso nunca menor del 20 %. Por otro lado, al contrario de lo que ocurre en el MSKFCM según se aumenta el tamaño de la vecindad las tasas de VP y FP aumentan.

Basándonos en estos resultados, podemos decir que el SKFCM prioriza la información espacial que contienen los píxeles que pertenecen al cluster que se etiqueta como nódulo, haciendo que cuanto mayor sea la vecindad más crezca el tamaño de éste. Sin embargo, en el MSKFCM el efecto es el contrario, teniendo más peso los píxeles que pertenecen al pulmón, provocando que cuanto mayor sea la ventana mayor es el número de píxeles asignado a este cluster.

En la Tabla 3 se muestra un resumen de los mejores resultados obtenidos para cada uno de los algoritmos analizados, buscando su comparación con el FCM.

Tabla 2: Resultados del SKFCM utilizando diferentes vecindades. Parámetros: $\sigma = 550,\ \alpha = 0.5,\ m = 2$.

		SKFCM					
		3x3		5x5		7x7	
		2D	3D	2D	3D	2D	3D
1	VP	49,1	63,6	49,1	63,6	60,0	63,6
	FP	5,0	24,2	5,2	29,4	5,8	31,3
2	VP	34,6	51,9	34,6	51,9	35,8	53,1
	FP	3,7	26,3	3,7	29,6	3,7	31,1
3	VP	95,7	100	100	100	100	100
	FP	3,5	10,1	8,0	19,2	14,9	23,4
4	VP	100	100	100	100	100,0	100
	FP	30,8	23,8	32,8	26,6	34,3	27,9
5	VP	89,5	89,5	89,5	89,5	89,5	92,1
	FP	20,9	35,2	26,1	38,5	28,7	40,1
6	VP	97,9	97,9	99,0	97,9	99,0	97,9
	FP	31,1	21,3	33,5	27,5	35,3	30,2
7	VP	43,8	43,8	43,8	43,8	43,8	43,8
	FP	9,7	8,4	15,3	15,9	21,0	21,2
8	VP	100	100	100	100	100	100
	FP	6,0	5,2	8,6	10,1	11,7	14,2
9	VP	97,0	97,8	100	100	100	100
	FP	2,4	5,9	13,0	14,7	17,1	19,2

Tabla 3: Resultados comparativos de los diferentes algoritmos.

		FCM	MSKFCM	SKFCM
1	VP	58,2	61,8	49,1
	FP	6,4	7,3	5,0
2	VP	19,8	48,1	34,6
	FP	5,1	5,7	3,7
3	VP	91,3	95,7	95,7
	FP	3,2	3,6	3,5
4	VP	90	95,0	100,0
	FP	6,3	6,3	30,8
5	VP	78,9	78,9	89,5
	FP	12	13,7	20,9
6	VP	84,5	89,7	97,9
	FP	2,8	5,7	31,1
7	VP	39,1	43,8	43,8
	FP	3,5	6,8	9,7
8	VP	70,2	97,6	100,0
	FP	2,3	3,8	6,0
9	VP	95,7	95,1	97,0
	FP	2,2	1,7	2,4

En ella se observa que los mejores resultados se obtienen con el algoritmo MSKFCM, presentando ratios de VP elevados y bajos índices de FP. Los peores resultados son los proporcionados por el SKFCM siendo su comportamiento muy variable entre estudios y presentando un alto índice de FP. El FCM presenta resultados medios siendo penalizado por su rigidez a la hora de controlar su comportamiento al carecer de parámetros variables que modifiquen su respuesta en un entorno de gran variabilidad como es este caso.

En las Figuras 3 y 4 se muestran dos ejemplos de los resultados variables obtenidos por los algoritmos MSKFCM y SKFCM en dos cortes de dos estudios diferentes.

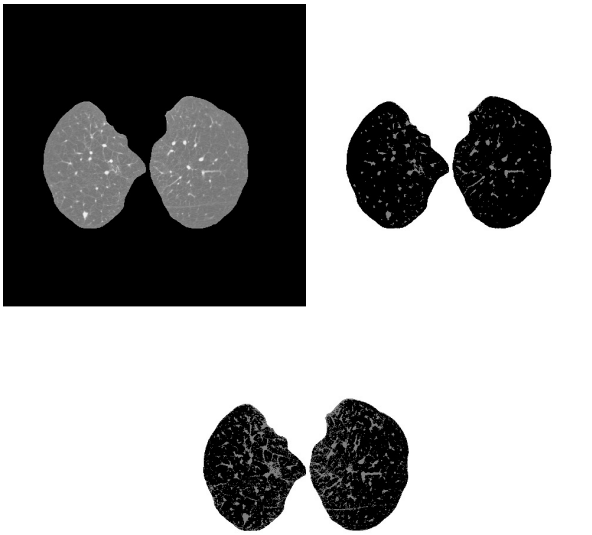


Figura 3: Resultados del MSKFCM y el SKFCM(abajo) para un corte del sexto estudio.

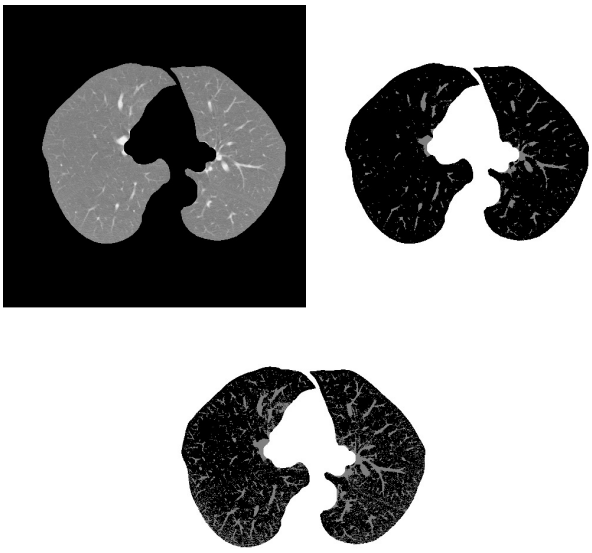


Figura 4: Resultados del MSKFCM y el SKFCM(abajo) para un corte del cuarto estudio.

5 CONCLUSIONES

En este trabajo se presenta una comparativa del comportamiento de dos algoritmos difusos kernelized espaciales a la hora de detectar regiones de interés en nueve estudios de alta resolución preprocesados. Haciendo uso de su factor de espacialidad se analiza el uso de vecindades espaciales tanto en 2D como en 3D.

El algoritmo SKFCM presenta resultados muy variables entre los diferentes estudios proporcionando tasas de falsos positivos que superan el 50 % en muchos casos, con lo cual afectaría gravemente a la futura clasificación de los posibles nódulos pulmonares. En cuanto al uso de vecindades en 3D se observa una gran variabilidad empeorando los resultados sobre todo en torno a los falsos positivos.

En el caso del MSKFCM se obtienen buenos resultados con tasas de verdaderos positivos elevadas y con porcentajes de falsos positivos contenidos en torno al 15 %. Al contrario que SKFCM y FCM, presenta un comportamiento estable en todo el conjunto de prueba al igual que en el uso de vecindades 3D, proporcionando los mejores resultados en este caso.

De acuerdo con este análisis, el uso de vecindades 3D mejora en algunos casos la detección de ROIs si bien no influye de manera decisiva en los resultados. Por otro lado, el aumento del tamaño de la vecindad no mejora en ningún caso los resultados provocando comportamientos diferentes entre los algoritmos analizados.

Por otra parte, pensamos que deberían desarrollarse nuevas funciones de pertenencia espaciales que tuviesen en cuenta vecindades de tres dimensiones. Como se comentó anteriormente, en algunos casos los resultados no mejoran, incluso llegan a empeorar; cuando se podría esperar que al disponer de mayor información y una mejor aproximación a la estructura real, en este caso del pulmón y nódulo, se mejorasen los resultados. En este sentido, estamos comenzando a realizar pruebas utilizando vecindades con diferentes distribuciones espaciales.

Agradecimientos

Este trabajo ha sido financiado por el proyecto 09SIN006PR de la convocatoria I+D Suma del Plan Galego de Investigación, Desenvolvemento e Innovación Tecnolóxica (INCITE), Xunta de Galicia.

Referencias

[1] J. Ahmedin et al, "Annual report to the nation on the status of cancer, 1975-2005, featuring trends in lung cancer, tobacco use, and tobacco control," *Journal of the National Cancer Institute*, vol. 100, no. 23, pp. 1672-1694, December 2008.

[2] S.G. Armato III, G. McLennan, M.F. McNitt-Gray for the Lung Image Database Consortium Research Group, "Lung Image Database Consortium: Developing a Resource for the Medical Imaging Research Community," *Radiology*, 232, pp. 739-748, 2004.

[3] J. C. Bezdek, Ehrlich, R. and Full, W. , "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, 10 (2), 191-203, 1984.

[4] A. Castro, C. Boveda, A. Rey and B. Arcay, "An analysis of various clustering algorithms for ROI detection in high resolution CT lung images," Lect. Notes Comput. Sc., vol. 6374, Warsaw, Poland, 2010.

[5] A. Castro and B. Arcay, "Comparison of various fuzzy clustering algorithms in the detection of ROI in lung CT and a modified kernelized-spatial fuzzy c-means algorithm," Proc. of 10th IEEE Int. Conf. on Inf. Tech. and Appl. in Biom., Corfu, Greece, 2010.

[6] T.M. Cover, "Geomeasureal and statistical properties of systems of linear inequalities in pattern recognition," *IEEE Trans. Electron. Comput*, vol. EC-14, pp. 326-334, Mar. 1965.

[7]] S. C. Chen and D. Q. Zhang. "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(4), 1, 2.2, 2004.

[8] K. Chuang, H. Tzeng,S. Chen, J. Wu, T.Chen, "Fuzzy C-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, 30, pp. 9-15, 2006.

[9] Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki, "Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique," *IEEE Trans. Med. Imaging*, vol. 27, no. 7, pp. 595-604, July 2001.

[10] I. Sluimer, A.M.R. Schilham, M. Prokop, and B.V. Ginneken, "Computer analysis of computed tomography scans of the lung: a survey," *IEEE Trans. Med. Imaging* , 2006, pp.385-405.

[11] D. Q. Zhang,] S. C. Chen, "A novel kernelized fuzzy C-means algorithm with application in medical image segmentation," *Artificial Intelligence in Medicine*, Volume 32, Issue 1, pp. 37-50, 2004.

DETECCIÓN DE BORDES EN IMÁGENES INTERVALARES

C. Lopez-Molina[1,2], **B. De Baets**[2], **E. Barrenechea**[1], **H. Bustince**[1]

[1]*Dpto. Automática y Computación, Universidad Pública de Navarra*, carlos.lopez@unavarra.es
[2]*Dept. of Mathematical Modelling, Statistics and Bioinformatics, Ghent University*, bernard.debaets@ugent.be

Resumen

Una imagen digital es una aproximación de una situación real, y por tanto conlleva cierta incertidumbre. En este trabajo modelamos la ambiguedad realacionada con la intensidad de un pixel usando una representación intervalar de la misma. Aplicamos el método de Sobel para la detección de bordes en la versión intervalar de la imagen, dando lugar una representación de los bordes en forma de conjunto intervalo valorado difuso. Para terminar, comprobamos experimentalmente cómo el uso de imágenes intervalares puede mejorar el rendimiento del método de Sobel.

Palabras Clave: Procesamiento de imagen, Detección de bordes, método de Sobel, Conjunto intervalo-valorado difuso,

1 Introducción

Cualquier proceso de discretización se basa en la medida de un hecho de naturaleza continua, y por tanto genera cierta pérdida de información. Además de este error, debemos considerar que los datos pueden estar contaminados de muchas maneras. En el caso de las imágenes digitales, hay que tener en cuenta diferentes factores, tanto inherentes al modelo (número limitado de tonos) or ajenos a él (ruido, malas condiciones de iluminación). De esta manera, no podemos tener nunca certeza absoluta acerca del tono de un píxel. En la detección de bords, estos problemas se manifiestan de manera clara, ya que a veces ni siquiera dos humanos pueden ponerse de acuerdo sobre dónde está el límite de un objeto.

La lógica difusa ha sido usada con éxito en muchos campos. Se han publicado diferentes aplicaciones, especialmente en tareas en las que la ambiguedad o la información parcial (o contradictoria) se da asiduamente. Dada la naturaleza de las imágenes, la lógica difusa aparece como una opción natural para manejar la incertidumbre [1, 2]. De hecho, dentro del procesamiento de imagen, la detección de bordes es una de las tareas en las que ha sido más prolífica. En la literatura encontramos la lógica difusa para cada subproceso de la detección de bordes, desde la misma interpretación de la imagen [3, 4] a la reconstrucción de bordes una vez se han caracterizado [5]. Se han empleado diferentes técnicas basadas en lógica difusa, como por ejemplo los sistemas de inferencia [6] o la morfología difusa [7]. Sin embargo, un hecho en el que coinciden la gran mayoría de los autores es en usar conjuntos difusos clásicos, siendo el uso de las extensiones minoritario [8].

En este trabajo presentamos y justificamos la representación invervalar de una imagen, diseñada para mejorar la captura de la ambigüedad. Además, presentamos la extensión directa del método de Sobel para la detección de bordes para su aplicación en imagenes intervaloradas. Este método genera una representación de los bordes de la imagen en forma de conjunto intervalo-valorado difuso (CIVD). Para terminar, mostramos un método de conversión de dicho CIVD en una imagen binaria de bordes.

El resto del trabajo se articula como sigue: La Sección 2 se dedica al análisis y modelado de la ambigüedad en imágenes digitales. La Sección 3 incluye la extensión del método de Sobel para la detección de bordes en imágenes intervalares. Mostramos en la Sección 4 los resultados del experimento de comparación del uso de imágenes clásicas e intervalares en el método de Sobel. Para terminar, incluimos unas conclusiones en la Sección 5.

2 Imágenes y representación de la ambigüedad

2.1 Representación de una imagen

Consideramos que una imagen es una matriz de M filas y N columnas, siendo $P = \{1,\dots,M\} \times \{1,\dots,N\}$ el conjunto de sus posiciones En una imagen I, el valor de un pixel en

la posición $p \in P$, I_p es un valor en $\{0,\ldots,255\}$. Además, denotamos como $n(p) \subset P$ al conjunto de posiciones en una ventana 3×3 centrada en (p).

2.2 Imágenes y ambiguedad inherente

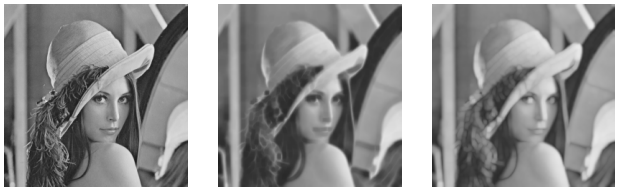
Las imágenes digitales son el resultado de la discretización de la realidad. Es decir, un muestreo discreto de un hecho continuo con infinita precisión. Por lo tanto, contiene diferentes fuentes de incertidumbre y ambiguedad que deben ser consideradas cuando se procesa una imagen. La mayoría de esas fuentes de incertidumbre son contextuales, en el sentido de que pueden estar presentes (o no) en una imagen, dependiendo de las condiciones en las que fue tomada. Algunos ejemplos de esta contaminación son el ruido, la iluminación excesiva o las sombras. Sin embargo, hay un tipo de incertidumbre inherente al mismo modelo de imagen digital: el error de medida. Las imágenes digitales discretizan la realidad en dos sentidos diferentes, el espacial y el tonal, y cada uno de ellos produce un error de medida diferente:

- *Error espacial:* Los objetos de la realidad tiene infinita precisión. Sin embargo, se vuelven discretos en la representación basada en píxeles. Por lo tanto, debido a la manera en que la información se almacena, podemos estar desplazando un objeto hasta 1 posición en cualquier dirección.
- *Error tonal:* Las imagenes almacenan un número limitado de tonos. Por ejemplo, hay 2^8 tonos en una imagen clásica en escala de grises, y 2^{24} en una RGB. Sin embargo, incluso si usamos una precisión mayor, siempre tendremos un cierto límite. Por tanto el error tona se manifiesta como $\pm$ 1 tono.

De hecho, ambos errores se derivan de la incapacidad de un modelo discreto para medir cada una de las dimensiones del problema (espacial y tonal). Una representación escalar de la intensidad de un pixel no es suficiente para capturar dichos errores. Por lo tanto, usaremos una representación basada en intervalos. La construcción de dichos intervalos está directamente derivada de los dos puntos anteriores. A partir de una imagen I, generamos una imagen intervalar W tal que el valor en cada pixel $p \in P$ es

$$W_p = \left[\text{máx}(0, \min_{p' \in n(p)} I_{p'} - 1),\ \text{mín}(255, \max_{p' \in n(p)} I_{p'} + 1)\right]. \quad (1)$$

Es decir, asignamos a cada posición un intervalo que abarca las intensidades en su vecindario (asumiendo el error espacial), modificados con $\pm$ 1 tone (por el error tonal). La Figura 1 incluye un ejemplo de la imagen de Lena junto a los límites inferior y superior de su representación intervalar. Dado que la construcción es similar a la dada en [9], es de esperar que la diferencia entre el límite superior e inferior sea una imagen difusa de bordes.



(a) Imagen original (b) Representación intervalar

Figura 1: Imagen de Lena junto a los límites inferir y superior de su representación intervalar.

Op. horizontal

-1		1
-2		2
-1		1

Op. vertical

1	2	1
-1	-2	-1

Figura 2: Operadores de Sobel para la detección de bordes.

Hay que tener en cuenta que la construcción de imágenes intervalares es bastante robusta frente a ciertos tipos de ruido (como el Gaussiano), pero tambien relativamente sensible a otros (como el S&P). Sin embargo, debemos entender que tratar con diferentes fuentes de contaminación no es el principal objetivo de esta metodología, ya que sólo tratamos de capturar factores intrínsecamente ligados al modelo de representación de imagen.

Las imágenes intervalares proveen una interpretación realista de la imagen, ya que el error de medida es inevitable. Sin embargo, hay pocas técnicas en la literatura del procesamiento de imagen preparadas para tratar con este tipo de imágenes. Dedicamos el resto de este trabajo a generaliar el resultado del método de Sobel para la detección de bordes en imágenes intervalares.

3 El método de Sobel y su aplicación en imágenes intervalares

3.1 El método de Sobel

El método de Sobel para detección de bordes se basa en el filtrado de la imagen con dos operadores, cada uno de los cuales estima el cambio de intensidad en la dirección de un eje (horizontal y vertical). Estas estimaciones se usan como componentes ortogonales del *gradiente*, una representación vectorial del cambio de intensidad en cada punto de la imagen.

Los operadores de Sobel (en la Figura 2) son filtros discretos de convolución para la aproximación de las derivadas parciales en cada dirección. Por ejemplo, en el caso de la

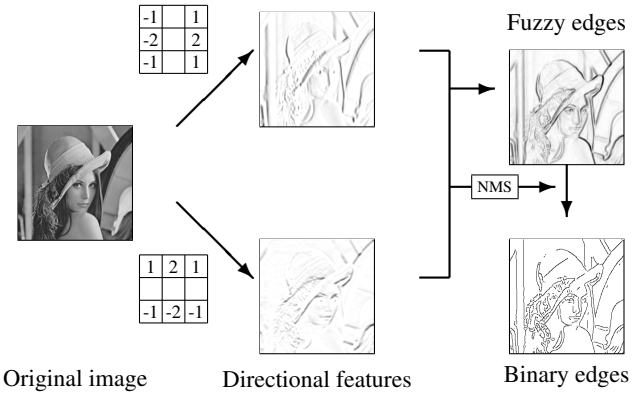


Figura 3: Representación esquemática del método de Sobel para la detección de bordes.

convolución horizontal $H_p \in [-1,1]$, tenemos

$$H_p = \frac{1}{4 \cdot 255} \sum \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * I(n(p)), \qquad (2)$$

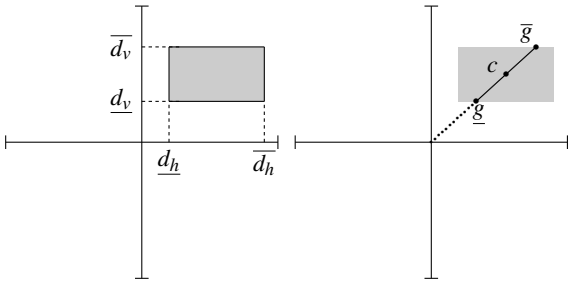
donde $*$ es el operador convolución. La estimación vertical (V_p) se calcula análogamente.

El gradiente en cada posición se calcula a partir de las dos estimaciones ortogonales de cambios, tal que $\mathbf{G}_p = (H_p, V_p)$. Una vez cada posición de la imagen se asocia con un gradiente, disponemos de muchas técnicas para obtener una representación binaria de los bordes. La Figure 3 incluye un ejemplo de la aplicación del método de Sobel en la imagen de Lena. La magnitud de los gradientes se ha usado para calcular la representación intermedia, en forma de conjunto difuso clásico. Para ello, hemos elegido la norma Euclidea normalizada $||\mathbf{G}_p|| = \frac{1}{\sqrt{2}}\sqrt{H_p^2 + V_p^2}$ [10]. Para terminar la conversión a una representación binaria hemos usado la supresión de no-máximos (NMS [11]) y el método de Rosin para thresholding [12].

El método de Sobel, así como el de Prewitt [13] ha tenido un gran impacto en la literatura. De hecho, se consideran como los métodos pioneros en la aproximación a la detección de bordes basada en derivadas [14], que ha propocionado alguno de los métodos más reconocidos en la actualidad, como el de Canny [11] o el de Marr-Hildreth [15] .

3.2 Aplicación a imágenes intervalares

Una vez tenemos las imágenes intervalares, existen diferentes maneras de aplicar los operadores de Sobel. Por ejemplo, podemos agregar los extremos de cada invervalo para generar una imagen escalar, o aplicarlo individualmente en cada uno de los extremos del intervalo. Sin embargo, esto choca con la idea de usar intervalos para mantener la mayor



(a) Area-based interpretation (b) Interval-valued gradient

Figura 4: Area gradient

cantidad posible de información. Por ello, vamos a aplicar la convolución directamente en imágenes intervalares, usando la aritmética clásica de intervalos [16]:

- Suma de intervalos: $[\underline{a},\overline{a}] + [\underline{b},\overline{b}] = [\underline{a}+\underline{b},\overline{a}+\overline{b}]$;
- Diferencia de intervalos: $[\underline{a},\overline{a}] - [\underline{b},\overline{b}] = [\underline{a}-\overline{b},\overline{a}-\underline{b}]$;
- Producto de un escalar positivo s y un intervalo: $s \cdot [\underline{a},\overline{a}] = [s \cdot \underline{a}, s \cdot \overline{a}]$

De esta manera, los componentes del gradiente en ambas direcciones vienen expresados como intervalos (en concreto, subintervalos de $[-1,1]$). Como consecuencia, el gradiente asociado con cada posición no es un punto en el espacio $[-1,1]^2$, sino un área en ese mismo espacio, como se ilustra en la Figura 4. Este área es el resultado de la propagación de la incertidumbre acerca del valor exacto de cada pixel en la image. Es por otro lado lógico que la incertidumbre acerca de los valores iniciales de un problema se traslade (y aumente) al tomar medidas sobre ellos. Tomaremos la interpretación de que el gradiente real existe, y que la variación de la intensidad en la imagen puede ser representada por medio de un vector. Entendemos que dicho gradiente viene representado por algun punto dentro del área sombreada de la Figura 4(a).

Nuestra interpretación geométrica de dicha área se basa en el cálculo del centro de gravedad de la misma (c), calculado tal que

$$c = \left[\frac{\underline{d_h} + \overline{d_h}}{2}, \frac{\underline{d_v} + \overline{d_v}}{2} \right] \qquad (3)$$

A partir del centro de gravedad, tomaremos los límites del segmento formado en la intersección del área con la linea que pasa por c y el origen de coordenadas. Estos dos puntos se llamarán $\underline{g}$ (el punto más cercano al origen) y $\overline{g}$ (el más alejado), tal y como se ilustra en la Figura 4(b). Esto definirá un CIVD tal que para cada punto se asocia con su intervalo $[||\underline{g}||, ||\overline{g}||]$.

Para poder aplicar las técnicas clásicas de binarización de imágenes, tenemos que obtener una representación difusa

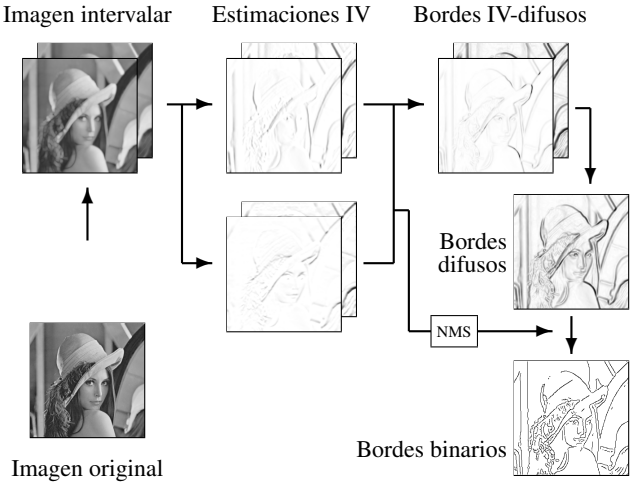


Figura 5: Esquema visual del método de Sobel intervalo-valorado.

clásica de los bordes. En nuestro caso usaremos el operador de Atanassov para convertir las pertenencias intervalares en pertenencias clásicas, definidos tal que $K_\alpha([\underline{a},\overline{a}]) = (\underline{a} + \alpha(\overline{a} - \underline{a}))$. De esta manera, $\alpha \in [0,1]$ mide el optimismo con respecto a la cantidad de bordes en la imagen. Una vez los bordes han sido modelados en forma de conjunto difuso clásico, el procesamiento es tal y como se ve en la Figura 3. Primero usaremos NMS para afinar los bordes, y luego usaremos el método de Rosin para binarizar. De esta manera, obtendremos bordes que satisfacen las restricciones de Canny [11].

Este método de detección, tal y como se ilustra en la Figura 5, se llama método de Sobel Intervalo-valorado (IV-Sobel).

A modo de muestra del rendimiento del método, hemos tomado la imagen de Lena de la Figura 1(a). Primero, hemos regularizado la imagen con un filtro Gaussiano de $\sigma = 1.80$. En el momento de crear la representación difusa, hemos experimentado con diferentes valores de α. Los resultados se incluyen en la Figura 6. Nótese que los límites inferior y superior del CIVD se obtienen con $\alpha = 0.0$ y $\alpha = 1.0$, respectivamente En esta figura observamos cómo un incremento del optimismo (representado por un incremento de α) aumenta la cantidad de puntos seleccionados como bordes. Sin embargo, no queda claro si esto es bueno o malo, en la medida en que se incluyen bordes reales pero también respuestas erróneas. Por ejemplo, en la Figura 6 observamos cómo la silueta del sombrero casi se completa con valores altos de α, pero estos también implican muchas detecciones falsas, como las de la región del pelo.

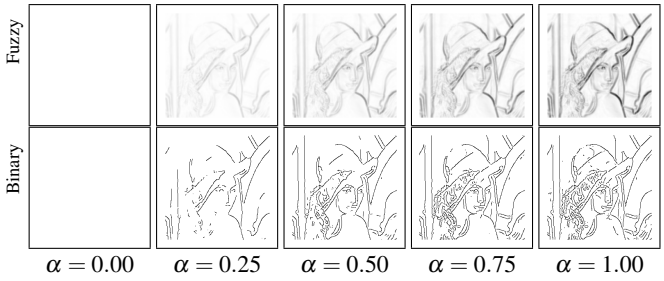


Figura 6: Bordes difusos y binarios obtenidos usando diferentes valores de α en la conversión de conjunto invervalo-valorado difuso a conjunto difuso clásico.

4 Resultados experimentales

4.1 Objetivo del experimento

En este experimento queremos comparar los resultados obtenidos por el método de Sobel con los que obtiene el IV-Sobel. Para ello, hemos ejevutado ambos algoritmos sobre un conjunto idéntico de imágenes, para posteriormente cuantificar la calidad de sus resultados.

En este experimento hemos usado el test set del Berkeley Segmentation Dataset (BSDS) [17], que contiene 100 imágenes. Estas imágenes tienen una resolución de 481×321 (o 321×481) píxeles, y se proveen en escala de grises de 256 tonos. Además, cada una de ellas tiene de 5 a 10 segmentaciones manuales. Como estas segmentaciones se dan en forma de regiones completas, las hemos usado para generar la solución al problema de la detección de bordes.

4.2 Método de comparación

La métrica Delta de Baddeley (BDM) es una medida inicialmente diseñada para la comparación de conjuntos binarios [18], pero también puede usarse para comparar imágenes. Esta medida intenta medir la similaridad de dos conjuntos de elementos en un mismo referencial. Como los bordes se proveen generalmente en forma de imágen binaria (satisfaciendo las restricciones de Canny [11]), esta medida puede usarse para nuestro propósito.

Sean B_1 y B_2 dos imágenes de bordes binarias y M, N y P sus dimensiones tal y como se han definido en la Sección 2. Dado un valor $1 < k < \infty$ la k-BDM entre B_1 y B_2 (brevemente $\Delta^k(B_1,B_2)$) es:

$$\Delta^k(B_1,B_2) = \left[\frac{1}{|P|} \sum_{p \in P} |w(d(p,B_1)) - w(d(p,B_2))|^k \right]^{\frac{1}{k}}, \quad (4)$$

donde $d(p,B_i)$ representa la distacia entre la posición p al borde más cercano de la imagen B_i y $w : [0,\infty] \to [0,\infty]$ es una función cóncava, creciente usada para modular. Nótese que asumimos que cada imagen tiene, al menos, un píx-

el seleccionado como borde. En nuestros experimentos usamos la distancia Euclídea como d. Por lo tanto, $d(p,B_i)$ se refiere a la mínima distancia Euclídea entre p y un punto borde de B_i. En cuanto al resto de parámetros, usamos $w(x) = x$ [19]. y $k = 2$.

La *puntuación* de un método de detección en una imagen es la distancia mínima entre su resultado y el conjunto de soluciones de la imagen. Es decir, la distancia a la imagen solución más *cercana*. A su vez, la puntuación en todo el conjunto de imágenes es la media de la puntuación en cada imagen.

4.3 Algoritmo

El Algoritmo 1 incluye todos los pasos dados en el procesamiento de los bordes. El algoritmo propuesta podría ser modificado para obtener mejores resultados. Por ejemplo, la umbralización con un sólo valor suele ser superada por técnicas más complejas (como la *hysteresis*), sobre todo en zonas de bajo contraste. De igual manera, la regularización Gaussiana podría sustituirse por otras técnicas que ofrecen mejor preservación de bordes [20]. Sin embargo, el objetivo de este experiento no es obtener los mejores resultados posibles, sino realizar una comparación entre el método de Sobel y el IV-Sobel.

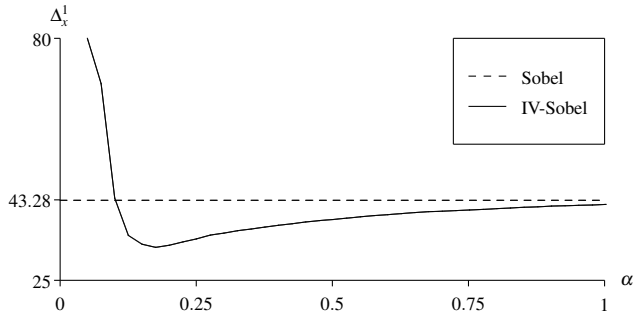
Data: Una imagen I, un indicador de optimismo α
Result: Una imagen binaria de bordes B
begin
  Regularizar I con un filtro Gaussiano de desviación típica $\sigma_1 = 1.0$;
  Crear la versión intervalorada de la imagen, I_w;
  Calcular las estimaciones intervalares de los cambios de intensidad H_w and V_w;
  Usar la técnica de la Fig. 4, para generar el CIVD E_w;
  Create the fuzzy edges as $E = K_\alpha(E_w)$;
  Obtener una versión de bordes finos (T) usando NMS [11];
  Generar la imagen binaria de bordes B a partir de T usando el método de Rosin [12];
end

Algorithm 1: Algoritmo del método IV-Sobel con indicador de optimismo α seleccionado por el usuario.

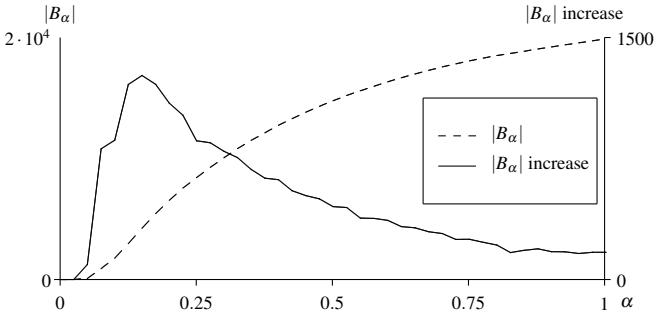
4.4 Resultados

Hemos ejecutado ambos algoritmos en las 100 imágenes del BSDS [17], usando diferentes $\alpha \in \{0, 0.025, \ldots, 1\}$ fijas. En la Figura 7 se muestra la media en el rendimiento em el conjunto de imágenes y su comparación con el rendimiento obtenido por el método de Sobel. Además, mostramos en un gráfico el número de píxeles que se han clasificado como bordes en cada caso, así como su incremento en cada variación de α.

Observamos que el resultado del método de Sobel puede ser claramente mejorado por el del método IV-Sobel. Hay



(a) Average BDM



(b) Average number of edge pixels

Figura 7: Resultados obtenidos por los métodos Sobel y IV-Sobel en el BSDS.

un rango inicial de valores de α que producen resultados muy malos, pero el rendimiento rápidamente mejora. Entonces, cuando α tiende a 1, el rendimiento de ambos métodos converge. Además, es interesante como los mejores resultados (en términos de BDM) coinciden con los valores de α que generan un mayor incremento de $|B_\alpha|$.

Por lo tanto, mantenemos la hipótesis de que una buena manera de determinar α es probar experimentalmente un conjunto de candidatos, y seleccionar aquel que produzca el mayor incremento de $|B_\alpha|$. El procedimiento del método IV-Sobel no-supervisado (uIV-Sobel) se encuentra en el Algoritmo 2. Obviamente, en este caso no existe la posibilidad de elegir $\alpha = 0$ ya que su resultado no puede ser comparado con ningún valor inferior de α.

5 Conclusiones

Hemos analizado el papel del error de medida en las imágenes digitales, proponiendo una representación intervalar de la imagen para reducir su impacto. Además, hemos propuesto una generalización del método de Sobel para la detección de bordes en imágenes intervalares. Para terminar, hemos realizado un experimento para mostrar la mejora del rendimiento del método de Sobel cuando se aplica en imágenes intervalares, lo cual nos lleva a pensar que la caracterización del error de medida puede resultar en una mejora de los métodos de procesamiento de imagen.

Data: Una imagen I, la variación del indicador de optimismo $\delta\alpha$
Result: Una imagen binaria de bordes B
begin
  Regularizar I con un filtro Gaussiano de desviación típica $\sigma_1 = 1.0$;
  Crear la versión intervalorada de la imagen, I_w;
  Calcular las estimaciones intervalares de los cambios de intensidad H_w and V_w;
  Usar la técnica de la Fig. 4, para generar el CIVD E_w;
  Generar el conjunto de αs candidatos,
  $A \leftarrow \{0, \delta\alpha, 2 \cdot \delta\alpha, \ldots, 1\}$;
  foreach $\alpha \in A$ **do**
    Crear el CDC representando los bordes $E_\alpha = K_\alpha(E_w)$;
    Obtener una versión de bordes finos (T_α) usando NMS [11];
    Generar la imagen binaria de bordes B_α a partir de T_α usando el método de Rosin [12];
  end
  $B \leftarrow$ El B_α produciendo el mayor incremento de $|B_\alpha|$;
end

Algorithm 2: Algoritmo del método IV-Sobel sin supervisión.

Referencias

[1] Aina Barcelo, Eduard Montseny, and Pilar Sobrevilla. Fuzzy texture unit and fuzzy texture spectrum for texture characterization. *Fuzzy Sets and Systems*, 158(3):239–252, 2007.

[2] J. Chamorro-Martínez, J.M. Medina, C.D. Barranco, E. Galán-Perales, and J.M. Soto-Hidalgo. Retrieving images in fuzzy object-relational databases using dominant color descriptors. *Fuzzy Sets and Systems*, 158(3):312–324, 2007.

[3] F. Jacquey, F. Comby, and O. Strauss. Fuzzy edge detection for omnidirectional images. *Fuzzy Sets and Systems*, 159(15):1991–2010, 2008.

[4] S. K. Pal and R. A. King. On edge detection of x-ray images using fuzzy sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5(1):69–77, 1983.

[5] T. Law, H. Itoh, and H. Seki. Image filtering, edge detection, and edge tracing using fuzzy reasoning. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(5):481–491, 1996.

[6] F. Russo. Edge detection in noisy images using fuzzy reasoning. In *Instrumentation and Measurement Technology Conference*, volume 1, pages 369–372, 1998.

[7] J.-A. Jiang, C.-L. Chuang, Y.-L. Lu, and C.-S. Fahn. Mathematical-morphology-based edge detectors for detection of thin edges in low-contrast regions. *IET Image Processing*, 1(3):269–277, 2007.

[8] A. Jurio, M. Pagola, R. Mesiar, G. Beliakov, and H. Bustince. Image magnification using interval information. *IEEE Trans. on Image Processing*, (99):1, 2011. In press.

[9] E. Barrenechea, H. Bustince Sola, B. De Baets, and C. Lopez-Molina. Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images. *IEEE Transactions on Fuzzy Systems*, -:1, 2011. Accepted paper.

[10] C. Lopez-Molina, J. Fernandez, A. Jurio, M. Galar, M. Pagola, and B. De Baets. On the use of quasi-arithmetic means for the generation of edge detection blending functions. In *Proceedings of the IEEE International Conference on Fuzzy Systems*, 2010.

[11] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[12] P. L. Rosin. Unimodal thresholding. *Pattern Recognition*, 34(11):2083–2096, 2001.

[13] J. M. S. Prewitt. *Object enhancement and extraction*, pages 75–149. Picture Processing and Psychopictorics. Academic Press, 1970.

[14] G. Papari and N. Petkov. Edge and line oriented contour detection: State of the art. *Image and Vision Computing*, 29(2-3):79–103, 2011.

[15] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London*, 207(1167):187–217, 1980.

[16] R. Moore. *Interval Analysis*. Prentince-Hall, 1996.

[17] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference on Computer Vision*, volume 2, pages 416–423, 2001.

[18] A. J. Baddeley. Errors in binary images and an L^p version of the Hausdorff metric. *Nieuw Archief voor Wiskunde*, 10:157–183, 1992.

[19] C. Lopez-Molina, H. Bustince, J. Fernandez, P. Couto, and B. De Baets. A gravitational approach to edge detection based on triangular norms. *Pattern Recognition*, 43(11):3730–3741, 2010.

[20] J. Weickert, B.M. ter Haar Romeny, and M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. on Image Processing*, 7(3):398–410, 1998.

Determinación de la mejor pareja t-norma-implicación en el gradiente morfológico

Manuel González-Hidalgo, **Sebastià Massanet**, **Arnau Mir**

Departament de Ciències Matemàtiques i Informàtica,
Universitat de les Illes Balears
Campus UIB, Crta. Valldemossa Km. 7.5,
07122 Palma, Mallorca, {manuel.gonzalez, s.massanet, arnau.mir}@uib.es

Resumen

En esta comunicación, se estudia el rendimiento del detector de contornos de la morfología matemática borrosa basada en t-normas. Se consideran distintas combinaciones de t-norma e implicación borrosa en el gradiente morfológico y se comparan los resultados mediante la medida de comparación objetiva de Pratt, *Pratt's figure of merit*. Se prueba que la combinación más utilizada en la literatura, la t-norma de Łukasiewicz y su implicación residual es, de hecho, la peor combinación tanto respecto a los resultados visuales como a la medida objetiva considerada.

Palabras Clave: Morfología matemática borrosa, detección de contornos, t-normas, implicaciones, histéresis.

1 INTRODUCCIÓN

La detección de contornos es una operación fundamental de bajo nivel en el procesamiento de imágenes que resulta esencial para desarrollar operaciones de alto nivel relacionadas, por ejemplo, con la visión por ordenador y el reconocimiento. Su rendimiento es crucial para los resultados finales de las técnicas de procesamiento de imágenes. En las últimas décadas, se han desarrollado un gran número de algoritmos de detección de contornos. Los diferentes enfoques van desde los algoritmos clásicos ([18]) basados en el uso de un conjunto de máscaras de convolución, a las nuevas técnicas basadas en conjuntos borrosos y sus extensiones ([4]).

Entre los enfoques borrosos, destaca la morfología matemática borrosa que generaliza la morfología binaria ([19]) usando conceptos y técnicas de la teoría de conjuntos borrosos ([2], [15]). Esta teoría permite un mejor tratamiento y una representación con mayor flexibilidad de la incertidumbre y la ambigüedad presentes en cada nivel de una imagen. Los operadores morfológicos son las herramientas básicas de esta teoría. Un operador morfológico P transforma una imagen A que se quiere analizar en una nueva imagen $P(A,B)$ mediante un elemento estructurante B. Las cuatro operaciones morfológicas básicas son dilatación, erosión, cierre y apertura. Gracias a que las imágenes en niveles de gris pueden ser representadas como conjuntos borrosos, se pueden usar herramientas borrosas para definir los operadores morfológicos borrosos. Así, se han usado conjunciones (normalmente t-normas continuas, pero también uninormas, ver [9]) y sus implicaciones residuales. Recientemente, se ha introducido una morfología matemática borrosa basada en t-normas discretas con buenos resultados en detección de contornos ([8]).

La morfología matemática borrosa basada en t-normas en $[0,1]$ fue introducida por De Baets en [6] y [7] estableciendo un marco general donde se usaban conjunciones e implicaciones para definir los operadores morfológicos. Después de analizar qué propiedades debían de satisfacer la t-norma y la implicación para obtener una morfología matemática borrosa que tuviese todas las propiedades algebraicas deseables, se concluyó que la pareja formada por la t-norma de Łukasiewicz T_{LK} y su implicación residual I_{LK} es la representante de la única familia de t-normas (las nilpotentes) que satisfacen todas las propiedades. En consecuencia, siempre se suele utilizar (T_{LK}, I_{LK}) para implementar un detector de contornos basado en esta morfología.

Sin embargo, la propiedad clave para definir un detector de contornos en esta teoría es la extensividad de la dilatación y la antiextensividad de la erosión. Esta propiedad se satisface con unas mínimas propiedades del elemento estructurante, de la t-norma y de la implicación. Así, muchas más t-normas e implicaciones pueden ser usadas para definir un gradiente morfológico útil para detectar contornos. En consecuencia, el objetivo de esta contribución es estudiar cuál es la mejor combinación de t-norma e implicación si el objetivo se limita a la detección de contornos. Para realizar una comparación de los resultados, se utilizará la medida objetiva de Pratt, *FoM* (ver el capítulo 15 de

[18]). Esta medida requiere que la imagen de contornos sea binaria y con los contornos de un píxel de anchura, en coherencia con las restricciones de Canny expuestas en [5]. Por lo tanto, una vez obtenida la imagen de contornos borrosos se implementarán el algoritmo de *thinning* conocido como *Non-Maxima Suppression* (NMS), expuesto por el mismo Canny y el recientemente introducido algoritmo automático de histéresis basado en la determinación de una "zona de inestabilidad" en el histograma propuesto en [14].

La comunicación se organiza de la siguiente forma. En la Sección 2, se recuerdan las definiciones de los operadores morfológicos y de los operadores borrosos que los definen. En la Sección 3, se presentan las t-normas e implicaciones consideradas, así como el algoritmo desarrollado con cada configuración. En la siguiente sección, se presentan y analizan los resultados. Finalmente, se comparten las conclusiones y el trabajo futuro que se quiere desarrollar.

2 PRELIMINARES

Los operadores morfológicos borrosos se definen a partir de operadores borrosos tales como las t-normas y las implicaciones. Más detalles sobre estos conectivos lógicos pueden encontrarse en [12] y [1], respectivamente.

Definición 1. Una *t-norma* es una función $T : [0,1]^2 \to [0,1]$ conmutativa, asociativa, monótona creciente y con elemento neutro 1.

Definición 2. Un operador binario $I : [0,1]^2 \to [0,1]$ es una *implicación borrosa* si es decreciente en la primera variable, creciente en la segunda y satisface $I(0,0) = I(1,1) = 1$ y $I(1,0) = 0$.

Así, se pueden definir los operadores morfológicos borrosos básicos como son la dilatación y la erosión. A partir de ahora, se usará la siguiente notación: I denotará una implicación, T una t-norma, A una imagen en niveles de gris, y B un elemento estructurante en niveles de gris.

Definición 3. La *dilatación borrosa* $D_T(A,B)$ y *la erosión borrosa* $E_I(A,B)$ de A por B son las imágenes en niveles de gris definidas como

$$\begin{aligned} D_T(A,B)(y) &= \sup_x T(B(x-y),A(x)), \\ E_I(A,B)(y) &= \inf_x I(B(x-y),A(x)). \end{aligned}$$

Como se ha comentado, con unas mínimas propiedades la siguiente proposición asegura la extensividad de la dilatación borrosa y la antiextensividad de la erosión borrosa.

Proposición 1. *Sean T una t-norma, I una implicación que satisface (NP), esto es, $I(1,y) = y$ para todo $y \in [0,1]$ y B un elemento estructurante en niveles de gris tal que $B(0) = 1$. Entonces se verifican las siguientes inclusiones:*

$$E_I(A,B) \subseteq A \subseteq D_T(A,B).$$

Así, como en el caso de la morfología clásica, la diferencia entre la dilatación borrosa y la erosión borrosa de una imagen en niveles de gris, $D_T(A,B) \setminus E_I(A,B)$, conocida como operador *gradiente borroso*, puede usarse en detección de contornos.

3 CONFIGURACIONES Y ALGORITMO A DESARROLLAR

Según la Proposición 1, cualquier t-norma y cualquier implicación que satisfaga (NP) son adecuadas para definir el gradiente borroso. Notar que la t-norma de Łukasiewicz y su implicación residual, la configuración estándar para definir el gradiente, es una configuración adecuada ya que cualquier implicación residual, o R-implicación, generada a partir de una t-norma satisface (NP) (ver Teorema 2.5.4 en [1]). Sin embargo, la propiedad (NP) es muy frecuente en los distintos tipos de implicaciones ya que la satisfacen además, las (S,N)-implicaciones, las implicaciones de Yager y las QL-operaciones (ver Proposiciones 2.4.3, Teorema 3.1.7 y 2.6.2 en [1] respectivamente). Estas últimas no son siempre implicaciones en el sentido de la Definición 2 ya que puede no verificarse el decrecimiento en la primera variable. A pesar de este hecho, siguen cumpliendo la antiextensividad de la erosión borrosa.

Se han considerado las t-normas de la Tabla 1 y las implicaciones y operaciones de la Tabla 2 generando 40 configuraciones diferentes en la pareja (T,I) que define el gradiente morfológico. Entre las implicaciones destacar que, además de las implicaciones más conocidas, se han incluido representantes de las familias más importantes de implicaciones. La implicación 8 es una implicación f-generada con $f(x) = \cos\left(\frac{\pi}{2}x\right)$, la operación 9 es la QL-operación, que no satisface el decrecimiento en la primera variable, generada por T_M, $S_M(x,y) = \max\{x,y\}$ y $N_C(x) = 1-x$ y la implicación 10 es la QL-implicación obtenida a partir de T_P, $S = S^2_{SS}$ la t-conorma de Schweizer-Sklar con $\lambda = 2$ y $N_K(x) = 1-x^2$.

Tabla 1: T-normas consideradas para calcular la dilatación.

Nombre	**Fórmula**
Łukasiewicz	$T_{LK}(x,y) = \max\{x+y-1,0\}$
Mínimo	$T_M(x,y) = \min\{x,y\}$
Producto	$T_P(x,y) = xy$
Mínimo Nilp.	$T_{nM}(x,y) = \begin{cases} 0 & \text{si } x+y \leq 1, \\ \min\{x,y\} & \text{en otro caso.} \end{cases}$

3.1 NMS e HISTÉRESIS AUTOMÁTICA

Para comparar los resultados obtenidos, se necesita alguna medida de rendimiento objetiva. Estas medidas necesitan, además de una imagen binaria de bordes con anchura un pí-

Tabla 2: Implicaciones y operaciones consideradas para el cálculo de la erosión.

Nombre	**Fórmula**	(S,N)	R	QL	**Yager's**
Łukasiewicz	$I_{LK}(x,y) = \text{mín}\{1, 1-x+y\}$	✓	✓	✓	✓
Yager	$I_{YG}(x,y) = \begin{cases} 1 & \text{si } x=0 \text{ y } y=0 \\ y^x & \text{si } x>0 \text{ o } y>0 \end{cases}$	X	X	X	✓
Reichenbach	$I_{RC}(x,y) = 1-x+xy$	✓	X	✓	✓
Kleene-Dienes	$I_{KD}(x,y) = \text{máx}\{1-x,y\}$	✓	X	✓	X
Gödel	$I_{GD}(x,y) = \begin{cases} 1 & \text{si } x \le y \\ y & \text{si } x > y \end{cases}$	X	✓	X	X
Goguen	$I_{GG}(x,y) = \begin{cases} 1 & \text{si } x \le y \\ \frac{y}{x} & \text{si } x > y \end{cases}$	X	✓	X	✓
Fodor	$I_{FD}(x,y) = \begin{cases} 1 & \text{si } x \le y \\ \text{máx}\{1-x,y\} & \text{si } x > y \end{cases}$	✓	✓	✓	X
Implicación 8	$I_{f_c}(x,y) = \frac{2}{\pi}\arccos\left(x\cos\left(\frac{\pi}{2}y\right)\right)$	✓	X	X	✓
Operación 9	$I_9(x,y) = \begin{cases} 1-x+x^2 & \text{si } x \le y, \\ 1-x+xy & \text{en otro caso.} \end{cases}$	X	X	✓	X
Implicación 10	$I_{PC}(x,y) = 1-(\text{máx}\{x(1+xy^2-2y),0\})^{\frac{1}{2}}$	X	X	✓	X

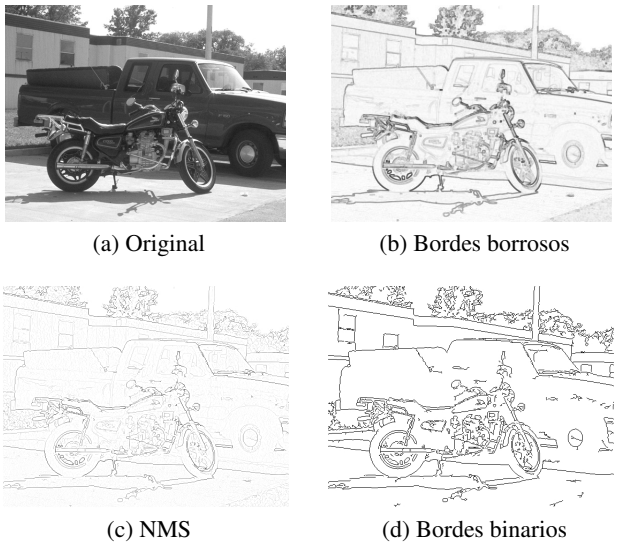
xel (DE) correspondiente a los bordes detectados, una imagen de bordes de referencia o *ground truth* (GT) que es una imagen binaria de bordes con anchura un píxel que contiene los contornos reales de la imagen original. Existen varias medidas de rendimiento para detección de contornos en la literatura (ver [16]). En este trabajo se usará para cuantificar la similaridad entre (DE) y (GT) la medida propuesta por Pratt, *Pratt's figure of merit*, definida como

$$FoM = \frac{1}{\text{máx}\{card\{DE\}, card\{GT\}\}} \cdot \sum_{x \in DE} \frac{1}{1+ad^2},$$

donde *card* es el número de píxeles de contorno de la imagen, a es una constante de escala y d es la distancia de separación entre un píxel de contorno obtenido respecto a uno ideal. En nuestro caso, se considerará $a = 1$ y la distancia Euclídea d. Un valor superior de FoM indica una mejor capacidad para detectar contornos.

Sin embargo, los métodos borrosos de detección de contornos generan una imagen donde el valor del píxel representa su pertenencia al conjunto de bordes. Esta idea contradice las restricciones de Canny ([5]), que fuerzan una representación de los contornos como imágenes binarias de un píxel de anchura. Por lo tanto, la imagen de bordes borrosos debe de ser adelgazada y binarizada. La imagen de bordes borrosos contendrá valores grandes donde haya un gradiente grande, pero para identificar los bordes, se han de adelgazar las vastas regiones con valores grandes del gradiente para que sólo permanezcan las magnitudes en aquellos puntos que son máximos locales. Non-Maxima Suppression (NMS), propuesto por Canny, realiza este trabajo eliminando todos los valores a lo largo de la recta del gradiente que no son valores extremos. Se ha usado la implementación de P. Kovesi en Matlab ([13]).

Finalmente, para binarizar la imagen se ha implementado una histéresis automática basada en la determinación de la zona de inestabilidad del histograma para encontrar los umbrales (ver [14]). La histéresis permite escoger aquellos píxeles significativos de ser considerados como bordes usando sus valores de pertenencia. Se usan dos umbrales U_1, U_2 con $U_1 \le U_2$. Todos los píxeles con grado de pertenencia mayor que U_2 son considerados bordes, mientras que aquellos menores que U_1 se descartan. Los píxeles con grado de pertenencia entre los dos umbrales se seleccionan si están 8-conectados con píxeles superiores a U_2. El método se ha implementado con el mismo conjunto inicial de candidatos usado en [14], $\{0.01, \ldots, 0.25\}$. En la Figura 1, se puede observar la secuencia del algoritmo.



(a) Original (b) Bordes borrosos

(c) NMS (d) Bordes binarios

Figura 1: Secuencia del algoritmo utilizado.

4 RESULTADOS Y ANÁLISIS

Como ya se ha comentado, se necesita una base de imágenes con sus respectivas imágenes de bordes de referencia para comparar las salidas obtenidas por las diferentes configuraciones. Así, se han usado las imágenes y sus imágenes de bordes de referencia de la base de imágenes pública de la University of South Florida [1] ([3]). En nuestro estudio, se han usado 15 de las 50 imágenes de la base de datos.

Los resultados, obtenidos todos ellos mediante el siguiente elemento estructurante isotrópico utilizado en [15]

$$B = \begin{pmatrix} 0.86 & 0.86 & 0.86 \\ 0.86 & 1 & 0.86 \\ 0.86 & 0.86 & 0.86 \end{pmatrix},$$

están recopilados en la Tabla 3. Como puede observarse en esta tabla, el hecho más significativo es el pobre rendimiento de la configuración (T_{LK}, I_{LK}), siendo la peor configuración en media y en particular, para 10 de las 15 imágenes. En la Figura 2, se puede apreciar que los resultados obtenidos por esta configuración contienen pocos bordes respecto al resto. Notar que la sola presencia de T_{LK} o de I_{LK} en una configuración determinada, empeora los resultados tanto visuales como la medida objetiva analizada. En cambio, las configuraciones a partir de las t-normas T_M y T_{nM} son las que dan mejores resultados sobre todo aquellas con las implicaciones I_{KD}, I_{GD} y I_{FD}. Los resultados obtenidos con las configuraciones usando T_M y T_{nM} son bastante parecidos debido a que ambas t-normas tienen la misma expresión en una zona de $[0,1]^2$ (ver la Tabla 1), así como a la naturaleza del elemento estructurante elegido.

Para reforzar el análisis anterior, se ha realizado un análisis de conglomerados para estudiar las relaciones entre las configuraciones. En primer lugar, se ha determinado el número de conglomerados óptimo según el llamado test F de reducción de la variabilidad ([17], página 223), dando lugar a 3 conglomerados. Finalmente, aplicando el algoritmo K-medias con el referido número de conglomerados se han obtenido los siguientes resultados:

- Conglomerado 1: T_M y T_{nM} con $I_{YG}, I_{KD}, I_{GD}, I_{GG}, I_{FD}$ y I_{PC}; T_P con I_{KD}, I_{GD} y I_{FD}.
- Conglomerado 2: T_M y T_{nM} con I_{LK}, I_{RC}, I_{f_c} y I_9; T_{LK} con I_{KD}, I_{GD} y I_{FD}.
- Conglomerado 3: T_{LK} y T_P con I_{LK}, I_{YG}, I_{RC}, I_{GG}, I_{f_c}, I_9 y I_{PC}.

Observando estos resultados podemos establecer un cierto orden de rendimiento entre los operadores lógicos borrosos considerados:

$$T_M, T_{nM} \succ T_P \succ T_{LK},$$

$$I_{KD}, I_{GD}, I_{FD} \succ I_{YG}, I_{GG}, I_{PC} \succ I_{LK}, I_{RC}, I_{f_c}, I_9,$$

donde $A, B \succ C$ indica que las configuraciones obtenidas a partir de A o B dan un mejor rendimiento que las obtenidas a partir de C.

Una posible explicación al rendimiento de la configuración (T_{LK}, I_{LK}) proviene de la propia definición de los operadores morfológicos. Se puede comprobar fácilmente que $T_{LK} \leq T_{nM}, T_M$ y $I_{LK} \geq I_{KD}, I_{GD}, I_{FD}$. Por lo tanto, se satisface

$$\begin{aligned} D_{T_{LK}}(A,B)(y) &= \sup_x T_{LK}(B(x-y), A(x)) \\ &\leq \sup_x T_{nM}(B(x-y), A(x)) = D_{T_{nM}}(A,B)(y), \\ E_{I_{LK}}(A,B)(y) &= \inf_x I_{LK}(B(x-y), A(x)) \\ &\geq \inf_x I_{KD}(B(x-y), A(x)) = E_{I_{KD}}(A,B)(y). \end{aligned}$$

Consecuentemente, el detector de contornos basado en la configuración (T_{nM}, I_{KD}) detecta muchos más posibles contornos que el basado en (T_{LK}, I_{LK}). Este hecho podría acarrear la presencia de falsos bordes y/o texturas en la imagen binaria de bordes resultante, sin embargo, los resultados visuales confirman que el algoritmo de histéresis automática discrimina correctamente los píxeles deseados.

Por otra parte, como se ha indicado anteriormente, recientemente se ha introducido una morfología matemática borrosa basada en t-normas discretas. Esta morfología trabaja con imágenes definidas por funciones de $\mathbb{Z}^2 \to L = \{0, 1, \ldots, 255\}$. En [8] y [10], se determinó que además de la t-norma discreta de Łukasiewicz, el mínimo nilpotente discreto así como toda una familia de t-normas discretas consistentes en una suma ordinal (con un único sumando) de la t-norma de Łukasiewicz en el recuadro $[a, n-a]^2$, $a \in L$ con $a \leq n-a$, truncado por 0, satisfacen todas las propiedades algebraicas requeridas en una morfología matemática para ser una buena teoría. A partir de este nuevo enfoque, se implementó en [8] y [11] un detector de contornos basado en la configuración $(T_{nM}, I_{T_{nM}})$ con buenos resultados. Sin embargo, $I_{T_{nM}}$ es de hecho I_{FD} y en la Tabla 3, la configuración (T_{nM}, I_{FD}) ocupa una de las primeras posiciones. Así, si existiera alguna reticencia por escoger configuraciones que no satisfacen todas las propiedades algebraicas de la morfología matemática borrosa basada en t-normas en $[0,1]$, se puede escoger la configuración (T_{nM}, I_{FD}) en la morfología matemática borrosa basada en t-normas discretas que sí cumple todas las propiedades algebraicas, incluida la idempotencia generalizada.

5 CONCLUSIONES Y TRABAJO FUTURO

En este trabajo, se ha llevado a cabo una comparación entre los gradientes morfológicos generados a partir de distintas configuraciones de t-norma e implicación, (T, I), en $[0,1]$, demostrando que la configuración más usada (T_{LK}, I_{LK}) es de hecho la peor configuración de todas las consideradas.

[1] Se puede descargar desde `ftp://figment.csee.usf.edu/pub/ROC/edge_comparison_dataset.tar.gz`

Tabla 3: Medidas estadísticas asociadas a los valores obtenidos de *FoM*.

Configuración		Media (± Desv.)	Imágenes		Configuración		Media (± Desv.)	Imágenes	
T-norma	Imp.		Mejor	Peor	T-norma	Imp.		Mejor	Peor
T_{LK}	I_{LK}	0.2738 (± 0.1248)	1	10	T_P	I_{LK}	0.3343 (± 0.1237)	0	0
	I_{YG}	0.3850 (± 0.0937)	1	0		I_{YG}	0.3969 (± 0.1044)	0	0
	I_{RC}	0.3329 (± 0.0975)	0	0		I_{RC}	0.3645 (± 0.1113)	0	0
	I_{KD}	0.4030 (± 0.0934)	0	0		I_{KD}	0.4351 (± 0.0851)	0	0
	I_{GD}	0.4058 (± 0.0870)	1	0		I_{GD}	0.4368 (± 0.0898)	1	0
	I_{GG}	0.3676 (± 0.1094)	0	0		I_{GG}	0.3775 (± 0.1122)	1	0
	I_{FD}	0.4007 (± 0.0922)	0	0		I_{FD}	0.4326 (± 0.0861)	1	0
	I_{f_c}	0.3002 (± 0.1147)	0	2		I_{f_c}	0.3485 (± 0.1158)	1	0
	I_9	0.3244 (± 0.0987)	0	0		I_9	0.3595 (± 0.1131)	0	0
	I_{PC}	0.3509 (± 0.0983)	0	0		I_{PC}	0.3669 (± 0.1079)	1	0
T_M	I_{LK}	0.3950 (± 0.0952)	0	2	T_{nM}	I_{LK}	0.3951 (± 0.0947)	0	0
	I_{YG}	0.4386 (± 0.0866)	1	0		I_{YG}	0.4391 (± 0.0863)	1	0
	I_{RC}	0.4155 (± 0.0826)	1	0		I_{RC}	0.4164 (± 0.0823)	0	0
	I_{KD}	0.4479 (± 0.0975)	**2**	0		I_{KD}	**0.4480** (± 0.0973)	1	0
	I_{GD}	0.4476 (± 0.0975)	**2**	0		I_{GD}	0.4471 (± 0.0974)	1	0
	I_{GG}	0.4361 (± 0.0905)	0	0		I_{GG}	0.4359 (± 0.0904)	0	0
	I_{FD}	0.4470 (± 0.0979)	**2**	0		I_{FD}	0.4466 (± 0.0976)	0	0
	I_{f_c}	0.3923 (± 0.0888)	0	1		I_{f_c}	0.3928 (± 0.0890)	0	0
	I_9	0.4177 (± 0.0783)	1	0		I_9	0.4182 (± **0.0775**)	1	0
	I_{PC}	0.4285 (± 0.0865)	0	0		I_{PC}	0.4286 (± 0.0865)	0	0

Además, se han argumentado las posibles causas de este hecho y se ha proporcionado una alternativa (la morfología basada en t-normas discretas) para usar una de las configuraciones más óptimas (T_{nM}, I_{FD}) sin perder ninguna propiedad algebraica de la morfología.

El próximo paso consiste en generalizar los operadores morfológicos usando una t-conorma y una t-norma en lugar de las operaciones sup e ínf respectivamente en la dilatación y la erosión. Como el máximo es la más pequeña de las t-conormas y el mínimo la más grande de las t-normas, esta generalización podría mejorar los resultados ya que se ampliaría el gradiente morfológico correspondiente permitiendo una mayor detección de contornos.

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto nacional MTM2009-10320 con el apoyo de fondos FEDER.

Referencias

[1] M. Baczyński and B. Jayaram. *Fuzzy Implications*, volume 231 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin Heidelberg, 2008.

[2] I. Bloch and H. Maître. Fuzzy mathematical morphologies: a comparative study. *Pattern Recognition*, 28:1341–1387, 1995.

[3] K. Bowyer, C. Kranenburg, and S. Dougherty. Edge detector evaluation using empirical ROC curves. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '99)*, volume 1, pages 354–359, 1999.

[4] H. Bustince, E. Barrenechea, M. Pagola, and J. Fernandez. Interval-valued fuzzy sets constructed from matrices: Application to edge detection. *Fuzzy Sets and Systems*, 160(13):1819–1840, 2009.

[5] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986.

[6] B. De Baets. Fuzzy morphology: A logical approach. In B. M. Ayyub and M. M. Gupta, editors, *Uncertainty Analysis in Engineering and Science: Fuzzy Logic, Statistics, and Neural Network Approach*, pages 53–68. Kluwer Academic Publishers, Norwell, 1997.

[7] B. De Baets. Generalized idempotence in fuzzy mathematical morphology. In E. E. Kerre and M. Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 2, pages 58–75. Physica-Verlag, New York, 2000.

[8] M. González-Hidalgo, S. Massanet, and J.Torrens. Discrete t-norms in a fuzzy mathematical morphology: Algebraic properties and experimental results.

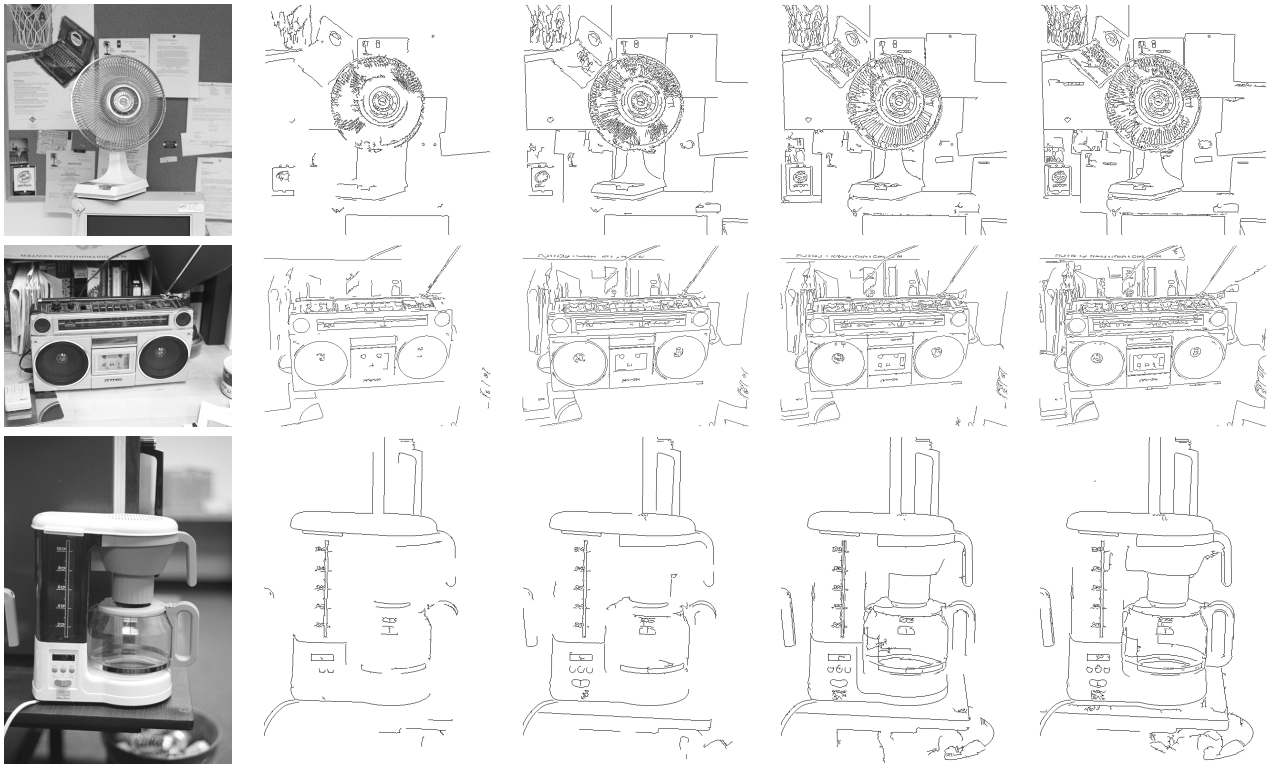


Figura 2: Algunas imágenes originales y de sus contornos obtenidos con diferentes configuraciones (T,I). De izquierda a derecha: (T_{LK},I_{LK}), (T_P,I_{YG}), (T_M,I_9) y (T_{nM},I_{KD}), respectivamente.

In *Proceedings of WCCI-FUZZ-IEEE*, pages 1194–1201, Barcelona, Spain, 2010.

[9] M. González-Hidalgo, A. Mir-Torres, D. Ruiz-Aguilera, and J. Torrens. Image analysis applications of morphological operators based on uninorms. In *Proceedings of the IFSA-EUSFLAT 2009 Conference*, pages 630–635, Lisbon, Portugal, 2009.

[10] M. González-Hidalgo and S. Massanet. Closing and opening based on discrete t-norms. Applications to natural image analysis. In S. Galichet, J. Montero, and G. Mauris, editors, *Proceedings of EUSFLAT-LFA-2011*, volume 17 of *Advances in Intelligent Systems Research*, pages 358–365. Atlantis Press, 2011.

[11] M. González-Hidalgo and S. Massanet. Towards an objective edge detection algorithm based on discrete t-norms. In S. Galichet, J. Montero, and G. Mauris, editors, *Proceedings of EUSFLAT-LFA-2011*, volume 17 of *Advances in Intelligent Systems Research*, pages 319–326. Atlantis Press, 2011.

[12] E. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, London, 2000.

[13] P. D. Kovesi. MATLAB and Octave functions for computer vision and image processing. Centre for Exploration Targeting, School of Earth and Environment, The University of Western Australia. Available from: <http://www.csse.uwa.edu.au/∼pk/research/matlabfns/>.

[14] R. Medina-Carnicer, R. Muñoz-Salinas, E. Yeguas-Bolivar, and L. Diaz-Mas. A novel method to look for the hysteresis thresholds for the Canny edge detector. *Pattern Recognition*, 44(6):1201 – 1211, 2011.

[15] M. Nachtegael and E. Kerre. Classical and fuzzy approaches towards mathematical morphology. In E. E. Kerre and M. Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 1, pages 3–57. Physica-Verlag, New York, 2000.

[16] G. Papari and N. Petkov. Edge and line oriented contour detection: State of the art. *Image and Vision Computing*, 29(2-3):79 – 103, 2011.

[17] D. Peña. *Análisis de datos multivariante*. McGraw-Hill, 2002.

[18] W. K. Pratt. *Digital Image Processing*. Wiley-Interscience, 4 edition, 2007.

[19] J. Serra. *Image analysis and mathematical morphology, vols. 1, 2*. Academic Press, London, 1982,1988.

APRENDIZAJE INTUITIVO DE LÓGICA FUZZY COMBINANDO TEORÍA Y PRÁCTICA

David P. Pancho, **Jose M. Alonso**, **Luis Magdalena**

European Centre for Soft Computing, Mieres, 33600, Spain
{david.perez,jose.alonso,luis.magdalena}@softcomputing.es

Resumen

Este trabajo describe el uso de la herramienta software GUAJE ("Generating Understandable and Accurate Fuzzy Models in a Java Environment") en el contexto de la asignatura para enseñanza de lógica fuzzy que se imparte en el máster internacional sobre "Soft Computing y Análisis Inteligente de Datos", adaptado al Espacio Europeo de Educación Superior impulsado por el proceso de Bolonia, y organizado conjuntamente por el European Centre for Soft Computing y la Universidad de Oviedo. Se propone la implementación de un tutorial interactivo que ayude a los alumnos a utilizar el software, con el fin de que esto les sirva de apoyo para asimilar de forma práctica los conceptos teóricos explicados a lo largo del curso. Para fomentar el carácter didáctico del tutorial, éste se centra en el diseño de sistemas fuzzy simplificando las tareas administrativas propias del software.

Palabras Clave: Enseñanza Práctica, Extracción y Representación de Conocimiento, Diseño de Sistemas Fuzzy.

1 INTRODUCCIÓN

La firma de la declaración de Bolonia[1], en 1999, supuso el comienzo de un proceso de convergencia encaminado a crear un nuevo Espacio Europeo de Educación Superior (EEES) [7, 8], que todavía hoy sigue en construcción [6, 9]. Los pilares básicos que sustentan el EEES son [5]:

- Adopción de un sistema único, estructurado en tres ciclos (grado, máster y doctorado).
- Establecimiento de un sistema común de créditos ("European Credit Transfer System", ECTS). Cada crédito ECTS corresponde a 30 horas que engloban tanto las horas de clase teóricas, como las clases prácticas, los trabajos a realizar por los alumnos, los seminarios asociados y los exámenes finales.
- Revisión y homogenización de las titulaciones universitarias europeas con el fin de simplificar la homologación y el reconocimiento de títulos oficiales expedidos por cualquier país adscrito al EEES.
- Promoción de la cooperación y movilidad de estudiantes, profesores, investigadores, etc.

Con la adaptación al EEES se pone énfasis en la evaluación continua del trabajo diario. Es decir, la calificación final debe reflejar el trabajo hecho por el alumno a lo largo de todo el curso y no depender sólo del resultado de un examen final. Por tanto, los nuevos cursos, tanto de grado como de posgrado, deben conjugar de forma efectiva enseñanzas teóricas y prácticas.

El European Centre for Soft Computing (ECSC) y la Universidad de Oviedo (UNIOVI) organizan conjuntamente (desde el curso 2009/2010) un máster de posgrado internacional adaptado al EEES, impartido integramente en inglés, que forma a los estudiantes para que estos puedan realizar un doctorado, o incorporarse a la industria. El máster tiene un acuerdo bilateral para intercambio de estudiantes con el instituto Tecnológico de Monterrey (Méjico), y está patrocinado por numerosas empresas (CajAstur, Indra, Repsol, HC Energía, Fundación Ramón Areces y SAS).

El objetivo de esta contribución es describir cómo se conjuga teoría y práctica (cumpliendo con el espíritu de Bolonia) en una de las asignaturas impartidas en el citado máster, en concreto en el curso sobre lógica fuzzy. Así mismo, se ha propuesto la mejora del sistema utilizado en la parte práctica, aumentando la capacidad didáctica de la herramienta utilizada para facilitar el estudio.

[1] Declaración conjunta firmada por los ministros de educación de la mayoría de países europeos (Bolonia, 19 de Junio de 1999) <http://www.bolognabergen2005.no/Docs/00-Main_doc/990719BOLOGNA_DECLARATION.PDF>

El resto del trabajo se ha estructurado de la siguiente forma. La sección 2 introduce el contexto en que se ha desarrollado este trabajo. Empieza presentando los objetivos básicos y la estructura en cursos del máster. Después se centra en los detalles del curso sobre lógica fuzzy. A continuación, presenta la herramienta software GUAJE para diseño de sistemas fuzzy, que se utiliza en la parte práctica de la asignatura. Y acaba destacando el uso dado y los puntos débiles de dicha herramienta desde un punto de vista didáctico. La sección 3 describe la propuesta de este trabajo para que GUAJE se pueda usar eficazmente como una herramienta didáctica. Por último, la sección 4 cierra el artículo con unas breves conclusiones.

2 PRELIMINARES

2.1 MÁSTER EN SOFT COMPUTING Y ANÁLISIS INTELIGENTE DE DATOS

El objetivo esencial de este máster internacional, organizado conjuntamente por ECSC y UNIOVI, es formar profesionales capaces de desempeñar tareas que requieren alta cualificación, tanto en el sector público como privado. Al finalizar el máster los alumnos deben estar capacitados para resolver problemas complejos (predicción, optimización, clasificación, regresión, minería de datos, etc.) en aplicaciones reales de todo tipo (logística, banca, automoción, control, diagnóstico médico, etc.) mediante el uso de técnicas de Soft Computing y análsis inteligente de datos imprecisos e inciertos.

En los tres años de existencia del máster, el número medio de estudiantes matriculados ha sido 13.33, siendo el 32.48 % de ellos extranjeros. Para obtener el título de máster los alumnos deben superar al menos 60 créditos ECTS, correspondientes a cursos obligatorios (33 ECTS), cursos optativos (15 ECTS) y proyecto fin de máster (12 ECTS). Los cursos del máster se imparten a lo largo de un año académico completo (entre Septiembre y Junio) y están estructurados en tres grandes bloques:

- **Fundamentos**: Lógica fuzzy, redes neuronales, computación evolutiva, razonamiento probabilístico, minería de datos, etc.
- **Temas avanzados**: Estadística fuzzy, técnicas de clasificación, laboratorios de análisis de datos, etc.
- **Aplicaciones**: Visión por ordenador, robótica, computación con palabras, recuperación de información, etc.

2.2 CURSO SOBRE LÓGICA FUZZY

Esta contribución se centra en uno de los cursos obligatorios, el de enseñanza de lógica fuzzy, del máster presentado en la sección previa. Con este curso se pretende que los alumnos adquieran los conocimientos necesarios para:

- Entender las ventajas e inconvenientes de la lógica fuzzy frente al resto de técnicas de Soft Computing.
- Comprender los conceptos teóricos esenciales que sustentan este tipo de lógica.
- Analizar y diseñar sistemas basados en reglas fuzzy.
- Resolver problemas reales mediante el uso de herramientas basadas en lógica fuzzy.

Este curso ofrece una visión muy completa, tratando desde los conceptos teóricos más básicos hasta los temas de investigación más novedosos, pasando por la resolución de problemas en aplicaciones reales, a nivel teórico y práctico. El temario abarca los siguientes contenidos: Conjuntos fuzzy; razonamiento aproximado; diseño y análisis de sistemas basados en reglas fuzzy; modelado de expresiones lingüísticas; análisis de interpretabilidad; herramientas software para sistemas fuzzy; y aplicaciones.

El curso equivale a 3 créditos ECTS, es decir 90 horas de esfuerzo académico distribuidas de la siguiente forma: 32 horas de clases presenciales, y 58 horas de trabajo personal de los alumnos fuera del aula. De estas horas, la parte correspondiente al estudio de sistemas fuzzy consta de 8 horas en el aula (4 horas teóricas y 4 horas prácticas) y 14 horas de trabajo individual aproximadamente.

2.3 GUAJE

GUAJE[2] es el acrónimo de "Generating Understandable and Accurate Fuzzy Models in a Java Environment" [1, 2]. Es una herramienta software gratuita, distribuida en forma de código abierto bajo licencia GPL[3] ("GNU General Public License") y en el registro de propiedad intelectual de SaveCreative[4]. Esta herramienta está pensada para simplificar las tareas de extracción y representación de conocimiento asociadas al diseño e implementación de sistemas fuzzy interpretables.

En última estancia, se pretende que el diseño sea sistemático e intuitivo, permitiendo llevar a cabo el diseño en distintos niveles de abstracción según el nivel de conocimiento y experiencia del diseñador. GUAJE implementa la metodología de diseño conocida como HILK ("Highly Interpretable Linguistic Knowledge") [3, 4] pensada para construir sistemas fuzzy con un buen equilibrio entre interpretabilidad y precisión, a partir de la integración de conocimiento experto e inducido de forma automática a partir de datos experimentales.

Al terminar el curso sobre lógica fuzzy, los alumnos deben entender (y diseñar) todos los elementos representados en

[2] http://www.softcomputing.es/guaje
[3] http://www.gnu.org/copyleft/gpl.html
[4] https://www.safecreative.org/work/1103168736763

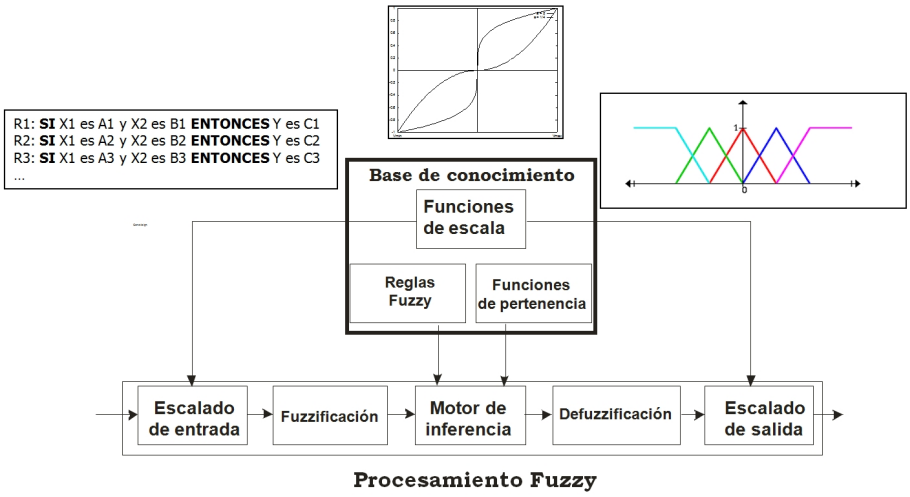


Figura 1: Diagrama de componentes de un sistema fuzzy.

la Figura 1 para construir un sistema fuzzy capaz de resolver un problema concreto. El diseño ha de realizarse con cuidado, considerando tanto los detalles relativos al procesamiento fuzzy como a la construcción de la base de conocimiento.

La herramienta software implementa todos los métodos necesarios para que dicha tarea sea sencilla e intuitiva, tanto para usuarios avanzados como para usuarios no expertos. Para ello, GUAJE ofrece una interfaz gráfica amigable en la que todos los métodos presentan una configuración por defecto que permite alcanzar resultados satisfactorios de manera rápida. Obviamente, el usuario experto puede refinar los resultados haciendo los cambios necesarios para encontrar la configuración óptima para cada problema.

2.4 USO DE GUAJE PARA ENSEÑANZA DE LÓGICA FUZZY

GUAJE no fue pensada en sus orígenes como una herramienta didáctica, sino como una herramienta de diseño. Por ello, aunque su manejo es bastante intuitivo para un usuario experto en lógica fuzzy, tanto su instalación como la primera toma de contacto puede resultar difícil y algo tediosa para usuarios inexpertos, como son la mayoría de alumnos del máster al comienzo del curso sobre lógica fuzzy.

El objetivo de las prácticas es evaluar la capacidad de los alumnos para construir sistemas fuzzy con un buen equilibrio entre precisión e interpretabilidad. Sin embargo, se debe remarcar el carácter didáctico de las prácticas. Así, se puso énfasis en la comprensión de los distintos algoritmos explicando el efecto de usar distintos parámetros (y no sólo la configuración por defecto sugerida por GUAJE), y no en encontrar la mejor solución posible para cada problema. Con este cometido se pidió a los alumnos llevar a cabo las siguientes tareas:

1. Construir particiones fuertes de 3 y 5 términos lingüísticos para cada variable de entrada, comparando los distintos algoritmos para generación automática de particiones proporcionados por GUAJE.

2. Construir bases de reglas con los distintos algoritmos de inducción implementados en GUAJE: Wang and Mendel, Fuzzy Decision Trees, y Fast Prototyping Algorithm.

3. Evaluar precisión (en entrenamiento y test) e interpretabilidad para cada una de las bases de conocimiento construidas.

4. Representar todas las soluciones obtenidas en una gráfica 2D, y seleccionar la base de conocimiento que a juicio del alumno (según los índices de calidad computados previamente) exhibe el mejor equilibrio entre interpretabilidad y precisión.

5. Simplificar la base de conocimiento seleccionada, y comparar la relación interpretabilidad-precisión antes y después de la simplificación.

6. Escribir un informe resumiendo los resultados obtenidos, explicando las decisiones de diseño tomadas, y exponiendo las principales conclusiones derivadas del estudio.

Para más detalles sobre los algoritmos y posibilidades que ofrece el software se puede acudir a la documentación de la herramienta [2].

La experiencia de los dos primeros años del máster, utilizando GUAJE en las prácticas de laboratorio de la asignatura para enseñanza de lógica fuzzy, nos permitió extraer conclusiones interesantes sobre su usabilidad:

- Los alumnos tienen dificultades para relacionar lo visto en la teoría con la parte práctica.
- Muchas opciones presentadas por defecto en GUAJE, como la selección del número de particiones a 3, no se indican a la hora de ejecución, con lo que el alumno pierde el control sobre el sistema diseñado.
- El seguimiento de la clase a través de una demostración del profesor no es lo más adecuado, ya que las múltiples opciones que proporciona GUAJE pueden dar lugar a resultados distintos según el alumno.
- La ayuda html estática proporcionada con la herramienta es poco utilizada por los alumnos, por lo tedioso de la búsqueda de información concreta.

En resumen, GUAJE tiene potencial suficiente para ser usada como una herramienta didáctica, es decir, como una plataforma software con la que los alumnos puedan experimentar y afianzar los contenidos teóricos presentados a lo largo del curso. No obstante, hay que simplificar su proceso de instalación y suavizar la curva de aprendizaje inicial, es decir, el esfuerzo que los alumnos (usuarios no expertos en lógica fuzzy al comenzar el curso, y empezar a usar GUAJE) deben hacer para familiarizarse con el manejo básico de la herramienta.

3 PROPUESTA

Con el fin de utilizar GUAJE, de forma eficaz, como una herramienta didáctica para enseñanza de lógica fuzzy, en este trabajo hemos diseñado un tutorial interactivo, integrado en la propia herramienta, que permite a los alumnos familiarizarse con ella de una forma sencilla, rápida e intuitiva. Dicho tutorial guía al usuario, paso a paso y de forma activa, a través de cada una de las etapas que componen el proceso de diseño de un sistema fuzzy.

El proceso completo (ver Figura 1) debe tener en cuenta tanto la etapa de procesamiento fuzzy como la de construcción de la base de conocimiento. No obstante, lo habitual es que la etapa de procesamiento fuzzy sea la misma para varios problemas y que en cada problema específico el mayor esfuerzo corresponda a la etapa de construcción de la base de conocimiento.

Con esta idea en mente se han diseñado dos módulos que permitan al usuario inexperto, obviando las tareas administrativas de creación y guardado del trabajo intermedio, la creación de un sistema fuzzy en dos supuestos:

- Con la información de un experto (sección 3.1)
- Con el conocimiento inducido a partir de un conjunto de datos (sección 3.2).

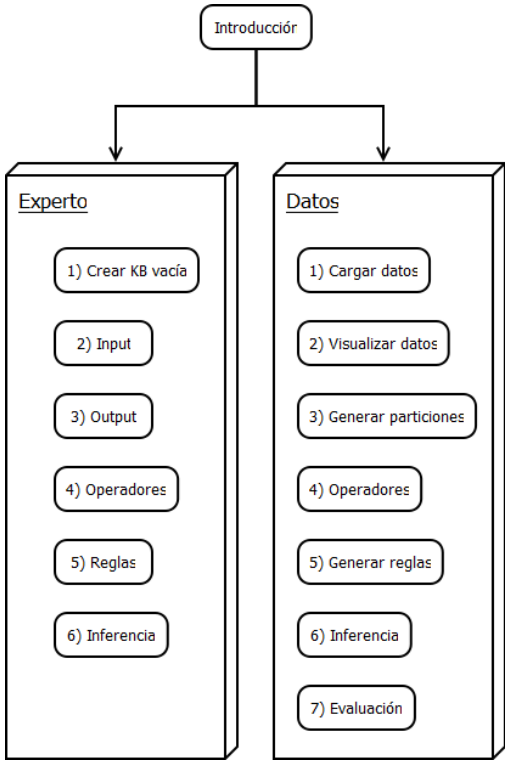


Figura 2: Flujo de iteración de los tutoriales

El tutorial se presenta con una serie de ventanas que nos proporcionan información de la tarea de diseño a realizar y los diálogos a utilizar en GUAJE para llevarla a cabo. Así mismo, y para facilitar la toma de contacto con la herramienta, se proveerá a los usuarios una forma de realizar los pasos de una determinada tarea de forma automática. En la Figura 3 se puede ver la estructura de una de las ventanas del tutorial a modo de ejemplo. Los botones NEXT y BACK permiten navegar por el tutorial, DO realiza los pasos de una tarea automáticamente, mientras que HELP nos muestra la ayuda específica para una tarea.

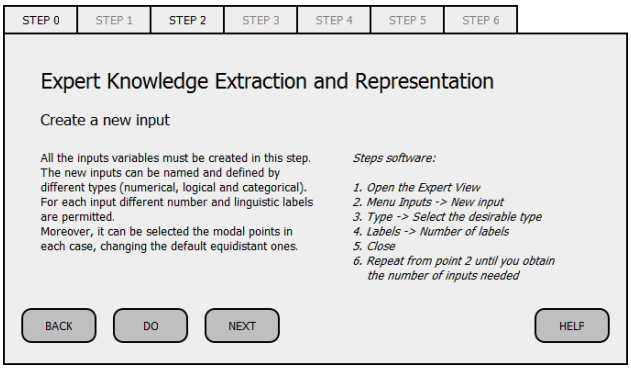


Figura 3: Ventana de definición de variables de entrada

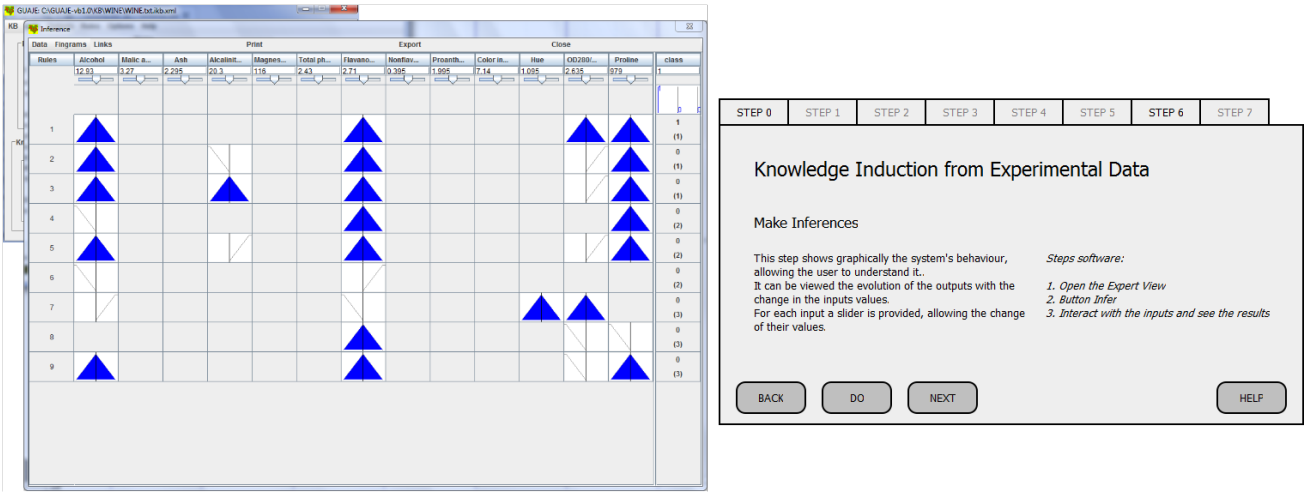


Figura 4: Ventana de inferencia

3.1 CREACIÓN DE UN SISTEMA DESDE CONOCIMIENTO EXPERTO

El primero de los tutoriales tiene como cometido la creación de un sistema experto definiendo todos los elementos que componen el mismo: variables de entrada, de salida, operadores utilizados y reglas. Así mismo, una vez completada la construcción se podrá ver, de forma dinámica, el comportamiento global del sistema a nivel de inferencia al variar la entrada.

Los pasos llevados a cabo por el usuario (alumno) al utilizar el software son las siguientes:

1. **Crear base de conocimiento vacía:** Sobre ésta se crearán todos los elementos del sistema fuzzy.

2. **Definición de variables de entrada:** Crear las variables de entrada que formarán parte del sistema, sus características y particiones. Puede verse esta ventana en la Figura 3.

3. **Definición de variables de salida:** Crear las variables de salida a obtener con el sistema y características de las mismas.

4. **Selección de operadores:** Seleccionar las t-normas, t-conormas y operador de defuzzificación a utilizar.

5. **Creación de reglas:** Construir las reglas que relacionarán la entrada con la salida.

6. **Inferencia:** Se permitirá la manipulación de la entrada para obtener la salida en cada caso con el sistema construido. Esta fase será fundamental para la comprensión del sistema construido y para familiarizar al alumno con los sistemas fuzzy.

3.2 CREACIÓN DE UN SISTEMA INDUCIDO DE UN GRUPO DE DATOS

Este segundo tutorial, más avanzado que el anterior, proporciona al alumno una visión de la creación de sistemas fuzzy a partir de un conjunto de datos de entrenamiento a los que ajustarse.

Dado el carácter didáctico del tutorial la experimentación correspondiente a esta parte se realizará utilizando el conjunto de datos Iris (disponible en *http://archive.ics.uci.edu/ml/datasets/Iris*), que es uno de los conjuntos de datos más simples y conocidos en la comunidad científica. Se trata de un problema de clasificación que cuenta con 150 instancias (datos de flores), con 4 variables (largo y ancho de pétalo y sépalo), que han de ser clasificadas en una de las 3 clases propuestas (tipos de flores).

Las fases guiadas de este tutorial son las siguientes:

1. **Carga de datos:** Inicialmente se cargarán los datos del conjunto Iris. La creación de ficheros, carga de datos y separación en train y test será transparente para el usuario. Se utilizarán el 80 % de los datos para train y el restante 20 % para test.

2. **Visualización de los datos:** Una vez se han cargado los datos, el usuario obtendrá un resumen detallado de los mismos.

3. **Generación de particiones:** Se generarán las particiones de 3 etiquetas lingüísticas que mejor se ajusten a los datos. Se seleccionarán particiones regulares, kmeans o hfp para buscar la que mejor se ajusta a los datos.[5]

[5]Pueden ampliarse detalles de los métodos mencionados en [3, 4]

4. **Selección de operadores:** Seleccionar las t-normas, t-conormas y operador de defuzzificación a utilizar.

5. **Generación de reglas:** Se inducirán reglas utilizando el algoritmo de Wang and Mendel, al ser el más sencillo.

6. **Inferencia:** Se permitirá la manipulación de la entrada para obtener la salida en cada caso con el sistema construido. Se puede ver un ejemplo de la herramienta junto al tutorial proporcionado en la Figura 4.

7. **Evaluación:** Se obtendrá la precisión (de train y test) y la interpretabilidad con distintos índices para el sistema creado.

4 CONCLUSIONES

Este trabajo presenta un nuevo conjunto de tutoriales interactivos para el diseño de sistemas fuzzy para usuarios sin conocimientos avanzados de lógica fuzzy. Estos tutoriales trabajan junto a la herramienta GUAJE aumentando su usabilidad no sólo en el ámbito didáctico sino también en todo tipo de aplicaciones prácticas.

La propuesta realizada reduce la curva de aprendizaje de los usuarios inexpertos. Así mismo, los usuarios, estudiantes u otros, que tengan poca experiencia en el diseño de sistemas fuzzy centrarán su atención en el diseño, obviando las tareas administrativas propias de la herramienta.

De igual forma, las ventanas y diálogos creados facilitan el uso del software guiando todo el proceso y reduciendo el tiempo empleado en interactuar con la herramienta. La disponibilidad de un proceso guiado reduce el número de errores y automatiza las tareas a cubrir.

A menudo, los usuarios de GUAJE diseñan sistemas fuzzy para resolver problemas cotidianos del mundo que los rodea, y para ello cuentan con la inestimable ayuda de expertos de dominio (que desafortunadamente en la mayoría de casos no saben nada de lógica fuzzy). Por ejemplo, en el diseño de sistemas inteligentes para diagnóstico médico la interpretabilidad de los sistemas fuzzy es muy apreciada, pero los potenciales usuarios no se pueden considerar expertos en lógica fuzzy. Por ello, herramientas como GUAJE tienen un potencial nicho de aplicación muy importante.

Se propone ampliar el número de tutoriales, creando materiales específicos para distintos métodos proporcionados por la herramienta. Por ejemplo, la utilización de GUAJE para crear dos clasificadores basados en reglas fuzzy en los que poder apreciar la diferencia de interpretabilidad y precisión de ambos e ilustrar así la importancia del diseño para conseguir un sistema fuzzy interpretable.

Agradecimientos

Este trabajo ha sido financiado por la Fundación para el progreso del Soft Computing (Mieres - Asturias, España), y el Ministerio de Ciencia e Innovación Español a través de los proyectos de investigación fundamental no orientada TIN2011-29824-C02-01 y TIN2008-06890-C02-01.

Referencias

[1] J. M. Alonso and L. Magdalena. GUAJE - a java environment for generating understandable and accurate models. In *XV Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)*, pages 399–404, 2010.

[2] J. M. Alonso and L. Magdalena. Generating understandable and accurate fuzzy rule-based systems in a java environment. In A.M. Fanelli, W. Pedrycz, and A. Petrosino, editors, *Lecture Notes in Artificial Intelligence - Fuzzy Logic and Applications (9th International WILF Workshop)*, volume LNAI6857, pages 212–219, 2011.

[3] J. M. Alonso and L. Magdalena. HILK++: an interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers. *Soft Computing*, 15(10):1959–1980, 2011.

[4] J. M. Alonso, L. Magdalena, and S. Guillaume. HILK: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *Int. Journal of Intelligent Systems*, 23(7):761–794, 2008.

[5] F. Campbell. The Bologna process & higher education in the context of Europe and globalization. *Bologna Process Official Seminar: "The Cultural Heritage and Academic Values of the European University and the Attractiveness of the EHEA"*, 2006.

[6] M. Kelo. *The Future of University. Translating Lisbon into Practice*. Bonn: Lemmens Verlags & Mediengesellschaft mbH, 2006.

[7] R. Sedgwick. Bologna declaration addresses higher education in european union. *World Education News & Reviews*, 13(1), 2000.

[8] R. Sedgwick. The Bologna process: How it is changing the face of higher education in europe. *World Education News & Reviews*, 14(2), 2001.

[9] B. Wächter. The Bologna process: developments and prospects. *The European Journal of Education*, 39(3), 2004.

Números borrosos y principio de extensión de Zadeh

Ana Belén Ramos-Guajardo, Itzíar García-Honrado

Universidad de Oviedo. Departamento de Estadística, Investigación Operativa y Didáctica de las Matemáticas,
ramosana@uniovi.es,garciaitziar@uniovi.es

Resumen

En este artículo se pretende explicar, dentro del marco de la enseñanza universitaria, la utilidad de los números borrosos y el posterior cálculo de operaciones básicas entre ellos, basándonos en el Principio de Extensión, introducido por Zadeh, y la representación por α-cortes de un número borroso.

Esto se ejemplificará sobre un problema real cuyos datos se representan por números borrosos, de los cuales se calculará la media aritmética utilizando la suma y el producto por escalares.

Palabras Clave: Números borrosos, Alfa cortes, Principio de Extensión.

1. INTRODUCCIÓN

Para representar fenómenos que albergan imprecisión es usual el empleo de conjuntos borrosos, que rompen las barreras de los conjuntos clásicos cuya principal característica es que dado un elemento cualquiera del Universo de Discurso, X, podemos decir si tal elemento pertenece o no al conjunto. La lógica borrosa recoge la posibilidad de que los elementos pertenezcan al conjunto con cierto grado. Un conjunto borroso, U, se puede representar por la función

$$U : X \to L,$$

que asigna a cada elemento el grado en que pertenece al conjunto. Dicho grado se define sobre el conjunto parcialmente ordenado, L, siendo L el intervalo unidad en el caso de los conjuntos borrosos definidos por Zadeh [16]. En este caso particular existe un orden total en L que hace que todos los elementos del universo de discurso X sean comparables a través de su imagen por U.

El Universo de Discurso, X, puede ser cualquier población. Por ejemplo, se puede considerar como Universo de Discurso a los habitantes de Valladolid. Dentro de este universo se puede definir el colectivo (o conjunto borroso) de aquellos que son altos. Dado que alto es un predicado gradual, no podemos discriminar clásicamente aquellos que son altos de los que no haciendo una partición del universo en dos conjuntos. Por lo tanto, definiremos un conjunto borroso que a cada vallisoletano le asigne un grado en una escala L de "ser alto". No obstante, la altura se puede caracterizar por una magnitud numérica definida dentro del intervalo $[0, 250]$ (dado en centímetros), y así podemos simplificar el universo de discurso complejo de los habitantes de Valladolid en un intervalo de números reales.

Respecto a las posibles escalas L en las que se puede definir el grado de pertenecia, nos encontramos con un problema de diseño para cada caso particular que es la elección de la escala $(L, \leq)$ más conveniente para el problema. Ésta es una elección no trivial y en los problemas tecnológicos suele ser una escala numérica. En este trabajo vamos a utilizar la escala $L = [0, 1]$.

Siendo el conjunto $L = [0, 1]$, manejamos la definición de conjunto borroso dada por Zadeh. Dentro de los conjuntos borrosos trabajaremos con un subconjunto de estos, los números borrosos, cuya definición introduciremos posteriormente. Un ejemplo de número borroso podría ser "Aproximadamente 3". Cabe notar también que los números precisos son un caso particular de números borrosos. Así, se establecería la relación que se presenta en la Figura 1.

Una vez diseñados los elementos de nuestro modelo, surge la necesidad de operar con ellos. Con este fin, se estudiará el Principio de Extensión definido inicialmente por Zadeh [17] que nos permitirá definir ciertas operaciones entre números borrosos de un modo operativo.

También se analizará un caso práctico de la vida real en el que se justificará el uso de los números borrosos para construir un modelo que se ajuste al problema anterior. Además, se mostrará una simplificación del Principio de

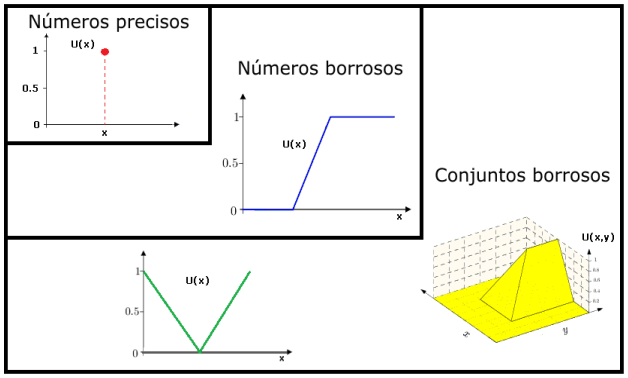


Figura 1: Conjuntos borrosos.

Extensión en el caso particular de trabajar con números borrosos.

Todos las nociones presentadas en este trabajo se dirigen a estudiantes universitarios con el fin de iniciarlos en el ámbito de la lógica borrosa. Los conceptos previos que se requieren para entender los conocimientos aquí presentados se centran en el manejo del análisis matemático (sobre todo el manejo de funciones reales) a nivel de enseñanza secundaria.

2. EL PROBLEMA

Una de las ventajas del uso de números borrosos radica en aplicaciones del mundo real de los mismos. En muchos experimentos reales aparecen diferentes tipos de incertidumbre, en el sentido general de falta de conocimiento preciso. Uno de ellos es la aleatoriedad o incertidumbre estocástica del resultado experimental que se formaliza a partir del Cálculo de Probabilidades y que aparece a la hora de seleccionar una muestra a partir de una población dada. Otro tipo es la imprecisión o incertidumbre no estocástica que puede deberse a la falta de información, a una percepción no exacta de la realidad, a un proceso de agregación, etc. Por ejemplo, ésta se presenta al evaluar subjetivamente la calidad de un determinado producto.

Con el fin de manejar tal imprecisión, en muchas ocasiones ésta se resume en un valor real. Por ejemplo, la descripción de las percepciones de los expertos relativas a un determinado producto vienen dadas en general por un número real. Concretamente se suele utilizar la llamada escala de Likert (ver [9]) que asocia valores numéricos a categorías (generalmente de 1 a 5, donde 1 significa calidad muy baja y 5 calidad muy alta).

El problema radica en que a veces los expertos tienen dificultades a la hora de resumir sus percepciones con un único valor y, además, muchas veces dos expertos diferentes pueden no percibir una categoría de la misma manera. En este contexto, se propuso a los expertos flexibilizar sus opiniones por medio del empleo de conjuntos borrosos.

El concepto de conjunto borroso fue introducido por Zadeh [16], y ha sido estudiado ampliamente a lo largo de los últimos 30 años (ver, por ejemplo, [3, 4, 17, 18]). Inicialmente, el principal interés que presentaba la teoría de conjuntos borrosos se centraba en la representación de la incertidumbre de los procesos cognitivos humanos (ver [16]). Posteriormente ha tenido aplicaciones en muchos campos tales como la Estadística (ver [5, 6, 12]), la Inteligencia Artificial, Control de Sistemas ([10, 14]), ...

En este trabajo se incluirá la definición general de número borroso y se presentarán de un modo didáctico el principio de extensión introducido por Zadeh [16] y las principales operaciones entre números borrosos que se derivan de tal principio.

El enfoque se centra en el empleo de los números borrosos como un modo de resumir la información recogida en las percepciones personales acerca de determinadas cualidades de un objeto concreto. Ese es el caso, por ejemplo, del análisis sensorial relativo a las características de un producto. El análisis sensorial se ha utilizado con el fin de interpretar las percepciones de las personas en relación con determinadas características de un producto por medio de los sentidos de la vista, el olfato, el tacto, el gusto y el oído (ver, por ejemplo, [2]).

Las propiedades sensoriales de un producto determinan la selección y preferencias de los consumidores, por lo que resulta esencial desarrollar métodos que permitan medir estas propiedades de un modo adecuado. En este contexto, la empresa ALCE CALIDAD (ver http://www.alcecalidad.com/?q=node/28) lleva a cabo sesiones de cata con catadores expertos con el fin de certificar la Denominación de Origen del queso Gamonedo, fabricado en diversas áreas de Asturias (España). Las dificultades para conservar la calidad del queso debido a su sensibilidad y complejidad hacen necesario un sistema de catas robusto para estimar su calidad, lo que conllevaría a una producción más eficiente así como a una mejor calidad del producto. En la Fundación para el Progreso del Soft Computing se ha llevado a cabo un proyecto relativo a este estudio conjunto con el instituto ALCE CALIDAD.

En tal proyecto se ha extendido la escala categórica usual (de 1 a 7 en este caso) al trabajo con números borrosos. El procedimiento está basado en previos estudios llevados a cabo durante el II Curso de Verano de la Fundación para el Progreso del Soft Computing (Mieres, España) en Julio de 2008 para evaluar la calidad de los cursos (como ha sido introducido en [7]).

Para determinar si el queso Gamonedo merece mantener la Denominación de Origen, es necesario desarrollar un análisis estadístico profundo teniendo en cuenta las percepciones borrosas proporcionadas por los expertos, aunque

este tema no pertenece al ámbito del presente artículo.

Sin embargo, posteriormente se mostrará el cálculo de la media muestral teniendo en cuenta algunos de los datos obtenidos en el estudio anterior.

3. NÚMERO BORROSO

Un conjunto borroso $U \in [0,1]^{\mathbb{R}}$ (ver Figura 2) es un número borroso siempre que en un intervalo cerrado $[a,b] \subset \mathbb{R}$ se verifique:

- U(x)=1, para todo $x \in [a,b]$.
- Que exista una función $I : (-\text{ínf}, a) \to [0,1]$ que sea continua y no decreciente y tal que $U(x) = I(x)$, para todo $x \in (-\text{ínf}, a)$.
- Que exista una función $D : (b, +\text{ínf}) \to [0,1]$ que sea continua y no decreciente y tal que $U(x) = D(x)$, para todo $x \in (b, +\text{ínf})$.

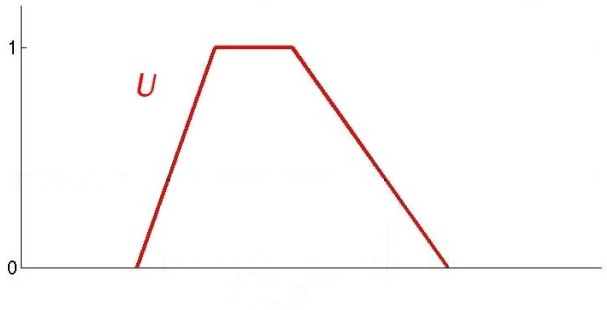


Figura 2: Número borroso.

Podemos representar los números borrosos a través de los llamados α-*cortes* o α-*niveles* del siguiente modo. Un número borroso U puede considerarse como una función $U : \mathbb{R} \to [0,1]$ donde cada α-corte

$$U_\alpha = \{x \in \mathbb{R}^p | U(x) \geq \alpha\}$$

es un intervalo compacto de la rectal real para todo $\alpha \in (0,1]$. Además, el 0-corte de U, U_0, será la clausura del soporte de U.

Así, en este caso un número borroso U va a quedar caracterizado por sus α-cortes. Es decir, dado U son conocidos sus α-cortes y, por otro lado, si conocemos todos los α-cortes de U, entonces podemos determinar quién es U. En la Figura 3 se ha representado un número borroso (en este caso trapezoidal) así como el correspondiente 0.4-corte.

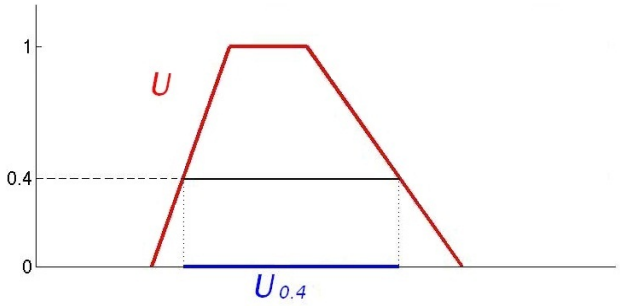


Figura 3: 0.4-corte del número borroso U.

4. OPERACIONES ENTRE NÚMEROS BORROSOS: EL PRINCIPIO DE EXTENSIÓN DE ZADEH

El uso de números borrosos requiere la consideración de operaciones elementales (especialmente la suma y el producto por valores reales) entre esos datos.

La mayoría de operaciones entre números naturales, enteros, reales,... son manejables, basándonos en ello, a partir del llamado *Principio de extensión de Zadeh* [17], definimos las operaciones entre números borrosos (ver Figura 4).

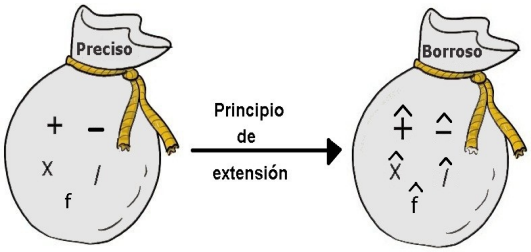


Figura 4: Número borroso.

Si tenemos una aplicación, que opera los elementos de X dando como resultado elementos de Y, $f : X \to Y$, un conjunto preciso (clásico) contenido en X, $A \subset X$, se tiene que $f(A) = \{y \in Y; f(a) = y, a \in A\}$ es la f-imagen de A en Y.

Se puede ver que

$$U_{f(A)}(y) = \sup\{U_A(x); f(x) = y\} =$$

$$\begin{cases} 1, & \text{si existe } x \in A \text{tal que } f(x) = y, \\ 0, & \text{en otro caso.} \end{cases}$$

Basándose en esta imagen, la aplicación $f : X \to Y$ se extiende al respectivo conjunto de partes por

$$\widehat{f} : \mathscr{P}(X) \to \mathscr{P}(Y), \quad A \mapsto f(A).$$

Del mismo modo dada la aplicación $f : X \to Y$, se puede extender al conjunto de fuzzy sets $[0,1]^X, [0,1]^Y$ por

$$\widehat{f} : [0,1]^X \to [0,1]^Y,$$

$$\widehat{f}(\mu)(y) = \sup\{\mu(x); f(x) = y\}, \text{ for all } y \in Y,$$

y $\widehat{f}$ es la extensión de f al conjunto de los fuzzy sets y la definición del *Principio de Extensión*.

Ejemplo 4.1. *Ejemplo extendiendo la operación +: Calculemos la suma para los siguientes conjuntos borrosos definidos sobre los números naturales $X = N$, en la escala $L = [0,1]$:*

- $U = 0/0 + 1/1 + 0/3 + 0/4$
- $V = 0/0 + 0/1 + 0.5/2 + 1/3 + 0/4$

$$U \widehat{+} V = 0/0 + 0/1 + 0/2 + 0.5/3 + 1/4 + 0.5/5 + 0/6$$

Dado que

- $(U \widehat{+} V)(0) = \underset{0=x+y}{Sup\ min}(U(x), V(y)) = Sup(min(U(0), V(0)))$.
- $(U \widehat{+} V)(1) = \underset{1=x+y}{Sup\ min}(U(x), V(y)) = Sup(min(U(1), V(0)), min(U(0), V(1))) = Sup(0,0) = 0$.
- $(U \widehat{+} V)(2) = \underset{2=x+y}{Sup\ min}(U(x), V(y)) = Sup(min(U(2), V(0)), min(U(0), V(2)), min(U(1), V(1))) = Sup(0,0,0) = 0$.
- $(U \widehat{+} V)(3) = \underset{3=x+y}{Sup\ min}(U(x), V(y)) = Sup(min(U(3), V(0)), min(U(0), V(3)), min(U(1), V(2)), min(U(2), V(1))) = Sup(0,0,0.5,0) = 0.5$.
- $(U \widehat{+} V)(4) = \underset{3=x+y}{Sup\ min}(U(x), V(y)) = Sup(min(U(4), V(0)), min(U(0), V(4)), min(U(1), V(3)), min(U(3), V(1)), min(U(2), V(2))) = Sup(0,0,1,0.5,0) = 1$.
- ...

En el siguiente ejemplo se mostrará la suma con sigo mismo del número borroso 3 representado en la figura 5, pudiéndose comprobar que el número borroso resultante es el representado en la Figura 5 que se considera una posible representación del número borroso 6.

Se observa que el número borroso resultante recoge la imprecisión de los dos conjuntos borrosos que se suman.

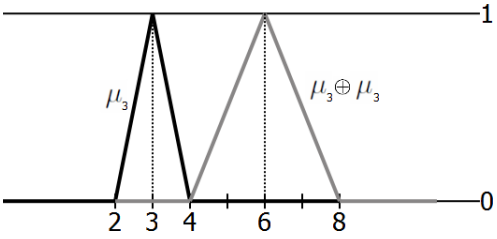


Figura 5: Aproximadamente 6

En el caso de tratar con números borrosos, los conjuntos X, Y serán conjuntos de números reales y definimos, como habitualmente las funciones f representarán una operación entre conjuntos números X será producto cartesiano de números reales, así se define,

Definición 4.2. *Principio de extensión para números borrosos: Sea $f : \mathbb{R}^k \longrightarrow \mathbb{R}$, y sean U_i números borrosos en $\mathbb{R}$, para $i \in \{1,\ldots,k\}$. Se define la* imagen (borrosa) de $(U_1,\ldots,U_k)$ inducida *por f, $V = f(U_1,\ldots,U_k)$, como un conjunto borroso en $\mathbb{R}$ dado por:*

$$V(y) = \widehat{f}(U_1,\ldots,U_k)(y)$$

$$= \begin{cases} \sup_{\substack{(x_1,\ldots,x_k)| \\ f(x_1,\ldots,x_k)=y}} \text{mín}\{U_1(x_1),\ldots U_k(x_k)\} & si\ f^{-1}(\{y\}) \neq \emptyset \\ 0 & si\ f^{-1}(\{y\}) = \emptyset. \end{cases}$$

En virtud de las conclusiones derivadas a partir de Nguyen [11], la aritmética entre números borrosos de $\mathbb{R}$ puede reducirse a la aritmética entre intervalos, en el sentido de que para $\alpha \in [0,1]$ los α-cortes de los resultados de las operaciones entre números borrosos pueden expresarse en términos de operaciones de sus α-cortes, que son intervalos reales. De este modo, si U y V son dos números borrosos, entonces se tiene que para todo $\alpha \in [0,1]$ y $\lambda \in (\mathbb{R} - \{0\})$ que:

$$(U+V)_\alpha = U_\alpha + V_\alpha = [\text{ínf}\, U_\alpha + \text{ínf}\, V_\alpha, \sup U_\alpha + \sup V_\alpha],$$

$$(\lambda U)_\alpha = \lambda U_\alpha = \begin{cases} [\lambda \cdot \text{ínf}\, U_\alpha, \lambda \cdot \sup U_\alpha], & \text{si } \lambda > 0, \\ [\lambda \cdot \sup U_\alpha, \lambda \cdot \text{ínf}\, U_\alpha], & \text{si } \lambda < 0. \end{cases}$$

Para entender mejor las operaciones que se acaban de definir, es necesario reseñar que un α-corte U_α viene dado en función del ínfimo y supremo del intervalo que forma, es decir:

$$U_\alpha = [\text{ínf}\, U_\alpha, \sup U_\alpha].$$

En la práctica la metodología a seguir para operar con números borrosos se centra en las correspondientes operaciones con los α-cortes de dichos números.

En las Figuras 6 y 7 se muestran la suma de dos números borrosos, así como el producto por un escalar negativo de uno de ellos.

De este modo, si por ejemplo $U_\alpha = [2,3]$ y $V_\alpha = [2.5,5]$, se verifica que $(U+V)_\alpha = [4.5,8]$ y $-2 \cdot U_\alpha = [-6,-4]$.

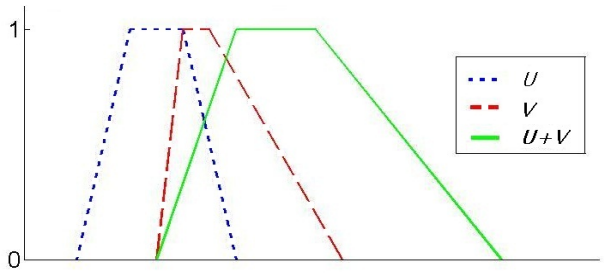


Figura 6: Suma de dos números borrosos U y V.

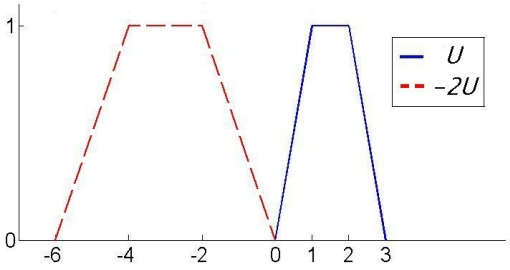


Figura 7: Producto de U por el escalar -2.

5. Aplicación práctica

Durante el estudio mencionado en la Sección 2 se han recogido una serie datos en relación con la opinión de un catador experto sobre la calidad del olor de una muestra de 25 quesos diferentes de la marca Gamonedo. Los datos vienen dados por números borrosos trapezoidales y se denotarán por $Tra(\text{ínf}U_0, \text{ínf}U_1, \sup U_1, \sup U_0)$, donde $[\text{ínf}U_0, \sup U_0]$ y $[\text{ínf}U_1, \sup U_1]$ son el 0-corte y 1-corte de U, respectivamente. Tales datos se muestran en la Tabla 1.

Las operaciones introducidas a lo largo de la presente sección nos permitirán calcular la media muestral de los datos anteriores de un modo operativo. La media muestral de un conjunto de n números borrosos, $\{U^1, \ldots, U^n\}$, viene dada por la suma de los n números borrosos multiplicada por el escalar $1/n$. En este caso particular en el que estamos trabajando con números borrosos trapezoidales, la media muestral de todos ellos será aquel número borroso trapezoidal $\overline{U} = Tra\left(\overline{\text{ínf}U_0}, \overline{\text{ínf}U_1}, \overline{\sup U_1}, \overline{\sup U_0}\right)$ cuyos parámetros, $\overline{\text{ínf}U_0}$, $\overline{\text{ínf}U_1}$, $\overline{\sup U_1}$ y $\overline{\sup U_0}$, son las correspondientes medias muestrales de los parámetros correspondientes de $\{U^1, \ldots, U^n\}$.

Así, la media muestral de los datos recogidos en la Tabla 1, vienen dados por el número borroso trapezoidal $Tra(56.4, 61.32, 66.84, 73.48)$.

Opinión	Experto 1
1	$Tra(47,50,53,60)$
2	$Tra(65,67,70,80)$
3	$Tra(50,55,60,65)$
4	$Tra(70,74,80,90)$
5	$Tra(50,60,67,75)$
6	$Tra(40,47,53,60)$
7	$Tra(74,80,84,90)$
8	$Tra(64,68,72,80)$
9	$Tra(70,74,80,90)$
10	$Tra(50,54,60,65)$
11	$Tra(50,55,60,66)$
12	$Tra(64,70,74,80)$
13	$Tra(36,44,50,56)$
14	$Tra(60,65,74,80)$
15	$Tra(70,75,82,86)$
16	$Tra(75,80,86,90)$
17	$Tra(45,49,56,60)$
18	$Tra(73,78,83,86)$
19	$Tra(50,55,60,68)$
20	$Tra(44,49,53,60)$
21	$Tra(50,50,56,66)$
22	$Tra(53,57,61,70)$
23	$Tra(70,77,84,84)$
24	$Tra(30,36,43,50)$
25	$Tra(60,64,70,80)$

Tabla 1: Muestra de opiniones de un experto acerca de la calidad del olor de varios quesos

6. Conclusiones

Este artículo no pretende ser más que una explicación del Principio de Extensión de Zadeh a nivel universitario, enmarcándolo dentro de un contexto en el que se requiere realizar la suma de números borrosos con fines estadísticos.

En primer lugar, se muestra la necesidad de trabajar con números borrosos en lugar de con números reales porque nos permiten representar la imprecisión presente en problemas cotidianos donde se trabaja con percepciones humanas.

En segundo lugar, se muestra la herramienta que nos permite operar con números borrosos a partir de las operaciones con números reales, el conocido principio de extensión de Zadeh.

Por último, se ejemplifican los cálculos de suma y producto por escalares de números borrosos en el cómputo de la media de una colección de datos.

Como proyecto futuro, se podrían analizar dentro del mismo contexto problemas en los que la escala $(L, \leq)$ no fuera el intervalo $[0,1]$, analizando las ventajas e inconvenientes que conllevaría tal elección.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la fundación para el progreso del Soft Computing (ECSC), la Universidad de Oviedo, y por el proyecto de Ministerio de Ciencia e Innovación (MICINN) con referencia TIN2008-06890-CO2-01.

Referencias

[1] C. Alsina, E. Trillas: On the symmetric difference of fuzzy sets. *Fuzzy Sets and Systems* 153, pp. 181–194, 2005.

[2] M.A. Amerine, R.M. Pangborn, E.B. Roessler: *Principles of sensory evaluation of food*. Academic Press, New York, 1965.

[3] R. Bellman, L.A. Zadeh: Decision-making in a fuzzy environment. *Management Science*, 17, pp. 141–164, 1970.

[4] D. Dubois, H. Prade, H. Farreny, R. Martin-Clouaire, C. Testemale: *Possibility Theory–An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[5] G. González-Rodríguez, M. Montenegro, A. Colubi, M.A. Gil: Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data. *Fuzzy Sets and Systems*, 157, pp. 2608–2613, 2006.

[6] G. González-Rodríguez, A. Blanco, N. Corral, A. Colubi: Least squares estimation of linear regression models for convex compact random sets. *Advances in Data Analysis and Classification*, 1, pp. 67–81, 2007.

[7] G. González-Rodríguez, A. Colubi, M.A. Gil: Fuzzy data treated as functional data: a one-way ANOVA test approach. ´Comput. Statist. Data Anal., doi:10.1016/j.csda.2010.06.013.

[8] P.C. Fishburn: *Utility Theory for Decision Making*. Wiley, Nueva York, 1970.

[9] R. Likert: A technique for the measurement of attitudes. *Archives of Psychology*, 140, pp. 1–55, 1932.

[10] E.H. Mamdami: Advances in the linguistic synthesis of fuzzy controllers. *Int. Jr. Man-Machine Studies*, 8, pp. 669–678, 1976.

[11] H.T. Nguyen: A note on the extension principle for fuzzy sets. *J. Math. Anal. Appl.*, 64, pp. 369–380, 1978.

[12] M.L. Puri, D.A. Ralescu: Fuzzy random variables. *Journal of Mathematical Analysis and Applications*, 114, pp. 409–422, 1986.

[13] A.B. Ramos-Guajardo, M.A. Lubiano: K-sample tests for equality of variances of random fuzzy sets. *Comput. Statist. Data Anal.*, doi:10.1016/j.csda.2010.11.025.

[14] M. Sugeno, T. Murofushi: Helicopter flight control based on fuzzy logic. In: *Proc. 1st Int. Fuzzy Engineering Symposium*, pp. 1120–1121, Yokohama, Japan, 1991.

[15] R.R. Yager: E-Z OWA weights. En: *Proceedings of the 10th IFSA World Congress*, Estambul (Turquía), pp. 39–42, 2003.

[16] L.A. Zadeh: Fuzzy sets. *Information and Control*, 8 (3), pp. 338–353, 1965.

[17] L.A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning, Part 1. *Information Sciences*, 8, pp. 199–249; Part 2. *Information Sciences*, 8, 301–353; Part 3. *Information Sciences*, 9, pp. 43–80, 1975.

[18] H.J. Zimmermann: *Fuzzy Set Theory and its Applications*. Third edition. Ed. Kluwer Academic Publishers, 1996.

Usos docentes del análisis exploratorio gráfico de datos borrosos

Luciano Sánchez[1], **Inés Couso**[2], **José Otero**[1]

[1] *Universidad de Oviedo, Departamento de Informática, 33204 Gijón, España*, [luciano,jotero]@uniovi.es
[2] *Universidad de Oviedo, Departamento de Estadística, 33204 Gijón, España*, couso@uniovi.es

Resumen

El análisis exploratorio gráfico para datos borrosos permite representar en un mapa a un conjunto de individuos, con características percibidas de forma imprecisa, de modo que sus disimilaridades sean compatibles con las distancias entre sus representaciones. En este estudio se discutirá el uso de esta herramienta para analizar de forma conjunta la evolución de un grupo de estudiantes durante un curso, y para seleccionar el personal de una empresa más adecuado para recibir un curso de formación, de acuerdo con un catálogo de competencias y considerando la fiabilidad de las fuentes de información.

Palabras Clave: Análisis Exploratorio Gráfico, Datos de Baja Calidad, Imprecisión de Datos.

1. Introducción

El análisis exploratorio gráfico consiste en la proyección de un conjunto de individuos en un plano, de forma que las similaridades entre pares de individuos sean compatibles con las distancias entre sus correspondientes representaciones [6]. Existen distintas técnicas para realizar esta proyección, en función del modelo empleado para asociar las similaridades entre objetos y las distancias entre sus representaciones en el mapa. El método más frecuente se aplica a objetos descritos mediante un vector de propiedades numéricas, y emplea la distancia euclídea tanto para calcular las distancias entre los objetos como entre sus proyecciones.

Ha habido diferentes generalizaciones del análisis exploratorio gráfico a datos imprecisos, y en particular al caso en que cada objeto se caracteriza mediante un vector de números borrosos [2, 3, 4, 10]. Así como en el análisis exploratorio de datos nítidos cada individuo se asocia a un punto del mapa, la proyección de un vector borroso es una figura geométrica que depende de las transformaciones entre las distancias espaciales y las distancias en el mapa.

Las técnicas de análisis exploratorio de datos nítidos son parte del conocimiento general y se emplean de forma rutinaria en multitud de problemas de descubrimiento de conocimiento [8]. Sin embargo, el uso del análisis exploratorio de datos difusos no está muy extendido. En relación con la minería de información en un contexto docente, el algoritmo definido en [10] se ha aplicado al análisis de tests resueltos por grupos de niños con dificultades de aprendizaje (diagnóstico precoz de la dislexia [9]) y al análisis de cuestionarios de seguimiento realizados por los alumnos de diferentes estudios de grado y master universitario. En ambos casos, además de obtenerse una información visual acerca de cuántos tipos diferentes de estudiantes hay en cada grupo, es posible medir las variaciones entre el éxito del aprendizaje de los diferentes subgrupos, y la evolución de sus rendimientos relativos durante el transcurso del curso. En este estudio se revisan estas aplicaciones y se introduce una nueva, la selección de las personas idóneas para asistir a cursos de formación dentro de una empresa, teniendo en cuenta para ello diferentes fuentes de información acerca de las competencias de cada empleado (declaración en el curriculum vitae, exámenes internos, cuestionarios, trabajos desempeñados con anterioridad, etc.) que están, a su vez, afectadas por la credibilidad de cada fuente.

La estructura de este estudio es la siguiente: en la sección 2 se describe la utilidad de las técnicas de análisis gráfico de datos en un contexto docente. En la sección 3 se discute la necesidad de una representación basada en conjuntos borrosos, y se describen algunos aspectos técnicos del algoritmo. En la sección 4 se discuten varios casos prácticos. El trabajo concluye en la sección 5, con las conclusiones y el trabajo futuro.

2. Utilidad del análisis exploratorio gráfico en problemas docentes

Al organizar cursos de formación para los empleados de una empresa, es importante estudiar qué carencias de formación deben suplirse, para así poder elegir los contenidos más adecuados para los cursos. Un estudio similar sirve también para elegir a los mejores asistentes a un curso con un número de plazas limitado, o bien, comparando los resultados del estudio antes y después de finalizar el curso, para comprobar si lo enseñado en el mismo se traduce en una mejora de la capacidad global de la empresa.

Como se ha mencionado en la introducción, los diferentes algoritmos de análisis exploratorio gráfico (mapas de Sammon, Análisis de Componentes Principales (PCA), Escalado Multidimensional (MDS), Mapas Autoorganizados (SOM), etc. [2]) son técnicas estadísticas que projectan objetos como puntos en un plano, de forma que la proximidad de las proyecciones de dos objetos en el mapa refleja la similitud entre las propiedades de los mismos. Si cada uno de estos objetos se asocia con un empleado, y se admite que las propiedades del mismo consisten en mediciones numéricas de su nivel de aptitud en un catálogo de competencias, el análisis exploratorio gráfico es una de las técnicas más adecuadas para evidenciar grupos de empleados con diferentes habilidades. Añadiendo individuos ficticios (empleados hipotéticos con conocimiento perfecto de unas tecnologías y nulo acerca de otras) las posiciones de los empleados reales con respecto a las de los empleados ficticios permiten detectar carencias de formación. Por último, comparando varios mapas de los mismos individuos en diferentes fechas, es posible evaluar el impacto de los cursos recibidos.

No obstante lo anterior, estas técnicas no son directamente aplicables al caso de que haya incertidumbre en alguna de las propiedades de un individuo (por ejemplo, un valor perdido) y además no contemplan la fiabilidad de las diferentes fuentes de información empleadas para caracterizar los individuos. Por ejemplo, una de las fuentes más accesibles para comprobar el nivel de formación de un grupo es el cuestionario de seguimiento [7]. A diferencia del examen o la entrevista, en esta modalidad es el propio estudiante o empleado quien declara sus conocimientos, por lo que esta información puede ser inexacta: un individuo puede declarar un conocimiento "avanzado" de inglés en el curriculum o haber superado un examen de ese nivel; en el primer caso, la incertidumbre acerca del dominio del idioma es mayor.

Una forma sencilla de conocer la incertidumbre asociada a un cuestionario consiste en asociar diferentes preguntas a un mismo item; la dispersión de las respuestas correspondientes es una indicación de la fiabilidad del test. Por ejemplo, si se encuesta a un estudiante acerca de sus conocimientos de teoría de la probabilidad, éste puede declarar un conocimiento "alto" del conocimiento del concepto "función de densidad" y "nulo" de "derivada de Radon-Nikodym", mientras que otro puede responder "medio" a ambas preguntas: la dispersión de las respuestas asociadas a un mismo item es una indicación de la fiabilidad de las mismas. Es evidente que, si se elige una medida de tendencia central para resumir todas las respuestas asociadas a una misma propiedad, se pierde una información valiosa.

Otro problema frecuente de la fase de recogida de información está en los valores perdidos [5]. Las soluciones más frecuentes son la eliminación del individuo del análisis o la imputación de un valor a la característica desconocida. Esta última es la solución preferida en organizaciones con pocos miembros, y generalmente consiste en buscar los individuos más cercanos, para calcular el promedio de los valores de la propiedad en cuestión para estos. De nuevo, la variabilidad de estos valores está siendo ignorada, lo cual puede significar que el valor imputado está lejos de la realidad, con lo que el análisis estaría siendo falseado.

3. Uso de conjuntos borrosos en el análisis de competencias

El uso de conjuntos borrosos permite dotar de una representación homogénea a todos los tipos anteriores de incertidumbre en los datos. Por medio de la interpretación posibilística de la función de pertenencia de un conjunto borroso [1], cada atributo de un objeto se puede asociar a un conjunto borroso $\widetilde{X}$, cuyos cortes $[\widetilde{X}]_\alpha$ son intervalos de confianza anidados, con $P(x \in [\widetilde{X}]_\alpha) \geq 1-\alpha$. Así, por ejemplo, un valor perdido puede reemplazarse con un conjunto borroso que modele la distribución de la característica correspondiente en otros objetos similares (aun cuando estos, a su vez, sean percibidos de forma imprecisa), y las declaraciones de un conocimiento "avanzado" de inglés o "medio" de teoría de la probabilidad estarán asociadas a sendos números borrosos, cuyas especificidades estarán vinculadas a las fiabilidades de las fuentes de información.

Por tanto, en este estudio se considerará que el conocimiento acerca de un individuo puede cuantificarse mediante un vector de números borrosos. Esta generalización tiene dos consecuencias fundamentales de cara a realizar un análisis gráfico de la población:

1. Las coordenadas espaciales de cada individuo son desconocidas, salvo por una familia anidada de intervalos de confianza multivariados.

2. La distancia euclídea entre dos individuos cuyas coordenadas son inciertas, es incierta a su vez.

De la primera afirmación puede concluirse que las proyecciones de cada individuo en un mapa no serán puntos, sino familias de conjuntos anidados, cuya forma dependerá de la distorsión de la geometría espacial, propia de cada técnica, en la proyección plana. De la segunda, se deduce que no es razonable usar una distancia entre conjuntos borrosos para calcular una matriz numérica de distancias entre individuos, ni tampoco entre sus proyecciones. En general, mediante una de estas distancias no podrá establecerse un orden total entre las proyecciones que sea coherente con el existente entre los auténticos valores de las propiedades de los individuos, ya que tal orden es parcialmente desconocido.

En trabajos anteriores, se han hecho distintas simplificaciones para conseguir una proyección aproximada. Por ejemplo, en el método descrito en [2, 3] se extiende el escalado multidimensional (MDS) para permitir que las matrices de distancias contengan intervalos o números borrosos. La versión estándar de MDS consiste en encontrar la nube de puntos que minimiza una función de esfuerzo, definida por la diferencia cuadrática entre las matrices de distancias de los datos y de la nube. En la versión generalizada, se define una función de esfuerzo borroso-valorada, que mide el ajuste entre el conjunto de distancias compatibles con las figuras del mapa y el conjunto de distancias entre las descripciones borrosas de los individuos. En este método se supone que las proyecciones son círculos en el mapa, lo que no siempre es acertado, ya que se no se permite que los atributos tengan diferentes niveles de incertidumbre. Extensiones posteriores [10] eliminaron esta restricción, de la forma que se explica a continuación.

3.1. Determinación de la forma de una proyección

Sean $\widetilde{X}_i = (\widetilde{X}_{i1} \times \ldots \times \widetilde{X}_{if})$ y $\widetilde{X}_j = (\widetilde{X}_{j1} \times \ldots \times \widetilde{X}_{jf})$ dos tuplas de números borrosos, que modelan el conocimiento disponible acerca de los f atributos de los individuos i-ésimo y j-ésimo. A su vez, sean $[\widetilde{X}_i]_\alpha = [x_{i1}^-, x_{i1}^+] \times \ldots \times [x_{if}^-, x_{if}^+]$ y $[\widetilde{X}_j]_\alpha = [x_{j1}^-, x_{j1}^+] \times \ldots \times [x_{jf}^-, x_{jf}^+]$, dos cortes al mismo nivel α de $\widetilde{X}_i$ y $\widetilde{X}_j$.

Nuestro conocimiento acerca de las distancias entre $[\widetilde{X}_i]_\alpha$ y $[\widetilde{X}_j]_\alpha$ viene dado por el conjunto

$$d([\widetilde{X}_i]_\alpha, [\widetilde{X}_j]_\alpha) = \Big\{ \sqrt{\textstyle\sum_{k=1}^{f} (x_{ik} - x_{jk})^2} \mid x_{ik} \in [x_{ik}^-, x_{ik}^+], x_{jk} \in [x_{jk}^-, x_{jk}^+], 1 \le k \le f \Big\}. \quad (1)$$

De acuerdo con [10], se admitirá que la proyección de cada $[\widetilde{X}_i]_\alpha$ es un polígono (ver figura 1) en el que las

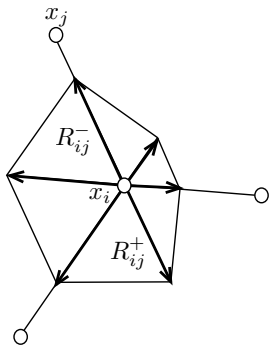


Figura 1: Los cortes de las proyecciones de cada individuo son polígonos, definidos por las distancias R_{ij} en las rectas que unen las proyecciones de sus centroides.

distancias entre centro y vértices, R_{ij}^+ y R_{ij}^-, dependen de las cotas de las distancias definidas en la ecuación 1.

Supóngase ahora que se desea realizar un mapa de N individuos $\{\widetilde{X}_1, \ldots, \widetilde{X}_N\}$. Sea $\overline{x}_i$ es el punto modal de $\widetilde{X}_i$, y sea $\{(z_{11}, \ldots, z_{1r}), \ldots, (z_{N1}, \ldots, z_{Nr})\}$ la proyección, en un mapa de dimensión r, de los N puntos modales. Los radios R_{ij}^+ y R_{ij}^- dependen de la distancia entre el conjunto $[\widetilde{X}_i]_\alpha$ y el valor $\overline{x}_j$ (ver figura 2 para una explicación gráfica) como sigue:

$$R_{ij}^+ = d_{ij}\left(\frac{\delta_{ij}^+}{\delta_{ij}} - 1\right) \quad R_{ij}^- = d_{ij}\left(\frac{\delta_{ij}}{\delta_{ij}^-} - 1\right) \quad (2)$$

donde $d_{ij} = \sqrt{\sum_{k=1}^{r}(z_{ik} - z_{jk})^2}$, $\delta_{ij} = \{d(\overline{x}_i, \overline{x}_j)\}$, $\delta_{ij}^+ = \text{máx}\{d(x, \overline{x}_j) \mid x \in [\widetilde{X}_i]_\alpha\}$, y $\delta_{ij}^- = \text{mín}\{d(x, \overline{x}_j) \mid x \in [\widetilde{X}_i]_\alpha\}$.

3.2. Función de esfuerzo

De acuerdo con lo anterior, el conocimiento disponible acerca del valor de la función de esfuerzo asociada a la proyección de los datos viene dado por la función borroso-valorada siguiente, definida por sus cortes:

$$S_\alpha = \Big\{ \sum_{i=1}^{N} \sum_{j=i+1}^{N} ||d(t,u) - \beta|| \mid t \in [\widetilde{X}_i]_\alpha, u \in [\widetilde{X}_j]_\alpha, \beta \in [d_{ij} - R_{ij}^- - R_{ji}^-, d_{ij} + R_{ij}^+ + R_{ji}^+] \Big\}. \quad (3)$$

Como alternativa a minimizar la función borroso-valorada anterior, es posible definir una medida que cuantifique cómo de diferentes son las ordenaciones de las matrices de distancias espaciales y las de su proyección. Sea $D(\widetilde{X}_i, \widetilde{X}_j)$ el conjunto borroso cuyos cortes son los conjuntos de distancias entre los individuos

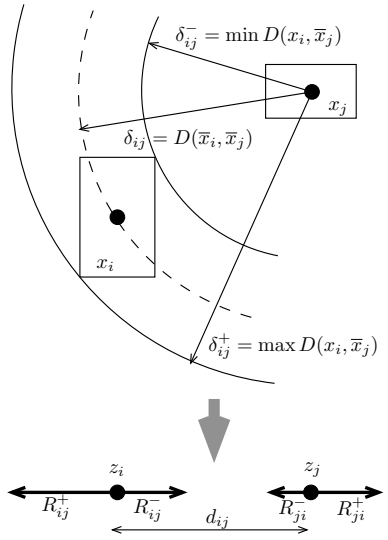


Figura 2: La distancia entre las proyecciones de $[\widetilde{X}_i]_\alpha$ y $[\widetilde{X}_j]_\alpha$ está entre $d_{ij} - R_{ij}^- - R_{ji}^-$ y $d_{ij} + R_{ij}^+ + R_{ji}^+$

$[\widetilde{X}_i]_\alpha$ y $[\widetilde{X}_j]_\alpha$,

$$[D(\widetilde{X}_i, \widetilde{X}_j)]_\alpha = d([\widetilde{X}_i]_\alpha, [\widetilde{X}_j]_\alpha) \quad (4)$$

y sea $[D'(\widetilde{Z}_i, \widetilde{Z}_j)]_\alpha$ el conjunto de distancias entre las proyecciones Z_i y Z_j de $[\widetilde{X}_i]_\alpha$ y $[\widetilde{X}_j]_\alpha$,

$$[D'(\widetilde{Z}_i, \widetilde{Z}_j)]_\alpha = [d_{ij} - R_{ij}^- - R_{ji}^-, d_{ij} + R_{ij}^+ + R_{ji}^+]. \quad (5)$$

Los valores $D(\widetilde{X}_i, \widetilde{X}_j)$ son los términos de la matriz de distancias entre los individuos, mientras que $D'(\widetilde{Z}_i, \widetilde{Z}_j)$ son los términos de la matriz de distancias entre sus proyecciones.

Si el mapa refleja correctamente las distancias entre los individuos de la población, debe cumplirse que el rango de la distancia que separa a los objetos i-ésimo y j-ésimo sea el mismo que el rango de la distancia entre sus proyecciones, en la matriz correspondiente. Por tanto, el número de parejas de objetos para los que esto no es cierto define una función de coste alternativa.

En el caso tratado en este estudio, el rango de una distancia borroso-valorada dentro de su matriz no está completamente definido, al haber parejas de distancias no comparables. No obstante, para cada nivel α puede definirse una relación de equivalencia entre los intervalos que definen los α-cortes de las distancias; dadas dos distancias $[d^-, d^+]$ y $[e^-, e^+]$, estas son no comparables cuando

$$[d^-, d^+] \parallel [e^-, e^+] \iff (d^+ \geq e^-) \wedge (e^+ \geq d^-). \quad (6)$$

En otro caso, podremos afirmar que $[d^-, d^+] \prec [e^-, e^+]$ o lo contrario, $[e^-, e^+] \prec [d^-, d^+]$. Por tanto, cada clase estará compuesta por intervalos no comparables entre sí, pero comparables con todos los intervalos de otras clases. El rango de una distancia vendrá definido, para el nivel α, por el número de distancias en otras clases que sean, con seguridad, mayores que aquélla. El objetivo del algoritmo numérico (que no se detallará en este trabajo, por razones de espacio) es el de obtener un mapa para el que los rangos de cada uno de los términos de las matrices de distancias entre individuos y entre proyecciones coincidan para todo valor de α. El valor de la función de esfuerzo es el máximo valor de α para el que dicha coincidencia no se produzca.

4. Análisis de casos prácticos

En esta sección se exponen los resultados obtenidos tras analizar cuestionarios de seguimiento de alumnos de máster de la Universidad de Oviedo. Este estudio tiene dos objetivos:

1. Identificar las diferencias globales entre los miembros de la misma asignatura en diferentes especialidades.

2. Evaluar los resultados del aprendizaje, viendo la evolución individual de los alumnos en un mismo mapa, de forma que pueda apreciarse en un único gráfico si se han conseguido reducir las diferencias entre grupos y la dirección de la formación.

4.1. Variación entre las capacidades de grupos de diferentes especialidades

El estudio se ha aplicado a alumnos de la asignatura "Informática" en primer curso de Ingeniería Técnica Industrial (especialidad en Electricidad), segundo curso de Ingeniería Técnica Industrial (especialidad en Química) y segundo curso de Ingeniería Industrial. Estos estudiantes reciben su clase por separado, aunque con programas similares. Las clases son impartidas por el mismo profesor. Estos datos se han adquirido al comienzo del segundo cuatrimestre, en Febrero de 2010. Cada especialidad se ha codificado por medio de un color diferente. El profesor ha evaluado, como en el caso anterior, los conocimientos previos de los alumnos en asignaturas que son prerrequisito para cursar la suya. Las coordenadas elegidas son "Cálculo", "Algebra", "Lenguajes de Programación" e "Internet". Cada estudiante se ha representado mediante una forma (el soporte del borroso correspondiente), y las posiciones de los individuos ficticios se han marcado con etiquetas. Estos son: "ONLY A" (todas las preguntas de la materia "A" bien contestadas, todas las restantes erróneas), "NO A" (todas las preguntas excepto las de "A" están bien contestadas, posición diametralmente opuesta), "NOTHING" (todas las respuestas erróneas) y "EVERYTHING" (todas las respuestas correctas).

La proximidad a este último punto es, aproximadamente, la nota media del test.

A la vista del gráfico (ver figura 3), los alumnos de la especialidad "Química" (azul) consideran que tienen una formación básica peor (por cercanía a "NOTHING") que los alumnos de la especialidad "Electricidad" (rojo) e Ingeniería Industrial (verde). Los alumnos de todos los grupos tienen una orientación neutra hacia las asignaturas de matemáticas, mientras que los que han elegido la especialidad en Electricidad se consideran mejor formados en lenguajes de programación y uso de internet que los que han elegido la especialidad en Química.

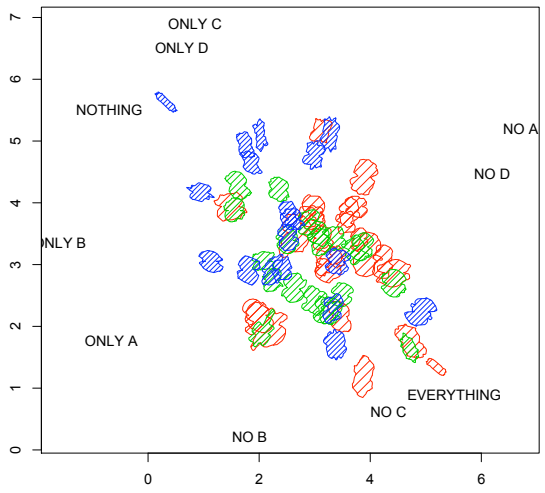


Figura 3: Variación en las capacidades individuales de alumnos de tres grupos

4.2. Evaluación del aprendizaje

En el segundo estudio se han analizado diez titulados en Informática, Física y Matemáticas, de nacionalidades española, estadounidense, italiana y polaca, que han cursado una asignatura de modelado inteligente en un máster de investigación de la Universidad de Oviedo. La formación de estos estudiantes es muy dispar. Se han seguido 36 materias, agrupadas globalmente en las coordenadas de evaluación "Algoritmos de Control", "Análisis Estadístico de Datos", "Algoritmos numéricos" y "Modelos Lineales".

En la parte superior de la figura 4 se observa que hay una gran dispersión en los conocimientos iniciales de los alumnos. Los alumnos de carreras técnicas, como los graduados en informática (parte inferior derecha de la figura), se han asignado a sí mismos las calificaciones más bajas, mientras que los alumnos de física y, con variaciones, de matemáticas, se consideran mejor preparados para dichas materias.

La misma encuesta, realizada al final del curso, arroja los resultados esperados. Todos los alumnos (figura 4) se han desplazado a la izquierda, a posiciones más cercanas del punto característico "EVERYTHING".

El desplazamiento ha sido más pronunciado para los alumnos del grupo de la derecha. Esto se percibe con mayor claridad en la parte inferior de la misma figura, donde se han retirado las figuras correspondientes a la parte final del curso y se han reemplazado por flechas que parten de la situación inicial y terminan en estas posiciones. La longitud de estas flechas está asociada al progreso del alumno durante el curso, y su dirección puede relacionarse (aunque quizás, de forma menos clara) con las materias en que el aprovechamiento ha sido superior. En este caso, el conocimiento se ha incrementado de forma relevante, para los alumnos de formación informática, en la materia "Análisis Estadístico de Datos"; los restantes alumnos han tenido una progresión menos diferenciada.

5. Conclusiones

En este trabajo se ha propuesto el uso de mapas exploratorios gráficos para analizar las competencias de grupos de individuos, cuando las características de los mismos tienen una incertidumbre derivada de las inconsistencias en la recogida de datos o en valores perdidos. El mapa de un grupo consta de varias figuras y de una lista de puntos característicos. La proximidad de un individuo a uno de estos puntos significa que este individuo tiene una formación cuyo balance se asemeja al valor representado por este indicador. Esta técnica puede ser usada para corroborar la mejora de las capacidades tras recibir un curso de formación: combinando en una misma gráfica los resultados de dos tests, separados en el tiempo, es posible determinar el desplazamiento de cada individuo hacia otros puntos característicos, y detectar por tanto a los individuos que mejor han aprovechado el curso.

Agradecimientos

Este estudio se ha financiado con cargo al proyecto de investigación TIN2011-24302, del Ministerio de Ciencia e Innovación

Referencias

[1] Couso, I., Sánchez, L., Higher order models for fuzzy random variables. Fuzzy Sets and Systems 159. 237-258. 2008.

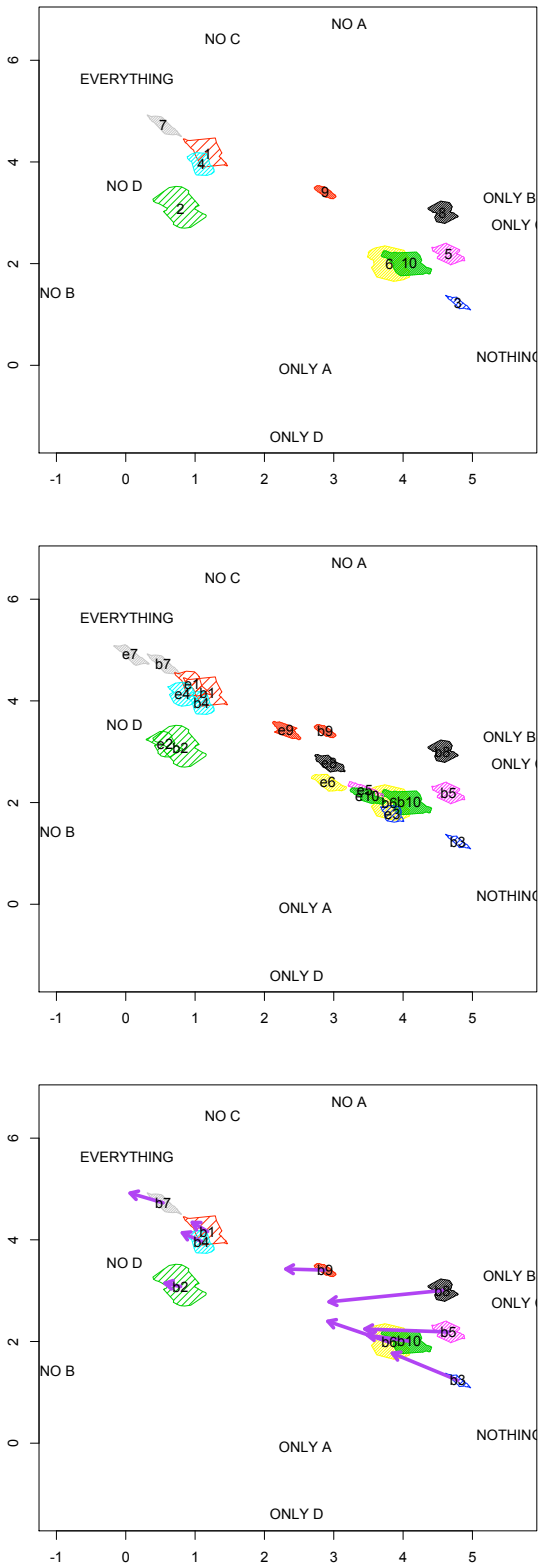


Figura 4: Variación en las capacidades individuales de alumnos de tres grupos

[2] Denoeux, T., Masson, M.-H., Multidimensional scaling of interval-valued dissimilarity data. Pattern Recognition Lett. 21, 83-92. 2000

[3] Hebert, P. A., Masson, M. H., Denoeux, T. Fuzzy multidimensional scaling. Computational Statistics and Data Analysis 51. 335-359. 2006.

[4] Honda, K., Ichihashi, H. Fuzzy local independent component analysis with external criteria and its application to knowledge discovery in databases. International Journal of Approximate Reasoning 42(3), 159-173. 2006.

[5] Kim, W., Choi, B., Hong, E-K., Kim S-K. A Taxonomy of Dirty Data. Data Mining and Knowledge Discovery 7, 81-99, 2003.

[6] Kruskal, J.B., Nonmetric multidimensional scaling: a numerical method. Psychometrika 29, 115-129. 1964.

[7] Nuhfer, E. Keep The Knowledge Survey: A Tool for All Reasons Edward Nuhfer, University of Colorado at Denver and Delores Knipp, United States Air Force Academy: To Improve the Academy, 21, 59-78. 2003.

[8] Mazza, R., Milani, C. Exploring usage analysis in learning systems: Gaining insights from visualisations. In Workshop on usage analysis in learning systems at 12th International Conference on Artificial Intelligence in Education, New York, USA 1-6. 2005.

[9] Palacios, A., Sánchez, L., Couso, I. Diagnosis of dyslexia with low quality data with genetic fuzzy systems. International Journal on Approximate Reasoning 51, 993-1009. 2010.

[10] Sánchez, L., Couso, I., Otero, J., Palacios, A. Assessing the evolution of learning capabilities and disorders with a graphical exploratory analysis of surveys containing missing and conflicting answers. Neural Network World 20:7, 825-838, 2010

KEEL: UNA HERRAMIENTA DOCENTE PARA SISTEMAS DIFUSOS

Joaquín Derrac[1], **Julián Luengo**[2], **Alberto Fernández**[3], **Salvador García**[3], **Jesús Alcalá-Fdez**[1]

[1]*Departamento de Ciencias de la Computación e Inteligencia Artificial,CITIC-UGR, Universidad de Granada, 18071 Granada, Spain*, {jderrac,jalcala}@decsai.ugr.es

[2]*Departamento de Ingeniería Civil, EPS, Universidad de Burgos, 09006, Burgos, España*, jluengo@ubu.es

[3]*Departamento de Ciencias de la Computación, Universidad de Jaén, 23071 Jaén, España*, {alberto.fernandez,sglopez}@ujaen.es

Resumen

En la actualidad, los sistemas difusos están cobrando cada vez más relevancia en los planes de estudio de asignaturas de minería de datos y aprendizaje automático. En este contexto, las herramientas software de código abierto ofrecen a los estudiantes la posibilidad de experimentar con las técnicas abordadas, reforzando así su aprendizaje. Entre ellas, destacamos KEEL como una herramienta versátil para evaluar y analizar el funcionamiento de los algoritmos estudiados.

El objetivo de este trabajo es presentar la faceta educativa de la herramienta software KEEL: Un módulo educativo de utilidad para reforzar la docencia en asignaturas de introducción al aprendizaje automático y a los sistemas difusos. Este módulo proporciona al usuario la capacidad de visualizar el progreso de los algoritmos, ayudándole a evaluar, ajustar y comprender el funcionamiento de las técnicas clásicas del área.

Palabras Clave: Minería de Datos, Aprendizaje Automático, Sistemas Difusos, Java, Herramienta Software, Software Educativo.

1 Introducción

En los últimos años es habitual la inclusión de asignaturas de Minería de Datos y Aprendizaje Automático como parte de la formación de los estudiantes de Ciencias de la Computación. Este tipo de materias, cruciales para su formación científica [11], les ayudan a obtener una serie de conceptos claves en su formación como profesionales, aportándoles importantes herramientas para el futuro.

Pese a ello, el aprendizaje de las técnicas derivadas de estos campos no siempre es sencillo. Este aprendizaje suele requerir habilidades de análisis que permitan al estudiante entender las características más importantes de cada técnica. En el Aprendizaje Automático, el problema subyacente consiste en que la implementación de la mayoría de los métodos relacionados requiere de ciertos conocimientos de programación, así como de una cantidad considerable de tiempo y esfuerzo. Así, se corre el riesgo de acabar empleando las horas dedicadas a la enseñanza de esta materia como meras *clases de programación*, en lugar de centrarse en el análisis de las características de los algoritmos, desde un punto de vista teórico y práctico.

Una manera de aliviar este problema consiste en emplear herramientas software que incluyan dichas implementaciones. Actualmente es posible encontrar en la Web un número elevado de ellas. Sin embargo, la mayoría están destinadas a abordar una determinada aplicación - siendo difíciles de aprovechar desde un punto de vista docente.

Sin embargo, en los últimos años han aparecido cierto número de herramientas de propósito general. Aunque muchas son comerciales, unas pocas (como Weka [5] o JavaML [1]) son distribuidas de forma libre, alcanzando gran popularidad dentro de la comunidad investigadora gracias a su condición de *software libre* [12] [1]).

KEEL [2, 3] es una herramienta software libre desarrollada completamente en Java. KEEL permite al usuario emplear una gran cantidad de técnicas de Aprendizaje Automático en diferentes tipos de problemas: Regresión, clasificación, agrupamiento, asociación, etc., incluyendo una gran recopilación de los Sistemas Difusos existentes. Además de como herramienta para investigación, KEEL ha sido diseñado también con características educativas.

En este trabajo se describen los principales aspectos de KEEL como herramienta docente. Su módulo educativo es presentado como una herramienta capaz de ofrecer una visión del funcionamiento de los algoritmos, permitiendo al estudiante comprender y adaptar su funcionamiento. Así,

[1](Más información en los siguientes enlaces: KDnuggets software http://www.kdnuggets.com/software The-Data-Mine site http://the-data-mine.com/bin/view/Software

esta herramienta puede usarse como apoyo práctico en todo tipo de asignaturas de Minería de Datos y Aprendizaje Automático, incluyendo aquellas con una fuerte componente de técnicas basadas en lógica difusa.

Es importante destacar que esta herramienta ya ha sido utilizada con éxito en varios programas de formación, tanto nacionales como internacionales. Por ejemplo, recientemente se empleó como apoyo al proyecto MIBISOC (Medical Imaging using Bio-Inspired and Soft Computing) [2].

El resto del trabajo se organiza como sigue: la Sección 2 describe brevemente las características generales de KEEL. La Sección 3 da una visión del módulo educativo, destacando sus capacidades más sobresalientes. La Sección 4 presenta un caso de uso considerando el análisis de varios Sistemas Difusos clásicos. Finalmente, en la Sección 5 se presentan las conclusiones del trabajo.

2 Descripción de KEEL

KEEL es una herramienta software para la preparación de algoritmos de minería de datos. La versión actual de KEEL está compuesta por los siguientes módulos (ver Figura 1):

Figura 1: Pantalla principal de KEEL

- *Tratamiento de datos* (Data Management): Este módulo contiene una serie de herramientas de tratamiento de datos: Importación, exportación, edición y visualización de datos, aplicación de transformaciones, etc.
- *Experimentos* (Experiments): Este módulo está dedicado al diseño de experimentos, proporcionando numerosas opciones: Tipo de validación, tipo de aprendizaje (clasificación, regresión, aprendizaje no-supervisado), etc.
- *Educativo* (Educational): Este módulo permite realizar experimentos interactivos. Con una estructura similar al módulo anterior, permite diseñar experimentos con propósitos educativos.
- *Módulos* (Modules): Adicionalmente, KEEL incluye un módulo para datos no-balanceados [4], un módulo de análisis estadístico no paramétrico [7, 9], y un módulo de aprendizaje multi-instancia [8].

Esta estructura hace que KEEL sea una herramienta interesante para distintos tipos de usuarios. Las principales características de KEEL son las siguientes:

- Posee una librería estadística para análisis de algoritmos. Los tests de esta librería permiten analizar la bondad de los resultados obtenidos, realizando comparaciones paramétricas y no paramétricas.
- Incluye algoritmos de aprendizaje de modelos predictivos, de preprocesamiento (discretización, selección de instancias, selección de características, etc.) y post-procesamiento. También incluye muchas propuestas del estado del arte de diferentes áreas de la Minería de datos, como por ejemplo árboles de decisión, sistemas difusos basados en reglas, etc.
- Ofrece al usuario una interfaz amigable, orientada al análisis de algoritmos.
- Permite crear experimentaciones conteniendo múltiples conjuntos de datos y algoritmos conectados entre sí. Los experimentos son generados mediante scripts independientes de la interfaz de usuario, para permitir una ejecución separada en la misma u otras máquinas.

Para más información sobre las características de la herramienta KEEL, pueden consultarse los trabajos [2, 3] o el sitio web del proyecto KEEL (`http://www.keel.es`).

3 Módulo Educativo

En el ámbito docente, las necesidades de un estudiante son muy diferentes de las de un investigador. Generalmente, sus necesidades no consisten en realizar experimentaciones con un gran número de problemas y métodos del estado del arte. Generalmente, los experimentos a realizar durante el aprendizaje serán más sencillos, orientados a comprobar la evolución y los resultados de los algoritmos más comunes de forma clara.

Siguiendo esta idea, el módulo educativo de KEEL ha sido diseñado para dar cabida a las propuestas más representativas de cada área. Así, ofrece las técnicas de preprocesamiento y aprendizaje más populares, junto con una amplia colección de problemas representativos en los ámbitos de clasificación y regresión.

[2] MIBISOC es una Red de Entrenamiento Internacional del programa Marie Curie International Training Network soportada por la Comisión Europea dentro del Séptimo Programa Marco (FP7 PEOPLE-ITN- 2008). http://www.softcomputing.es/mibisoc/

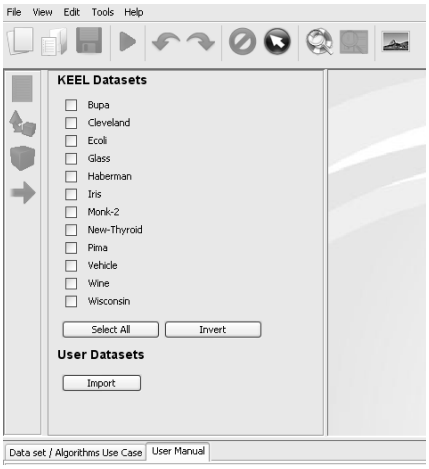


Figura 2: Selección de conjuntos de datos. Es posible añadir nuevos conjuntos, mediante el botón de importar.

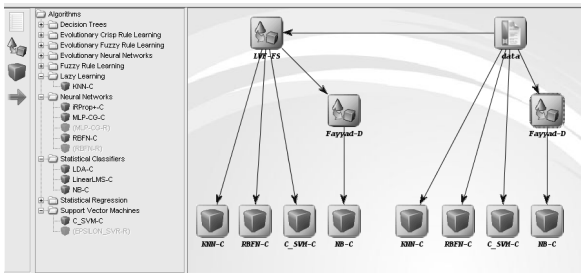


Figura 3: Un experimento para analizar el rendimiento del Filtro Las Vegas de selección de características sobre 4 clasificadores: K-NN, Naive Bayes, una red neuronal de base radial y una máquina de vectores soporte. Además se ha empleado el discretizador de Fayyad para adecuar los datos en el caso de Naive Bayes.

A la hora de crear un nuevo experimento con el módulo educativo, el primer paso consiste en seleccionar el tipo de problema (clasificación o regresión) y el esquema de validación a emplear (validación cruzada de k- particiones, validación cruzada 5x2 o sin validación). Una vez determinados ambos puntos, deben indicarse qué conjuntos de datos se desean emplear. La Figura 2 muestra un ejemplo en el contexto de clasificación, donde es importante destacar que el usuario puede incluir nuevos conjuntos mediante el botón de importar (*Import*).

La interfaz principal permite el diseño de experimentos de forma gráfica. La Figura 3 muestra un experimento diseñado para evaluar el comportamiento de varios clasificadores bajo conjuntos de datos preprocesados con un método de selección de características. Arrastrando y colocando los iconos que representan a cada técnica, los estudiantes pueden diseñar fácilmente este tipo de experimentos sin necesidad de emplear complicados procedimientos para establecer los parámetros, algoritmos y conjuntos de datos a utilizar típicos de un diseño experimental mayor.

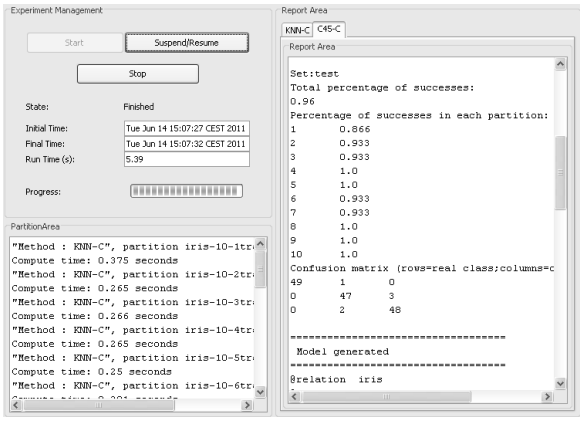


Figura 4: Ventana de gestión de experimentos. En ella se muestran los resultados obtenidos, los tiempos de ejecución y los modelos generados.

Esta facilidad de uso es especialmente importante si consideramos que el módulo educativo contiene una amplia colección de métodos pertenecientes a áreas muy diferentes entre sí. Por ello, disponer de un modelo sencillo para crear y configurar los experimentos que los incluyan es crucial, puesto que no se puede exigir a los estudiantes en estas etapas de su formación que adquieran un conocimiento profundo de cada una de las técnicas para poder emplearlas.

En el módulo educativo, el experimento se desarrolla en tiempo real. Una vez creado, el usuario puede escoger entre ejecutarlo o guardarlo en formato XML. Si se ejecuta, el sistema mostrará una ventana auxiliar para gestionar y visualizar el desarrollo del experimento (Figura 4). Una vez finalizado, esta ventana mostrará los resultados obtenidos por cada algoritmo, incluyendo información adicional como, por ejemplo, las matrices de confusión en clasificación o los errores medios obtenidos en regresión.

De esta manera, el usuario dispone de una visión del progreso de los algoritmos, pudiendo acceder a ellos directamente desde la propia interfaz. En el área de resultados se muestran los modelos generados, de gran utilidad como panorámica del conocimiento extraído a partir de los datos y su estructura intrínseca (por ejemplo, los algoritmos de árboles de decisión mostrarán una representación de los árboles construidos, mientras que para los sistemas basados en reglas se obtendrá la base de reglas generada).

Gracias a estas características, el módulo educativo es ideal para estudiantes que deseen analizar y mejorar los resultados obtenidos por un algoritmo dado. Los estudiantes pueden usarlo para comparar varios métodos relacionados sobre un conjunto fijo de problemas, e incluso para tratar de descubrir qué algoritmos o metodologías son más eficaces de cara a un problema concreto. Mediante la información obtenida en los informes de resultados, pueden tratar de determinar las causas del buen o mal comportamiento de un método para, más adelante, tratar de mejorarlo (adaptando, por ejemplo, la configuración de parámetros empleada).

Tabla 1: Conjuntos empleados en el estudio

Conjunto	# Ej.	# At.	# Cl.	Conjunto	# Ej.	# At.	# Cl.
Bupa	345	6	2	Monk-2	432	6	2
Ecoli	336	7	8	New-Thyroid	215	5	3
Glass	214	9	7	Pima	768	8	2
Haberman	306	3	2	Vehicle	846	18	4
Iris	150	4	3	Wine	178	13	3

4 Caso de estudio

Para mostrar las características del módulo educativo, se va a realizar un pequeño experimento involucrando dos algoritmos basados en sistemas difusos. El objetivo del estudio será destacar qué técnica ofrece mejores resultados en clasificación sobre un determinado conjunto de problemas de clasificación supervisada.

La Tabla 1 describe los 10 conjuntos seleccionados y sus principales características: **# Ej.** indica el número de ejemplos (instancias) del conjunto, **# At.** indica el número de atributos y **# Cl.** indica el número de clases. Estos conjuntos han sido tomados del repositorio KEEL-Dataset repository [3] [3], estando directamente disponibles en la instalación estándar del módulo educativo. Como esquema de validación, se ha seleccionado una validación cruzada en 10 partes (en general, el uso de este esquema es preferible a emplear una simple partición de entrenamiento y test, debido a que los resultados obtenidos suelen ser menos sensibles al sobreaprendizaje, y además permite que toda instancia disponible pertenezca al conjunto de test una vez).

Como algoritmos de Sistemas Difusos a estudiar, se ha seleccionado el algoritmo de Chi *et al.* [6], en sus versiones con 3 y 5 etiquetas por variable (**Chi-3** y **Chi-5**). Ambas versiones de este Sistema Difuso clásico serán comparadas con **SLAVE** [10], un competente Sistema Difuso Evolutivo. Puede encontrarse más información sobre ambos métodos tanto en la ayuda del propio módulo educativo como en el sitio web del proyecto KEEL, en la sección de algoritmos incluidos (*Included Algorithms*) [4]).

En KEEL, cada método ofrece una configuración de parámetros por defecto (configurada siguiendo las recomendaciones de sus autores). En este estudio, se emplearán dichos parámetros, con excepción del número de etiquetas para las 2 versiones del algoritmo de Chi *et al.*.

Una vez seleccionados los conjuntos de datos, algoritmos, parámetros y esquema de validación a emplear, el siguiente paso consiste en dibujar el grafo que representa al experimento (ver Figura 5). Una vez creado, el experimento puede comenzar. La interfaz de control (Figura 6) permitirá en este momento iniciar el experimento, así como pausarlo o terminarlo en caso de ser necesario.

Una vez finalizado el experimento, la interfaz mostrará los resultados para cada método y partición. Para cada uno, se muestra un informe completo incluyendo los resultados en acierto y las matrices de confusión obtenidas, tanto en entrenamiento como en test (ver Figura 7).

[3] `http://www.keel.es/datasets.php`
[4] `http://www.keel.es/algorithms.php`

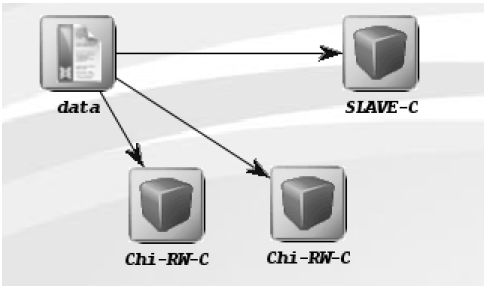


Figura 5: Grafo del experimento preparado para este caso de estudio

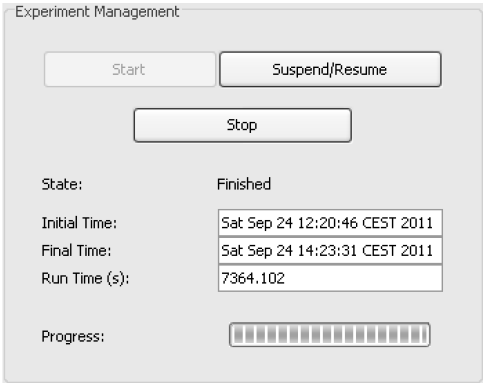


Figura 6: Interfaz de control para gestionar los experimentos del módulo educativo

```
Set:test
Total percentage of successes:
0.96
Percentage of successes in each partition:
1       0.933
2       0.933
3       1.0
4       1.0
5       1.0
6       0.933
7       0.933
8       0.933
9       1.0
10      0.933
Confusion matrix (rows=real class;columns=obtained class):
49      1       0
0       45      5
0       0       50
```

Figura 7: Resultados obtenidos por SLAVE en el problema Iris. El informe muestra el acierto obtenido en cada partición, el acierto medio y la matriz de confusión generada.

Chi-3

```
@Number of rules: 14

1: sepalLength IS L_0 AND sepalWidth IS L_1 AND petalLength IS L_0 AND petalWidth IS L_0: Iris-setosa with Rule Weight: 0.9946914141525106
2: sepalLength IS L_1 AND sepalWidth IS L_2 AND petalLength IS L_0 AND petalWidth IS L_0: Iris-setosa with Rule Weight: 1.0
3: sepalLength IS L_1 AND sepalWidth IS L_1 AND petalLength IS L_0 AND petalWidth IS L_0: Iris-setosa with Rule Weight: 0.9897080145841041
4: sepalLength IS L_0 AND sepalWidth IS L_0 AND petalLength IS L_0 AND petalWidth IS L_0: Iris-setosa with Rule Weight: 0.8743721633888051
5: sepalLength IS L_1 AND sepalWidth IS L_1 AND petalLength IS L_1 AND petalWidth IS L_1: Iris-versicolor with Rule Weight: 0.5474870380145459
6: sepalLength IS L_1 AND sepalWidth IS L_0 AND petalLength IS L_1 AND petalWidth IS L_1: Iris-versicolor with Rule Weight: 0.672414074237995
7: sepalLength IS L_0 AND sepalWidth IS L_0 AND petalLength IS L_1 AND petalWidth IS L_1: Iris-versicolor with Rule Weight: 0.7985932110408883
8: sepalLength IS L_1 AND sepalWidth IS L_1 AND petalLength IS L_2 AND petalWidth IS L_2: Iris-virginica with Rule Weight: 0.872291803923327
9: sepalLength IS L_2 AND sepalWidth IS L_1 AND petalLength IS L_2 AND petalWidth IS L_2: Iris-virginica with Rule Weight: 0.9575702462535564
10: sepalLength IS L_1 AND sepalWidth IS L_1 AND petalLength IS L_2 AND petalWidth IS L_1: Iris-virginica with Rule Weight: 0.18040318939382682
11: sepalLength IS L_2 AND sepalWidth IS L_1 AND petalLength IS L_2 AND petalWidth IS L_1: Iris-virginica with Rule Weight: 0.6509993328176151
12: sepalLength IS L_1 AND sepalWidth IS L_0 AND petalLength IS L_2 AND petalWidth IS L_1: Iris-virginica with Rule Weight: 0.1843982387098268
13: sepalLength IS L_1 AND sepalWidth IS L_1 AND petalLength IS L_1 AND petalWidth IS L_2: Iris-virginica with Rule Weight: 0.6473680231744567
14: sepalLength IS L_1 AND sepalWidth IS L_0 AND petalLength IS L_1 AND petalWidth IS L_2: Iris-virginica with Rule Weight: 0.6585359964748987
```

SLAVE

```
@Number of rules: 3

IF  petalWidth = ( L0 ) THEN Class IS Iris-setosa   W 1.0

IF  petalLength = ( L0 L1 L2 ) petalWidth = ( L0 L1 L2 ) THEN Class IS Iris-versicolor   W 0.4789917154426708

IF  petalWidth = ( L0 L1 L3 L4 ) THEN Class IS Iris-virginica   W 0.4123616236162361
```

Figura 8: Bases de reglas generadas por **Chi-3** (arriba) y **SLAVE** (abajo) en Iris. En ambos casos puede comprobarse la influencia de la longitud y anchura del pétalo para clasificar, si bien la base de reglas de SLAVE es mucho más compacta.

Tabla 2: Resultados obtenidos (Acierto en test)

Conjunto	**Chi-3**	**Chi-5**	**SLAVE**
Bupa	0.5790	**0.5940**	0.5850
Ecoli	0.7200	0.8180	**0.8510**
Glass	**0.5980**	0.5930	0.5790
Haberman	**0.7320**	0.7250	0.7120
Iris	0.9260	0.9530	**0.9600**
Monk-2	0.4280	0.4720	**0.9740**
New-thyroid	0.8410	**0.9110**	**0.9110**
Pima	**0.7300**	0.7170	0.7260
Vehicle	0.6070	**0.6400**	0.6180
Wine	**0.9380**	0.7520	0.9320
Average	0.7099	0.7175	**0.7848**

La Tabla 2 muestra los resultados medios obtenidos en acierto (expresados en el intervalo $[0,1]$). Los mejores resultados están destacados en **negrita**. Como puede verse, no existe un método que domine claramente la comparación en todos los casos, si bien SLAVE obtiene un acierto medio mayor. Este es un resultado interesante, ya que sugiere que los problemas seleccionados para el experimento poseen diferentes cualidades que los hacen más sencillos o complejos para cada algoritmo. Así, puede ser interesante tratar de analizar los motivos de estas diferencias.

Otro aspecto interesante consiste en analizar los modelos generados. El módulo educativo ofrece, junto a los resultados obtenidos, un listado de los modelos generados durante el entrenamiento. Esto es importante de cara a determinar cómo está funcionando un determinado algoritmo y por qué motivo está ofreciendo resultados más o menos comunes.

En el caso de los Sistemas Difusos basados en reglas, es posible obtener la base de conocimiento generada durante el entrenamiento. De esta manera, los estudiantes pueden tratar de interpretar el conocimiento extraído por el algoritmo, valorando los modelos generados y comparándolos entre sí. Esta tarea, íntimamente ligada a como se analizarían los algoritmos en un estudio más complejo, puede ser realizada fácilmente gracias al módulo educativo.

Por ejemplo, la Figura 8 muestra las bases de reglas generadas por **Chi-3** y **SLAVE** para el problema Iris. De éste problema se conoce que los atributos relativos a la longitud y anchura del pétalo son los más determinantes a la hora de establecer una clasificación. En efecto, si estudiamos la base de reglas obtenida por **Chi-3** podemos ver cómo los ejemplos de la clase *Iris-setosa* quedan caracterizados por aquellas reglas que asignan la etiqueta L_0 a los dos atributos del pétalo (*petalLength y petalWidth*). Los ejemplos de la clase *Iris-versicolor* se caracterizan por tener asignada la etiqueta L_1 a ambos atributos, mientras que los ejemplos de la clase *Iris-virginica* deben de tener asignada la etiqueta L_2 en al menos uno de estos 2 atributos clave.

La base de reglas de **SLAVE** también usa los atributos del pétalo, pero de forma más simple: La clase *Iris-setosa* queda caracterizada por los ejemplos cuya anchura de pétalo tenga asignada la etiqueta L_0. Si ambas medidas de pétalo tienen asignadas las etiquetas L_0, L_1 o L_2, la clase de la instancia será *Iris-versicolor* (con menor prioridad que los ejemplos de *Iris-setosa*). Finalmente, aquellos ejemplos cuya anchura de pétalo no sea L_2 y no hayan sido clasificados por las reglas anteriores, serán considerados como instancias de la clase *Iris-virginica*.

Ambos ejemplos contrastan en este caso, sobre todo si se tiene en cuenta que **SLAVE** ofrece un mayor acierto en test. En este caso, un alumno podría concluir que el motivo de la mejor precisión ofrecida por **SLAVE** se debe a la mayor generalidad de sus reglas (de 1 o 2 condiciones), lo cual permite clasificar de forma más consistente nuevos ejemplos no considerados durante el entrenamiento.

En resumen, este caso de estudio ha mostrado las principales características del módulo educativo de KEEL como apoyo a la docencia de sistemas difusos. El módulo permite a los estudiantes poder emplear y analizar métodos de diferentes áreas sin necesidad de implementarlos. Así, pueden dedicar su esfuerzo a analizar los resultados obtenidos y así tratar de encontrar los motivos de las diferencias entre métodos y/o la forma de optimizar su comportamiento.

5 Conclusiones

En este trabajo hemos presentado KEEL como herramienta de apoyo a la docencia de Sistemas Difusos. Mediante su módulo educativo, es posible acceder de forma sencilla a varios problemas y técnicas del estado del arte, permitiendo así su uso por alumnos y profesores en asignaturas y cursos introductorios al área.

Se ha desarrollado un caso de uso con el objetivo de demostrar las capacidades del módulo. Este caso de uso ha consistido en una comparación entre varios Sistemas Difusos sobre 10 problemas de clasificación. Tras su realización, se han analizado los resultados obtenidos a través de la herramienta, así como los modelos obtenidos. Esto ha permitido ilustrar cómo puede obtenerse información relevante acerca del comportamiento de un método mediante las herramientas disponibles en el módulo.

Gracias a ello, han podido extraerse varias conclusiones de forma sencilla, sin necesidad de implementar ninguno de los Sistemas Difusos considerados ni de realizar ningún esfuerzo de programación adicional para recuperar y analizar los resultados. De esta manera, pueden realizarse experimentos similares en la docencia de éstas u otras técnicas incluidas en el módulo, de forma asequible para los profesores o alumnos que lo necesiten.

Agradecimientos

Este trabajo ha sido soportado por los proyectos nacionales TIN2008-06681-C06-01 y TIN2011-28488, y por el proyecto andaluz TIC-2010-6858. J. Derrac posee una beca FPU del Ministerio de Educación.

Referencias

[1] T. Abeel, Y. V. de Peer, Y. Saeys: Java-ML: A machine learning library. *Journal of Machine Learning Research* 10, pp. 931–934, 2009.

[2] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, F. Herrera: KEEL: A software tool to assess evolutionary algorithms to data mining problems. *Soft Computing* 13:3, pp. 307–318, 2009.

[3] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera: Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing* 17:2–3, pp. 255–287, 2011.

[4] G. E. A. P. A. Batista, R. C. Prati, M. C. Monard: A study of the behaviour of several methods for balancing machine learning training data. *SIGKDD Explorations* 6:1, pp. 20–29, 2004.

[5] R. R. Bouckaert, E. Frank, M. A. Hall, G. Holmes, B. Pfahringer, P. Reutemann, I. H.Witten: Weka - experiences with a java open-source project. *Journal of Machine Learning Research* 11, pp. 2533–2541, 2010.

[6] Z. Chi, H. Yan, T. Pham: *Fuzzy algorithms with applications to image processing and pattern recognition* World Scientific, Nueva York, 1996.

[7] J. Derrac, S. García, D. Molina, F. Herrera: A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation* 1:1, pp. 3–18, 2011.

[8] T. Dietterich, R. Lathrop, T. Lozano-Pérez: Solving the multiple instance problem with axis-parallel rectangles. *Artifical Intelligence* 89:1–2, pp. 31–71, 1997.

[9] S. García, F. Herrera: An extension on S̈tatistical comparisons of classifiers over multiple data setsS̈for all pairwise comparisons. *Journal of Machine Learning Research* 9, pp. 2579–2596, 2008.

[10] A. González, R. Pérez: Selection of relevant features in a fuzzy genetic learning algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 31:3, pp. 417–425, 2001.

[11] E. Mjolsness, D. DeCoste: Machine learning for science: State of the art and future prospects. *Science* 293, pp. 2051–2055, 2001.

[12] S. Sonnenburg, M. Braun, C. Ong, S. Bengio, L. Bottou, G. Holmes, Y. LeCun, K.-R. Müller, F. Pereira, C. Rasmussen, G. Rätsch, B. Schölkopf, A. Smola, P. Vincent, J. Weston, R. Williamson: The need for open source software in machine learning. *Journal of Machine Learning Research* 8, pp. 2443–2466, 2007.

ON FUZZY SETS AND THEIR LINGUISTIC LABELLING

– A REFLEXION FOR THOSE TEACHING FUZZY SET THEORY –

Enric Trillas

European Centre for Soft Computing, Mieres, 33600, Spain,
enric.trillas@softcomputing.es

Abstract

This paper tentatively deals with the basic problem of knowing when a function in $[0,1]^X$ can represent a fuzzy set labelled *P*. It is neither for students, nor a divulgative one, but just for the professors own reflection.

1 INTRODUCTION

Fuzzy sets were introduced just to represent 'classes without sharp boundaries', and they usually appear in many applications with reference to the corresponding linguistic labels naming them. Nevertheless, in the textbooks on fuzzy sets and fuzzy logic they are just presented as being either functions from the universe of discourse into the unit interval, or sets of pairs constitued by elements in that universe and an associated number between 0 and 1. In this way the linguistic genesis of fuzzy sets seems to be lost when it is actually important not only for its current applications, but mainly for the evolution of fuzzy logic towards Computing with Words or Computing with Natural Language.

This short paper only has the pedagogical aim of going back towards the linguistic roots of fuzzy sets, with the goal of helping those who teach on the theoretical aspects of fuzzy logic to make a presentation more closely linked than usual to natural language. At this end it should be emphasized that it is not an immediate and clear fact that linguistic labels P, or names given to a property p (the predicate P), are measurable in the totally ordered scale given by the unit interval of the real line, a case in which the degrees up to which x and y are P are always comparable although this were not the case between the statements 'x is P' and 'y is P'.

It should be pointed out that the meaning of a predicate or linguistic label depends not only on some background knowledge on its use, but mainly on which universe of discourse it is applied to, and for which purpose it is done. For instance, a very well known predicate like it is 'short', could be non relevant in a universe of discourse constitued by people of the same height, like a basketball team, but full of meaning when these people do have different heights as it happens when the universe of discourse is, for instance, the European Union. Consequently, in this paper all predicates P will be considered jointly with the universe of discourse X where they are applied to.

Notice that the *meaning* of P in X comes from the management in language of the elemental statements 'x is P' for all x in X. Even if X is known, P alone cannot have meaning, and the set of statements {x is P; x in X} must be considered. Without considering these statements, it is impossible to consider the concept of a degree for P on X, since such 'degree' does refer to the totality of the degrees up to which is 'x is P'.

It should be also noticed that, to avoid theoretic complexities and as it is done in fuzzy logic, it will be supposed that the universe X is just a classical set, a supposition that, notwithstanding, covers a very large number if not the totality of the applications, as well as those currently foreseeable.

2 ON PREDICATES AND THEIR COLLECTIVES

If a predicate or linguistic label P is applicable to a universe of discourse X, and if *P* can be considered a *collective noun*, then it does exist a *collective*, or community, in some way constitued by the elements of X verifying the property p to some extent. The serious problem with this *abstract entity* lies precisely in describing in which way it is 'constituted'. For instance, if X = [0,10] and P = small, it is well known (Sorites Paradox) that it does not exist the subset {x in X; x is small}, since for the predicate 'small' the axiom of specificity of Set Theory [3] fails in non accepting a definition of 'small' of the type 'x is small *if and only if* such and such' [1, 8].

In a layman's speech in natural language, 'small' is used in the interval $[0,10]$ by following some rules not allowing

to obtain the former set. But nobody will doubt that there are numbers that can be qualified as small, and numbers that are not at all small, as for instance 0 and 10, respectively. The *collective* of small numbers in the interval [0,10] is 'constituted' by those that can be linguistically qualified as small, and clearly neither the set [0,10] is such a collective, nor it is a set. The collective of the small numbers is, nevertheless, 'something' inside the interval [0,10]. It is a kind of 'cloud' in [0,10].

Of course, for some collectives the axiom of specificity works perfectly well and they are just sets. This is the case, for instance, of P = *less than three*, for which the collective is reduced to the subset [0,3). It should be pointed out that all the numbers in [0,3) are equally less than three, that is, they are less than three up to the same extent, and this marks the main difference with the collectives that are not sets where the extent up to which they are P is not the same for all of its 'elements', whatever this term could mean.

In the case in which a collective is a subset A, the universe of discourse can be immediatelly partitioned by the pair of subsets A and its complement $A' : X = A \bigcup A'$, and $A \bigcap A' = \emptyset$. These properties depend, in the case of collectives that are not sets, on the meaning of the connectives *and* ($\bigcap$), *or* ($\bigcup$), and *not* ($'$). Such collectives are generated by the predicates that are gradable, or imprecise. Instead those that are sets are generated by crisp, rigid or bivaluate predicates.

It should be pointed out that also 'sets' are abstract entities. Think, for instance, in the subset of real numbers whose elements are the transcendent ones, for which a definition like 'x is transcendent if and only if such and such' exists and makes the collective a subset of the real line.

A subset A of X can also be defined by its characteristic or membership function,

$$\mu_A(x) : X \to [0,1],$$

given by

$$\mu_A(x) = 1, \text{ if } x \text{ is in } A, \text{ and } \mu_A(x) = 0 \text{ if } x \text{ is not in } A,$$

and, reciprocally, if f is a function *in* $\{0,1\}^X$, then it defines the unique subset $A = f^{-1}(1)$ whose characteristic function is exactly f. In this sense there is no difference between the subsets of X and the functions from X to the two-elements set $\{0,1\}$. The power set $P(X)$ can be identified with the set of functions $\{0,1\}^X$. Obviously, collectives that are not subsets of X cannot be identified with such functions. If the collective originated by a predicate *P cannot*, in most of the cases, *be defined*, the problem lies notwithstanding in *representing* it.

3 ON THE DEGREE OF A PREDICATE

3.1

Given a universe of discourse X and a predicate P on it, to capture *how P works in X* it is needed, at least, to know the *empirical* relation '*x is less P than y*' that can be translated into X by the algebraic relation $\leq_P \subset X \times X$, defined by:

$$x \leq_P y \ \textit{iff} \ x \text{ is less } P \text{ than } y,$$

allowing to consider:

1. The inverse relation $\leq_P^{-1}$ (*more P than*), by

$$x \leq_P^{-1} y \ \textit{iff} \ y \leq_P x \ \textit{iff} \ y \text{ is less } P \text{ than } y$$

2. The relation (*equally P as*)

$$x \equiv_p y \ \textit{iff} \ x \leq_P y \& y \leq_P x,$$

that can be interpreted as x and y are equally P

Of course, $\equiv_P$ is an equivalence relation provided $\leq_P$ is reflexive and transitive, that is, when $\leq_P$ is a preorder. In this case, it exists the quotient set $X/\equiv_P$ whose elements are the equivalence classes

$$[x] = \{y \in X; x \equiv_P y\}.$$

If $\leq_P$ is not a preorder, there also exist the relation $\equiv_P$ and the classes $[x]$ that, nevertheless, do not necessarily constitute a partition of X as it is the case of the quotient set.

Remarks 3.1. *1. Notice that it is not possible to state that the relation $\leq_P$ is always a preorder. This supposition is only a working hypothesis: $\leq_P$ can be called the elemental meaning of P in X. Predicate P is* meaningless *in X if $\leq_P = \emptyset$.*

2. The structure $(X, \leq_P)$ is added to X by the empirical relation 'less P than', and reflects the (syntactical) structure X adopts by the use of P in it. It translates into algebraic terms the common belief that when properly speaking, the universe of discourse is ordered in some way.

3.2

Let us suppose that $(L, \leq)$ is a partially ordered set (poset)

Definition 3.2. *A mapping $\mu : X \to L$ is a $(L, \leq)$-degree for P, provided: $x \leq_P y \Rightarrow \mu(x) \leq \mu(y)$.*

Proposition 3.3. *Provided $\leq_P$ is a preorder, any $(L, \leq)$-degree distinguishes between the elements in each class of $X/\equiv_P$.*

Proof. If $x \equiv_P y$, since it means $x \leq_P y$ and $y \leq_P x$, it follows $\mu(x) \leq \mu(y)$ and $\mu(y) \leq \mu(x)$. Since $\leq$ is a partial order, it follows $\mu(x) = \mu(y)$. □

Consequently, any $(L,\leq)$-degree μ of P in X, as much has no more values than the number of classes in $X/\equiv_P$.

Theorem 3.4. *Given $(X,\leq_P)$, with $\leq_P$ a preorder, at least it always exists a degree.*

Proof. Take $L = X/\equiv_P$, and $[x] \leq [y]\, iff\, x \leq_P y$, a definition that is not dependent on the representatives of the two classes $[x]$ and $[y]$. It is easy to prove that $(X/\equiv_P, \leq)$ is a poset. Defining $\mu : X \to L$ by $\mu(x) = [x]$, it is clear that if $x \leq_P y$, then $\mu(x) \leq \mu(y)$. Hence, μ is a $(L,\leq)$-degree for P in X. □

This degree (the *natural* one) shows that the concept of degree is not an empty one when $\leq_P$ is a preorder.

Once a degree μ is defined for P, the relation '$x \leq_\mu y$ iff $\mu(x) \leq \mu(y)$' is also a preorder in X that is larger than $\leq_P$, since

$$x \leq_P y \Rightarrow \mu(x) \leq \mu(y) \Leftrightarrow x \leq_\mu y,$$

that is $\leq_P \subset \leq_\mu$. When $\leq_P = \leq_\mu$ it is said that the degree *perfectly reflects* the meaning of P in X, but if $\leq_P \subset \leq_{\mu_P}$ it is said that the degree only *partially reflects* such meaning.

Remark 3.5. *When it is taken the unit interval of the real line instead of the natural poset, since the unit interval is a totally ordered set, also $\leq_\mu$ enjoys a 'total' character. In fact, since it will be always either $\mu(x) \leq \mu(y)$, or $\mu(y) \leq \mu(x)$, it is either $x \leq_\mu y$, or $y \leq_\mu x$, that is, there cannot exist elements x, y not comparable under the preorder $\leq_\mu$. In these cases and usually, the degree only will reflect partially the elemental meaning of P in X. Hence, and in principle, to well capture the elemental meaning, it can be better a poset than a totally ordered set. A totally ordered set could be, from this theoretical point of view, a drawback to represent the predicate.*

4 L-SETS AND FUZZY SETS

4.1

Suppose that P is a predicate on X with elemental meaning $(X,\leq_P)$, with $\leq_P$ a preorder, and with an $(L,\leq)$-degree $\mu \in L^X$, and let us go back towards a possible representation of the collective generated by P in X.

Is it now possible to say something on collectives? At least,

1. Collectives generated by predicates with, as much, the two degrees *min* $(L,\leq) = \alpha$, *max*$(L,\leq) = \omega$ – provided they exist–, are 'sets'.
2. Collectives generated by predicates with a degree having, at least, one value different from α and ω, are not 'sets'.

Hence, how can be represented the collectives that are not sets? The following is a way similar to the representation of sets by functions in $\{0,1\}^X$.

Definition 4.1. *Let P be a predicate on X with elemental meaning $(X,\leq_P)$ and an $(L,\leq)$-degree μ_P.*

The L-set $\mathbb{P}$ defined by P in X is given by the definition

1. $\mu_P(x) = r \Leftrightarrow x \in_r \mathbb{P}$ *($\in_r$= belonging to the degree $r \in L$)*
2. $\mathbb{P} = \mathbb{Q} \Leftrightarrow \mu_P(x) = \mu_Q(x), \forall x \in X$.

Notice that when $\mathbb{P}$ is a set, the new symbols $\in_r$ are reduced to either $\in_\alpha$ ($=\notin$), or to $\in_\omega$ ($=\in$). When P is not a set there exists, at least, a third symbol $\in_r$ with $r \neq \alpha$ and $r \neq \omega$. Notice also that there is actually no difference between μ_P and $\mathbb{P}$, this last being only a new form of reading the function μ_P, and representing the collective generated by P in X with respect to the scale of degrees $(L,\leq)$.

Consequently, *there is no a single representation of the collective generated* by P in X: For each $(L,\leq)$ and each $(L,\leq)$-degree there is one of such representations. Of course, since $(\{\alpha,\omega\},\leq) \subset (L,\leq)$ is obviously isomorphic with the totally ordered subset $(\{0,1\},\leq)$ of the unit interval $([0,1],\leq)$, all the L-sets in $\{\alpha,\omega\}^X$ are nothing else than the (crisp) subsets of X, that is, $P(X)$ is isomorphic with $\{\alpha,\omega\}^X$.

4.2

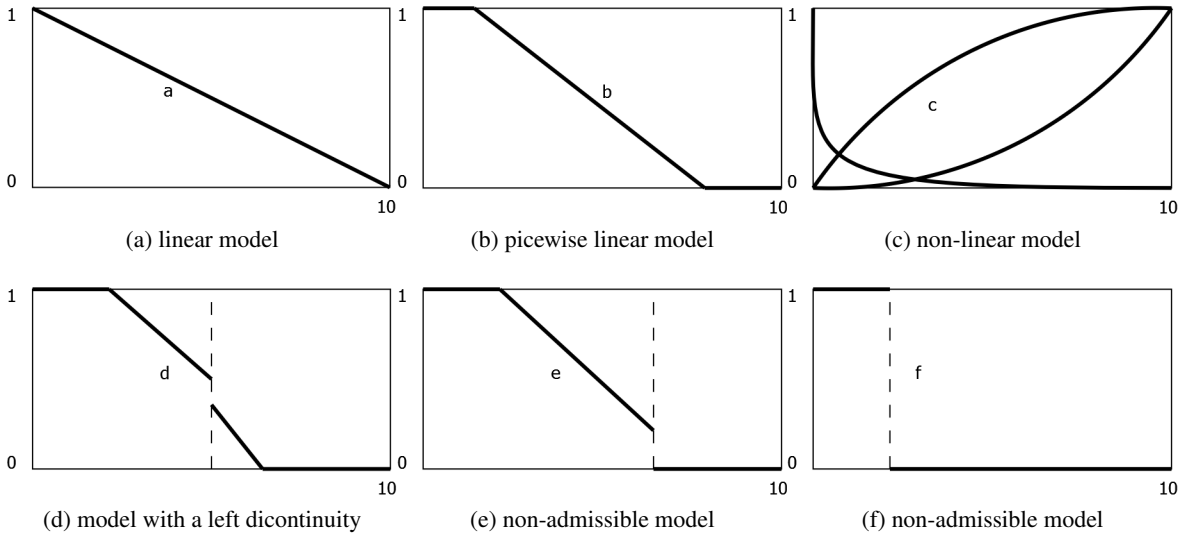
When $(L,\leq) = ([0,1],\leq)$ it is said that $\mu \in [0,1]^X$ is a *fuzzy set*, or a Zadeh's set. Since $(X,\leq_P)$ is not too often a totally ordered preset (preordered set), it will not often be the case that $\leq_P = \leq_\mu$, and μ only will partially reflect the elemental meaning of P in X. In addiction, to name $\mathbb{P}$ (the fuzzy set labelled P) the fuzzy set given by $\mu \in [0,1]^X$, it is now clear that it should be

$$x \leq_P y \Rightarrow \mu(x) \leq \mu(y),$$

that is, μ should be a $([0,1],\leq)$-degree for P in X.

5 CONCLUSION AND EXAMPLES

In the same way that $P(X)$, or $\{0,1\}^X$, collects all subsets of X without attributing a name to each one of them, the set $\mathfrak{F}(X) = [0,1]^X$ collects all possible fuzzy sets in X without attributing a name to each one of them. Anyway, in both cases many applications of both concepts are made under 'naming' the set by means of a linguistic label.



(a) linear model (b) picewise linear model (c) non-linear model

(d) model with a left dicontinuity (e) non-admissible model (f) non-admissible model

Figure 1

It should be noticed that all mathematical concepts can be translated into language, although some times this translation is neither easy, nor immediate. In the case of (crisp) sets, that naming is warranted by the *Axiom of Specificity*: Given on X a binary property p named P, it exists a unique set $\mathbb{P} \subset X$ such that

$$x \in \mathbb{P} \Leftrightarrow \mu_P(x) = 1 \Leftrightarrow x \text{ verifies (totally) the property } p.$$

What the axiom of specificity does not state is that any set A can be named by a single P. On the contrary, usually several names $P_1, ..., P_n$ can be attributed to the set, in which case all of them can be taken as *sinonyms*, as it is the case of $\{1,2,3\} \subset N$, for which it can be P_1 =The first three positive integers, or P_2 =The first three prime numbers.

The naming of the fuzzy sets $\mu \in [0,1]^X$ has a different, and approximate, character although with some similarity for what corresponds to the existence of sinonyms. Given $\mu \in [0,1]^X$ it can be named, or labelled, by P, provided μ is a $([0,1], \leq_P)$-degree for P in X, that is $\mu = \mu_P$. This, of course does not exclude that some other $\sigma \in [0,1]^X$ could also be a degree and, hence, taken as μ_P.

Example 5.1. *P = small and $X = [0,10]$. Since it can be agreed that in the common use of 'small' it is*

1. 0 *is completely* small
2. 10 *is completely* not small
3. *If x is small, and $0 < y < x$, then y is small*
4. *For any x that is small, there exists $\varepsilon > 0$ such that all $y \in [x, x+\varepsilon]$ are also small*

it can be concluded that $\leq_{small} = \leq^{-1}$ (with $\leq$ the total order of the real line restricted to [0,10]) and that all degrees μ for 'small' do verify

a. $\mu(0) = 0$, $\mu(10) = 1$
b. *If $x \leq y$, then $\mu(y) \leq \mu(x)$ (μ is decreasing)*
c. *If $\mu(x) > 0$, it is $\mu(y) > 0$ for all $y \in [x, x+\varepsilon]$.*

Hence, in figure 1 there are depicted some instances of fuzzy sets that can represent the collective of small numbers in $[0,10]$.

Notice that the function in figure (1f) represents the classical set specified in $[0,10]$ by the predicate P 'less than 3'. That is, the gradable predicate 'small' cannot be confused with *any one of the crisp predicates 'less than r', with $0 \leq r \leq 10$.*

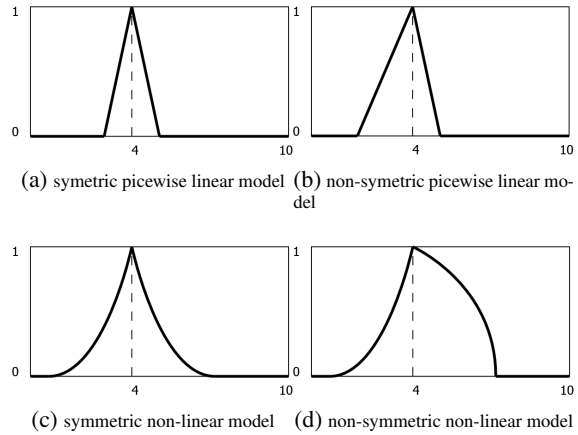
Example 5.2. *P = around four, $X = [0,10]$. Since it can be agreed that 'around four' is commonly used by the rules*

1. $\mu(4) = 1$, $\mu(0) = \mu(10) = 0$
2. *μ is decreasing between 4 and 10, and non-decreasing between 0 and 4,*

it can be concluded that

$$x \leq_P y \Leftrightarrow \begin{cases} x \leq y, & \text{if } x, y \in [0,4] \\ y \leq x, & \text{if } x, y \in [4,10] \end{cases}$$

Hence, next figures show some models of 'fuzzy sets' able to represent the collective of the numbers in [0,10] that are around four.



(a) symetric picewise linear model (b) non-symetric picewise linear model

(c) symmetric non-linear model (d) non-symmetric non-linear model

Figure 2

Remarks 5.3. *1. In the case of* small *(s), if μ_S is strictly decreasing, from $x \leq_S y \Leftrightarrow \mu_S(x) \leq \mu_S(y) \Leftrightarrow y \leq x \Leftrightarrow x \leq^{-1} y$, follows $\leq_{\mu_S} = \leq^{-1} = \leq_S$. That is, μ_S perfectly reflects $\leq_S$.*

2. *In the case of* around four *(a4), if x and y are in the base of the triangle (linear or not) in such a way that $x < 4 < y$, and $y \leq_{\mu_{a4}} x$. Since x and y are not comparable under $\leq_{a4}$, it follows that is not $\leq_{a4} = \leq_{\mu_{a4}}$, and μ_{a4} only reflects partially $\leq_{a4}$.*

3. *Notice that all curves in figures 1 and 2 keep the 'air of family' given, respectively, by the four rules in Example 5.1, and the two rules in Example 5.2.*

Acknowledgements

This work has been supported by the Foundation for the Advancement of Soft Computing (ECSC) (Asturias, Spain), and by the Spanish Department of Science and Innovation (MICINN) under project TIN2008-06890-C02-01.

The author is in debt with the comments made by the two reviewers.

References

[1] M. Black. Reasoning with loose concepts. *Dialogue*, 2:1–12, 1963.

[2] I. García-Honrado and E. Trillas. An essay on the linguistic roots of fuzzy sets. *Information Sciences*, 181:4061–4074, 2011.

[3] P.R. Halmos. *Naive set theory*. Springer-Verlag, New York, 1974.

[4] H.T. Nguyen and E.A. Walker. A First Course in Fuzzy Logic. 1997.

[5] E. Trillas. On the use of words and fuzzy sets. *Information Sciences*, 176(11):1463–1487, 2006.

[6] E. Trillas. On the genesis of fuzzy sets. *Agora*, 27(1):7–33, 2008.

[7] E. Trillas, C. Alsina, and J. M. Terricabras. *Introducción a la Lógica Borrosa*. Ariel, Barcelona, 1995 (In Spanish).

[8] E. Trillas and L.A. Urtubey. Towards the dissolution of the sorites paradox. *Applied Soft Computing*, 11(2):1506–1510, 2011.

Funciones de disimilaridad restringidas y funciones penalty

Humberto Bustince[1], **Javier Fernandez**[1], **Mikel Galar**[1], **Radko Mesiar**[2], **Ana Pradera**[3], **Gleb Beliakov**[4]

[1]*Universidad Pública de Navarra, Pamplona, Spain,*
[2]*Slovak University of Technology, Bratislava, Slovakia and UTIA CAS Prague, Czech Republik,*
[3]*Universidad Rey Juan Carlos, Madrid, Spain,*
[4]*Deakin University, Melbourne, Australia,*

Resumen

En este trabajo presentamos la definición de funció de disimilaridad restringida y relacionamos este nuevo concepto con otros, como las métricas. Mostramos en particular cómo podemos utilizar las funciones de disimilaridad restringidas para construir funciones penalty.

Palabras Clave: Agregación, Disimilaridad, Función penalty.

1 INTRODUCCIÓN

Dado un conjunto de datos o entradas, las funciones penalty (véase [4, 5, 7]) nos permiten determinar cuál es la función de agregación M_j que proporciona la salida y_j menos disimilar respecto al conjunto de entradas $\{x_1, \cdots, x_p\}$. Por otra parte, las funciones de disimilaridad son una herramienta que permite medir hasta qué punto son diferentes dos entradas dadas. Esta idea nos ha llevado a considerar el uso de funciones de disimilaridad restringidas [2] para construir funciones penalty. Con este objetivo en mente, investigamos en primer lugar bajo qué condiciones podemos construir funciones de disimilaridad restringidas que sean convexas o cuasi-convexas. Posteriormente analizamos la relación de estas funciones con las denominadas funciones de disimilaridad fieles (vease [7]).

La estructura de este trabajo es la siguiente. En la sección siguiente presentamos algunos conceptos y resultados preliminares. En la Sección 3 introducimos el concepto de función de disimilaridad restringida junto a algunos resultados relacionados con él. La Sección 4 se centra en el estudio de funciones de disimilaridad restringidas convexas o cuasi-convexas, y su relación con las métricas. La Sección 5 incluye algunos métodos de construcción de funciones de disimilaridad restringidas. Las Secciones 6 y 7 consideran funciones de disimilaridad restringidas fieles y su relación con funciones penalty. Terminamos con las conclusiones y referencias.

2 PRELIMINARES

Definición 1 *Una función $M : [0,1]^n \to [0,1]$ es una función de agregación si es monótona no decreciente en cada una de sus variables y se satisfacen las condiciones de frontera $M(0,0,\cdots,0) = 0$ y $M(1,1,\cdots,1) = 1$.*

Definición 2 *Una función de agregación M es una media si*

$$\min(x_1,\cdots,x_n) \leq M(x_1,\cdots,x_n) \leq \max(x_1,\cdots,x_n)$$

para todo $x_1,\ldots,x_n \in [0,1]$.

Cualquier función de agregación que sea una media es idempotente. De la misma forma, cualquier función de agregación idempotente es una media.

Definición 3 *Una función penalty es una aplicación $P : [0,1]^{n+1} \to [0,\infty]$ tal que:*

1. *$P(\mathbf{x},y) \geq 0$ para todo $\mathbf{x} \in [0,1]^n, y \in [0,1]$;*
2. *$P(\mathbf{x},y) = 0$ si y solo si $x_i = y$ para todo $i = 1,\cdots,n$;*
3. *$P(\mathbf{x},y)$ es cuasi-convexa en y para cualquier $\mathbf{x}$; es decir, para cada $\mathbf{x} \in [0,1]^n$ fijado, se verifica la desigualdad*

   $$P(\mathbf{x}, \lambda \cdot y_1 + (1-\lambda)\cdot y_2) \leq \max(P(\mathbf{x},y_1), P(\mathbf{x},y_2))$$

   para cualesquiera $\lambda \in [0,1]$ e $y_1, y_2 \in [0,1]$.

Definición 4 *Sea P una función penalty. Denominamos función basada en la función penalty P a la aplicación*

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x},y),$$

si y es el único mínimo, e $y = \frac{a+b}{2}$ si el conjunto de mínimos viene dado por un intervalo $[a,b]$.

Una función de agregación basada en una función penalty es siempre una media. El siguiente resultado establece que el converso también es cierto.

Teorema 1 *[4] Toda media puede representarse como una función basada en una función penalty en el sentido de la Definición 3.*

3 FUNCIONES DE DISIMILARIDAD RESTRINGIDAS

El concepto de *función de disimilaridad restringida* fue introducido en [2] como una herramienta para poder medir la disimilaridad entre dos datos dados. En el mismo trabajo se presentaron diversos teoremas de construcción y caracterización de tales funciones. En particular, se utilizaron las funciones de disimilaridad restringidas para construir distancias entre conjuntos difusos en el sentido de Liu [6].

Definición 5 *[2] Una aplicación $d_R : [0,1]^2 \to [0,1]$ es una función de disimilaridad restringida si:*

1. *$d_R(x,y) = d_R(y,x)$ para cada $x,y \in [0,1]$;*
2. *$d_R(x,y) = 1$ si y solo si $x = 0$ e $y = 1$ o $x = 1$ e $y = 0$; es decir, si $\{x,y\} = \{0,1\}$;*
3. *$d_R(x,y) = 0$ si y solo si $x = y$;*
4. *Para todo $x,y,z \in [0,1]$, si $x \le y \le z$, entonces $d_R(x,y) \le d_R(x,z)$ y $d_R(y,z) \le d_R(x,z)$.*

Diremos que d_R es una función de disimilaridad restringida estricta si para cualesquiera $x,y,z \in [0,1]$, si $x < y < z$, entonces $d_R(x,y) < d_R(x,z)$ y $d_R(y,z) < d_R(x,z)$.

Ejemplo 1 *La aplicación $d_R(x,y) = |x-y|$ proporciona un ejemplo simple de función de disimilaridad restringida que es estricta. Por otra parte, tenemos el siguiente ejemplo de función de disimilaridad restringida que no es estricta. Sea $c \in]0,1[$. Entonces*

$$d_R(x,y) = \begin{cases} 1 & \text{si } \{x,y\} = \{0,1\}\,; \\ 0 & \text{si } x = y; \\ c & \text{en otro caso.} \end{cases}$$

es una función de disimilaridad restringida. Nótese que ni siquiera es continua.

Recordemos que una negación difusa es una función no creciente $N : [0,1] \to [0,1]$ tal que $N(0) = 1$ y $N(1) = 0$. Una negación difusa se dice estricta si N es estrictamente decreciente. Una negación difusa estricta e involutiva se denomina negación fuerte.

Teorema 2 *Sea $d : [0,1]^2 \to [0,1]$ una función. Los siguientes enunciados son equivalentes.*

(i) $N : [0,1] \to [0,1]$ es una negación estricta y $d(x,y) = |N(x) - N(y)|$;

(ii) d es una función de disimilaridad restringida, $d(x,y) = |d(x,1) - d(y,1)|$ y $d(1,x)$ es estrictamente monótona decreciente;

(iii) d es una función de disimilaridad restringida y para todo $x,y,z \in [0,1]$ with $x \ge y \ge z$ se verifica que $d(x,y) + d(y,z) = d(x,z)$ y $d(1,x)$ es estrictamente monótona decreciente.

Dem. (i) $\Rightarrow$ (ii): Supongamos que $d(x,y) = |N(x) - N(y)|$. La simetría de d es obvia. Además, $d(x,y) = 0$ si y solo si $N(x) = N(y)$, y como N es estricta, esto ocurre si y solo si $x = y$. Por otra parte, $d(x,y) = 1$ si y solo si $\{N(x),N(y)\} = \{0,1\}$, y , de nuevo, al ser N estricta, esto equivale a $\{x,y\} = \{0,1\}$. Si $x \le y \le z$, entonces $d(x,y) = |N(x) - N(y)| = N(x) - N(y) \le N(x) - N(z) = |N(x) - N(z)| = d(x,z)$. El caso $d(y,z) \le d(x,z)$ es análogo. Obsérvese que, si $x < y < z$, entonces $d(x,y) < d(x,z)$ y $d(y,z) < d(x,z)$, por ser N estricta. Finalmente, $N(x) = |N(x) - N(1)| = d(x,1)$, luego $d(x,y) = |d(x,1) - d(y,1)|$.

(ii) $\Rightarrow$ (iii) Sean $x > y > z$. Entonces

$$\begin{aligned} d(x,y) + d(y,z) &= |N(x) - N(y)| + |N(y) - N(z)| \\ &= N(y) - N(x) + N(z) - N(y) = N(z) - N(x) \\ &= |N(x) - N(z)| = d(x,z)\,. \end{aligned}$$

La monotonicidad estricta de $d(1,x)$ se sigue de la simetría de d y de que $d(x,y) = |d(1,x) - d(1,y)| = 0$ si y solo si $x = y$.

(iii) $\Rightarrow$ (i) Sea $N(x) = d(1,x)$. En primer lugar, como $d(1,x)$ es estrictamente monótona, $d(1,1) = 0$ y $d(1,0) = 1$, se tiene que N es una negación estricta. Además, si $1 \ge y \ge z$, se sigue que

$$d(1,y) + d(y,z) = d(1,z)$$

luego $d(y,z) = d(1,z) - d(1,y)$, y como $y > z$, $d(y,z) = |d(1,y) - d(1,z)| = |N(y) - N(z)|$ □

La prueba del siguiente resultado es directa.

Proposición 1 *Sea d_R una función de disimilaridad restringida. Entonces*

1. *$N(x) = d_R(1,x)$ para todo $x \in [0,1]$ es una negación difusa;*
2. *$N(x) = d_R(1,x)$ es una negación involutiva si y solo si $d_R(1,d_R(1,x)) = x$ para todo $x \in [0,1]$;*
3. *Si $N(x) = d_R(1,x)$ y $d_R(x,y) = d_R(N(x),N(y))$ para todo $x,y \in [0,1]$, entonces $N(x) = d_R(0,N(x))$ para todo $x \in [0,1]$;*

4. *Si $N(x) = d_R(1,x)$ y $d_R(x,y) = d_R(N(x),N(y))$ para todo $x,y \in [0,1]$, entonces $d(1,d(1,x)) = x$ para todo $x \in [0,1]$; es decir, N es una negación involutiva.*

El ítem 1) de la Proposición 1 puede extenderse como sigue.

Proposición 2 *Sea d_R una función de disimilaridad restringida. Entonces, para cada $y \in]0,1]$, la aplicación*

$$N(x) = \frac{d_R(yx,y)}{d_R(0,y)}$$

es una negación difusa.

Dem. N está bien definida ya que estamos tomando $y \in]0,1]$ y si $y \neq 0$, entonces $d_R(0,y) \neq 0$. Además $0 \leq yx \leq y$ luego $d_R(yx,y) \leq d_R(0,y)$ y $0 \leq N(x) \leq 1$. Para terminar, $N(0) = \frac{d_R(0,y)}{d_R(0,y)} = 1$ y $N(1) = \frac{d_R(y,y)}{d_R(0,y)} = 0$ □

Respecto a la relación con automorfismos, tenemos el siguiente resultado.

Proposición 3 *Sea d_R una función de disimilaridad restringida continua y estricta. Entonces, para cada $y \neq 1$ la aplicación*

$$\varphi(x) = \frac{d_R((1-y)x+y,y)}{d_R(1,y)}$$

es un automorfismo del intervalo unidad.

Dem. Es un cálculo sencillo. Obsérvese que el denominador nunca se anula □

Obsérvese que los resultados de esta sección son relevantes de cara a considerar posibles métodos de construcción.

4 FUNCIONES DE DISIMILARIDAD RESTRINGIDAS CONVEXAS Y CUASI-CONVEXAS

Dado que vamos a relacionar las funciones de disimilaridad restringidas y las funciones penalty, el estudio de la convexidad y de la cuasi-convexidad es crucial.

Teorema 3 *Sea $d_R : [0,1]^2 \to [0,1]$ una función de disimilaridad restringida. Entonces d_R es cuasi-convexa en cada variable; es decir, para todo $x,y_1,y_2,\lambda \in [0,1]$*

$$d_R(x,\lambda \cdot y_1 + (1-\lambda)\cdot y_2) \leq \max(d_R(x,y_1), d_R(x,y_2))$$

Dem. Sean $x,\lambda \in [0,1]$. Sabemos que, para todo $y_1,y_2 \in [0,1]$, $\min(y_1,y_2) \leq \lambda \cdot y_1 + (1-\lambda)\cdot y_2 \leq \max(y_1,y_2)$. Hay tres posibilidades.

i) $x \leq \min(y_1,y_2)$. Entonces $d_R(x,\lambda \cdot y_1 + (1-\lambda)\cdot y_2) \leq d_R(x,\max(y_1,y_2)) \leq \max(d_R(x,y_1), d_R(x,y_2))$;

ii) $\min(y_1,y_2) \leq x \leq \max(y_1,y_2)$. En esta situación pueden ocurrir dos cosas:

a) $\min(y_1,y_2) \leq x \leq \lambda \cdot y_1 + (1-\lambda)\cdot y_2 \leq \max(y_1,y_2)$, luego $d_R(x,\lambda \cdot y_1 + (1-\lambda)\cdot y_2) \leq d_R(x,\max(y_1,y_2)) \leq \max(d_R(x,y_1), d_R(x,y_2))$, o

b) $\lambda \cdot y_1 + (1-\lambda)\cdot y_2 \leq x \leq \max(y_1,y_2)$, luego $d_R(x,\lambda \cdot y_1 + (1-\lambda)\cdot y_2) \leq d_R(x,\min(y_1,y_2)) \leq \max(d_R(x,y_1), d_R(x,y_2))$;

iii) $\max(y_1,y_2) \leq x$, que puede tratarse como el ítem i) □

Corolario 1 *Sea $d_R : [0,1]^2 \to [0,1]$ una función de disimilaridad restringida. Si $\lambda_0,\lambda_1 \in [0,1]$, entonces*

$$d_R(\lambda_0 \cdot x_1 + (1-\lambda_0)\cdot x_2, \lambda_1 \cdot y_1 + (1-\lambda_1)\cdot y_2) \leq \max(d_R(x_1,y_1), d_R(x_1,y_2), d_R(x_2,y_1), d_R(x_2,y_2))$$

para todo $x_1,x_2,y_1,y_2 \in [0,1]$.

Dem. De la cuasi-convexidad de cada una de las variables de d_R tenemos $d_R(\lambda_0 \cdot x_1 + (1-\lambda_0)\cdot x_2, \lambda_1 \cdot y_1 + (1-\lambda_1)\cdot y_2) \leq \max(d_R(\lambda_0 \cdot x_1 + (1-\lambda_0)\cdot x_2, y_1), d_R(\lambda_0 \cdot x_1 + (1-\lambda_0)\cdot x_2, y_2))$, y el resultado se sigue de la cuasi-convexidad en cada una de las variables. □

En lo que queda de este trabajo, cuando digamos que d_R es convexa o cuasi-convexa (cóncava o cuasi-cóncava) queremos decir que lo es en ambas variables simultáneamente. En otro caso lo indicaremos explícitamente.

Teorema 4 *Sea $d_R[0,1]^2 \to [0,1]$ una función de disimilaridad restringida cóncava en una variable. Entonces d_R es una métrica en $[0,1]$.*

Dem. Solo debemos comprobar si se satisface la desigualdad triangular, ya que las restantes propiedades demandadas a una métrica se verifican trivialmente. El único caso no trivial que debe comprobarse es cuando $0 \leq x < z < y \leq 1$. Entonces, la concavidad en cada variable asegura que, para cada $\lambda \in [0,1]$

$$\begin{aligned} d_R(x,\lambda x + (1-\lambda)y) &\geq \lambda d_R(x,x) + (1-\lambda) d_R(x,y) \\ &= (1-\lambda) d_R(x,y)\ . \end{aligned}$$

De la misma forma,

$$d_R(\lambda x + (1-\lambda)y, y) \geq \lambda d_R(x,y)\ .$$

Por tanto

$$\begin{aligned} d_R(x,y) &= \tfrac{y-z}{y-x} d_R(x,y) + \tfrac{z-x}{y-x} d_R(x,y) \\ &\leq d(x,z) + d(y,z)\ , \end{aligned}$$

lo que prueba que d_R es una métrica □

5 MÉTODOS DE CONSTRUCCIÓN DE FUNCIONES DE DISIMILARIDAD RESTRINGIDAS

Proposición 4 *Sea $\varphi : [0,1] \to [0,1]$ un automorfismo y $d_{R_1} : [0,1]^2 \to [0,1]$ una función de disimilaridad restringida. Entonces $d_R(x,y) = d_{R_1}(\varphi(x),\varphi(y))$ es una función de disimilaridad restringida. Además, si $d_{R_1}(x,1)$ es estrictamente monónotona, $d_R(x,y) = d_{R_1}(x,y)$ para todo $x,y \in [0,1]$ si y solo si $\varphi(x) = x$ para todo $x \in [0,1]$.*

Dem. El hecho de que d_R definida como en la proposición es una función de disimilaridad restringida es claro. Para probar el último punto, supongamos que $d_{R_1}(x,y) = d_R(x,y)$. Entonces, en particular, $d_{R_1}(1,y) = d_{R_1}(1,\varphi(y))$, y el resultado se sigue de la monotonicidad de $d_{R_1}(1,x)$ □

Proposición 5 *Si φ_1, φ_2 son dos automorfismos del intervalo unidad, entonces*

$$d(x,y) = \varphi_1(|\varphi_2(x) - \varphi_2(y)|)$$

es una función de disimilaridad estricta.

Proposición 6 *Sea φ un automorfismo y $K : [0,1] \to [0,1]$ una aplicación tal que:*

1. *$K(x) = 0$ si y solo si $x = 0$;*
2. *$K(x) = 1$ si y solo si $x = 1$;*
3. *K es creciente.*

Entonces

$$d_R(x,y) = K(|\varphi(x) - \varphi(y)|)$$

es una función de disimilaridad restringida.

Dem. Directa □

Corolario 2 *Sea d_R una función de disimilaridad restringida continua y estricta. Entonces, la aplicación*

$$d'_R(x,y) = d_R(0, |d_R(x,1) - d_R(y,1)|)$$

también es una función de disimilaridad restringida continua y estricta.

Proposición 7 *Sean d_R y d'_R definidas como en el Corolario anterior. Entonces, $d_R(x,y) = d'_R(x,y)$ para cada $x,y \in [0,1]$ si y solo si se verifica que*

$$d_R(x,y) = |d_R(x,1) - d_R(y,1)|$$

Dem. La necesidad es sencilla de ver. Basta observar que si $d_R(x,y) = d_R(0,|d_R(x,1) - d_R(y,1)|)$, entonces, para todo $x \in [0,1]$, $d_R(x,1) = d_R(0,d_R(x,1))$. Como $\varphi(x) = d_R(0,x)$ es un automorfismo del intervalo unidad y d_R es continua y estricta, vemos que $\varphi(x) = x$ para todo $x \in [0,1]$. La suficiencia es inmediata □

Tomando $y = 1$ vemos que

$$d'_R(x,1) = d_R(0,d_R(x,1))$$

Otros métodos de construcción de funciones de disimilaridad restringidas pueden encontrarse en [2].

6 FUNCIONES DE DISIMILARIDAD RESTRINGIDAS FIELES

En [4, 5, 7] se introdujo el siguiente concepto para construir funciones penalty.

Definición 6 *La función $p : X^2 \to \mathscr{R}_+$ se denomina función penalty fiel si $p(x,y) = 0$ si y solo si $x = y$ y además puede representarse como $p(t,s) = \mathbf{K}(h(t),h(s))$, donde $h : X \to \mathscr{R}$ es una función continua monótona (función de escalado) y $\mathbf{K} : \mathscr{R}^2 \to \mathscr{R}_+$ es convexa.*

Aquí, $X = [a,b]^n \subseteq \bar{\mathscr{R}} = [-\infty,+\infty]$. En este trabajo tomamos $p : [0,1]^2 \to [0,1]$, $h : [0,1] \to [0,1]$ y $\mathbf{K} : [0,1]^2 \to [0,1]$.

Proposición 8 *En las condiciones de la Proposición 4, si d_{R_1} es convexa, la función de disimilaridad restringida d_R es una función penalty fiel.*

Dem. Directa □

Una clase especial de funciones penalty fieles son las denominadas funciones de disimilaridad fieles, definidas en [7]. Una función de disimilaridad fiel p se escribe como

$$p(x,y) = K(h(x) - h(y))$$

con $K : \mathscr{R} \to \mathscr{R}$ convexa y con un único mínimo en $K(0) = 0$ y h una función continua y estrictamente monótona, $h : X \to \mathscr{R}$. En este trabajo tomamos $K : [-1,1] \to [0,1]$.

Usaremos la denominación *función de disimilaridad restringida fiel* para una función de similaridad fiel que también es función de disimilaridad restringida

Proposición 9 *En las condiciones de la Proposición 6, si K es convexa, entonces d_R es una función de disimilaridad restringida fiel.*

Dem. Directa □

Ejemplo 2
1. *Si $K(x) = x^2$, entonces $d_R(x,y) = (h(x) - h(y))^2$,*
2. *Si $K(x) = |x|$, entonces $d_R(x) = |h(x) - h(y)|$.*

Vamos a caracterizar un tipo particualr de función de disimilaridad restringida fiel.

Lema 5 *Sea d_R una función de disimilaridad restringida fiel tal que*

$$d_R(x,y) = K(h(x) - h(y))$$

con $K : [-1,1] \to [0,1]$ convexa y con un único mínimo en $K(0) = 0$ y $h : [0,1] \to [0,1]$ continua y estrictamente monótona. Entonces existe $K_0 : [-1,1] \to [0,1]$ convexa con un único mínimo en $K(0) = 0$ tal que si $0 < x < y$, entonces $K_0(x) < K_0(y)$ de modo que

$$d_R(x,y) = K_0(h(x) - h(y)) \, .$$

Dem. Dada $K : [-1,1] \to [0,1]$ como en el enunciado, si $0 < x < y$, $K(x) < K(y)$, no hay nada que probar. Supongamos que este no es el caso y h es creciente (el caso decreciente es análogo). La función

$$F : [0, h(1) - h(0)] \to [0,1] \text{ dada por}$$
$$x \to K(h(x) - h(0))$$

es una biyección ya que:

a) F es continua, dado que K es convexa y h es continua;

b) $F(0) = K(h(0) - h(0)) = K(0) = 0$;

c) $F(1) = K(h(1) - h(0)) = d_R(1,0) = 1$;

d) De la continuidad y los ítems b) y c) se sigue que F es sobreyectiva;

e) Si $F(x) = F(y)$, entonces $K(h(x) - h(0)) = K(h(y) - h(0))$; es decir,

$$x = d_R(x,0) = d_R(y,0) = y \, .$$

Por tanto, queda probada la inyectividad.

Consideremos la función:

$$r : [0,1] \to [0, h(1) - h(0)] \text{ dada por}$$
$$t \to (h(1) - h(0))t$$

donde r es creciente y $K \circ r$ es convexa. Además, como r es continua y biyectiva, también $K \circ r$ lo es. Si definimos $K_0 = K \circ r$ y tenemos en cuenta que K debe ser simétrica, la prueba está completa □

Teorema 6 *Sea $d_R : [0,1]^2 \to [0,1]$. Las siguientes afirmaciones son equivalentes.*

i) d_R es una función de disimilaridad restringida fiel tal que $d_R(x,y) = d_R(N(x),N(y))$ with $N(x) = d_R(1,x)$.

ii) Existe un automorfismo cóncavo φ del intervalo unidad tal que

$$d_R(x,y) = \varphi^{-1}(|\varphi(x) - \varphi(y)|)$$

Dem. i) $\Rightarrow$ ii) Como d_R es una función de disimilaridad restringida fiel, existe $K : [-1,1] \to [0,1]$ convexa tal que $K(x) = 0$ si y solo si $x = 0$ y existe $h : [0,1] \to [0,1]$ continua y estrictamente monótona tal que

$$d_R(x,y) = K(h(x) - h(y))$$

Por otra parte, $N(x) = d_R(1,x)$ es una negación fuerte. Por tanto, existe un automorfismo $\varphi : [0,1] \to [0,1]$ tal que

$$N(x) = d_R(1,x) = \varphi^{-1}(1 - \varphi(x))$$

y

$$d_R(1,x) = K(h(1) - h(x)) \, .$$

Además, $N(x) = d_R(1,x) = d_R(N(1),N(x)) = d_R(0,N(x))$. Como N es una negación fuerte, $x = d_R(0,x)$ para todo $x \in [0,1]$. De donde

$$x = d_R(0,x) = d_R(x,0) = K(h(x) - h(0)) \, .$$

Tomando $x = 1$, tenemos $1 = K(h(1) - h(0))$

Del Lema 5, si $0 < x < y$, entonces $K(x) \neq K(y)$.

Si K es estrictamente convexa, $h(1) - h(0) = 1$. Luego:

a) o h es un automorfismo, o

b) h es una negación estricta.

a) Si h es un automorfismo, entonces $x = d_R(x,0) = K(h(x) - h(0)) = K(h(x))$. Como K es inyectiva en $[0,1]$ se tiene que $K^{-1}(x) = h(x)$ es un automorfismo y como K es convexa, K^{-1} es cóncava.

Sea $\varphi(x) = K^{-1}(x)$. Entonces $h(x) = \varphi^{-1}(x)$ y

$$d_R(x,y) = \begin{cases} \varphi^{-1}(\varphi(x) - \varphi(y)) & \text{si } x \geq y \, ; \\ \varphi^{-1}(\varphi(y) - \varphi(x)) & \text{en otro caso.} \end{cases}$$

Esto es, $d_R(x,y) = \varphi^{-1}(|\varphi(x) - \varphi(y)|)$.

b) puede probarse análogamente □

7 FUNCIONES DE DISIMILARIDAD RESTRINGIDAS FIELES Y FUNCIONES PENALTY

Teorema 7 *Sea $M : [0,1]^n \to [0,1]$ una función de agregación y $d_R : [0,1]^2 \to [0,1]$ una función de disimilaridad restringida fiel. Entonces*

$$P : [0,1]^{n+1} \to [0,1] \text{ dada por}$$
$$P(\mathbf{x},y) = \underset{i=1}{\overset{n}{M}} d_R(x_i,y)$$

es una función penalty.

Dem. (1) $P(\mathbf{x},y) \geq 0$ para todo $\mathbf{x} \in [0,1]^n$ e $y \in [0,1]$. (2) Si $y = x_i$ para todo $i = 1,\cdots,n$, entonces $d_R(x_i,y) = 0$, luego $P(\mathbf{x},y) = 0$. (3) Como d_R es una función de disimlaridad restringida fiel, podemos escribirla como $d_R(x_i,y) = K(h(x_i) - h(y))$ con $K : [-1,1] \to [0,1]$ una función convexa con único mínimo en 0 ($K(0) = 0$) y $h : [0,1] \to [0,1]$ continua y estrictamente monótona. Luego d_R es cuasi-convexa.

Si h es creciente (el caso decreciente es análogo)

$$\begin{aligned} & P(\mathbf{x}, \lambda y_1 + (1-\lambda)y_2) \\ &= \underset{i=1}{\overset{n}{M}} d_R(x_i, \lambda y_1 + (1-\lambda)y_2) \\ &= \underset{i=1}{\overset{n}{M}} K(h(x_i) - h(\lambda y_1 + (1-\lambda)y_2)) \\ &\leq \begin{cases} MK(h(x_i) - h(\max(y_1,y_2))) \\ \text{ si } h(x_i) - h(\lambda y_1 + (1-\lambda)y_2) \leq 0; \\ MK(h(x_i) - h(\min(y_1,y_2))) \\ \text{ en otro caso.} \end{cases} \end{aligned}$$

En cualquier caso, esto es menor o igual que $\max(\underset{i=1}{\overset{n}{M}} K(h(x_i) - h(y_1)), \underset{i=1}{\overset{n}{M}} K(h(x_i) - h(y_2)))$ □

Corolario 3 *En las condiciones del Teorema 7, si $M : [0,1]^n \to [0,1]$ es una función de agregación tal que $M(x_1,\ldots,x_n) = 0$ si y solo si $x_1 = \cdots = x_n = 0$ y $M(x_1,\ldots,x_n) = 1$ si y solo si $x_1 = \cdots = x_n = 1$, entonces*

1. *$P(\mathbf{x},y) = 0$ si y solo si $\{x_i,y\} = \{0,1\}$ para todo $i = 1,\cdots,n$;*
2. *$P(\mathbf{x},y) = 1$ si y solo si $x_i = y$ para todo $i = 1,\cdots,n$;*
3. *Si $\mathbf{x} = (1,\cdots,1)$, entonces $P(\mathbf{x},y) = \underset{i=1}{\overset{n}{M}} n(y)$ con $n(y) = d_R(1,y)$;*
4. *$P(\mathbf{x},1) = \underset{i=1}{\overset{n}{M}} n(x_i)$ con $n(x_i) = d_R(1,x_i)$ para todo $i = 1,\cdots,n$;*
5. *Si φ es un automorfismo cóncavo, entonces*

$$P(\mathbf{x},y) = \underset{i=1}{\overset{n}{M}} \varphi^{-1}(|\varphi(x_i) - \varphi(y)|)$$

es una función penalty.

Dem. Directa □

Corolario 4 *En las condiciones del Teorema 7, si M es idempotente, entonces si $\mathbf{x} = (1,\cdots,1)$, $P(\mathbf{x},y) = n(y)$ con $n(y) = d_R(1,y)$.*

Dem. Directa □

Proposición 10 *Sea $M : [0,1]^n \to [0,1]$ una función de agregación idempotente y $d_R : [0,1]^2 \to [0,1]$ una función de disimilaridad restringida fiel. Entonces la función basada en la función penalty*

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x},y) = \arg\min_y \underset{i=1}{\overset{n}{M}} d_R(x_i,y)$$

es idempotente.

Dem. Directa □

8 CONCLUSIONES

En este trabajo hemos recordado el concepto de función de disimilaridad restringida y lo hemos relacionado con las funciones penalty a través del concepto de función de disimilaridad restringida fiel. También hemos analizado la relación con métricas

En el futuro pretendemos desarrollar los aspectos teóricos de este trabajo, analizando en particular qué tipos de medias pueden ser generadas por medio de funciones penalty construídas a partir de funciones de disimilaridad restringidas.

Agradecimientos

H. Bustince, J. Fernandez y M. Galar son apoyados por el proyecto TIN 2010-15055 del Gobierno de España. R. Mesiar está apoyado por las becas GACR P 402/11/0378, VEGA 1/0080/10 y TUCI-0601-66. A. Pradera está apoyada por los proyectos TIN2009-07901 y TEC2009-14587-C03-03 del Gobierno de España.

Referencias

[1] G. Beliakov, A. Pradera and T. Calvo: *Aggregation Functions: A Guide for Practitioners* (Studies In Fuzziness and Soft Computing, Springer); 2007.

[2] H. Bustince, E. Barrenechea and M. Pagola, Relationship between restricted dissimilarity functions, restricted equivalence functions and normal EN-functions: Image thresholding invariant, *Pattern Recognition Letters*, 29: 525-536, 2008 .

[3] T. Calvo, A. Kolesarova, M. Komornikova and R. Mesiar, Aggregation operators: properties, classes and construction methods. In T. Calvo, G. Mayor and R. Mesiar (Eds.): *Aggregation Operators New Trends and Applications*, pages 3–104, Physica-Verlag, Heidelberg, 2002.

[4] T. Calvo and G. Beliakov, Aggregation functions based on penalties, *Fuzzy Sets and Systems* 161:1420-1436, 2010.

[5] T. Calvo, R. Mesiar and R. Yager, A quantitative weights and aggregation, *IEEE Trans. Fuzzy Systems*, 12:62–69 2004.

[6] X. Liu, Entropy, distance measure and similarity measure of fuzzy sets and their relations, *Fuzzy Sets and Systems* 52:305–318, 1992.

[7] R. Mesiar, Fuzzy set approach to the utility, preference relations,and aggregation operators, *European J. Oper. Res* 176:414-422, 2007.

FUNCIONES T–S LINEALES CONVEXAS: UNA GENERALIZACIÓN DE LA ECUACIÓN DE FRANK

Tomasa Calvo[1], **Javier Martín**[2], **Gaspar Mayor**[2]

[1]*Universidad de Alcalá*, tomasa.calvo@uab.es
[2]*Universitat de les Illes Balears*, {javier.martin,gmayor}@uib.es

Resumen

En este trabajo se plantea y resuelve la ecuación funcional $T+\lambda S = \mathrm{Min}+\lambda\,\mathrm{Max}$, con $0 \leqslant \lambda \leqslant \infty$, para pares (T,S) con una t-norma T y una t-conorma S. Este tipo de ecuaciones aparecen al considerar operadores del tipo combinación lineal convexa de una t-norma y una t-conorma y plantear el problema de la intersección entre los segmentos determinados por el par $(\mathrm{Min},\mathrm{Max})$ y por otro cualquiera (T,S).

Palabras clave: Función de agregación, Combinación lineal convexa, Norma y conorma triangulares, Ecuación de Frank, Cópula, Cópula dual, Asociatividad.

1 INTRODUCCIÓN

La bien conocida ecuación funcional de Frank para t-normas y t-conormas

$$T(x,y)+S(x,y)=x+y,\ \ x,y\in[0,1] \tag{1}$$

surge de forma natural en el estudio de la asociatividad de una clase de operaciones binarias sobre distribuciones de probabilidad [2]. Si (T,S) es una solución de (1), debido a la monotonía de S, T tiene que ser una cópula, digamos C, y S la cópula dual asociada $\hat{C}$ dada por $\hat{C}(x,y)=x+y-C(x,y)$. Teniendo en cuenta las propiedades de $\hat{C}$, resolver la ecuación (1) es equivalente a hallar los pares $(C,\hat{C})$ con C cópula y $\hat{C}$ su dual, de manera que C y $\hat{C}$ sean simultáneamente asociativas.

Consideremos la familia de cópulas estrictas

$$C_\alpha(x,y)=-\frac{1}{\alpha}\log\left[1+\frac{(e^{-\alpha x}-1)(e^{-\alpha y}-1)}{e^{-\alpha}-1}\right],$$

$\alpha\neq 0$, y sus casos límites $C_{-\infty}=W$, $C_0=P$, $C_\infty=\mathrm{Min}$. En [3] se demuestra que un par (T,S) es solución de la ecuación (1) si, y solo si,

(a) $T=C_\alpha$, $\alpha\in[-\infty,\infty]$ y $S=\hat{T}$ (S es también T^*, dual de T), o bien

(b) T es una suma ordinal de miembros de la familia C_α, $\alpha\in[-\infty,\infty]$, y $S=\hat{T}$ ($S=T^*$ si, y solo si, el conjunto de idempotentes de T es simétrico respecto del punto $x=\frac{1}{2}$).

Esta ecuación, que juega un papel importante en diversos campos como lógica multivaluada, teoría de conjuntos borrosos y modelado de preferencias, una vez resuelta, ha sido objeto de atención en varios aspectos. En [1] se recogen varias extensiones de la misma, una para intervalos arbitrarios cerrados $[a,b]$, $-\infty\leqslant a<b\leqslant\infty$, y otra generalizando también la propia expresión en la forma

$$\Phi(T(x,y),S(x,y))=\Phi(x,y),\ \ x,y\in[a,b], \tag{2}$$

siendo Φ una función definida en $[a,b]^2$. Si Φ es conmutativa, el par $(\mathrm{Min},\mathrm{Max})$ es solución de (2) y, en algunos casos, por ejemplo cuando $\Phi(x,y)=\frac{x^2+y^2}{xy}$, ese par es la única solución. En [1] se considera el caso en que Φ es conmutativa y expresable en la forma

$$\Phi(x,y)=f(g(x)+g(y)), \tag{3}$$

donde g es una función continua y estrictamente creciente definida en $[a,b]$ y con rango un intervalo J, y f es invertible sobre el intervalo $J+J$. En este trabajo se plantea y resuelve la ecuación (2) para los casos

$$\begin{aligned}\Phi_1(x,y)&=\min\{x,y\}+k\max\{x,y\},\\ \Phi_2(x,y)&=k\min\{x,y\}+\max\{x,y\},\end{aligned} \tag{4}$$

siendo $0<k<1$. Como se comprobará más adelante, estas funciones no son de la forma (3).

El origen de las ecuaciones

$$\begin{aligned}T(x,y)+kS(x,y)&=\min\{x,y\}+k\max\{x,y\},\\ kT(x,y)+S(x,y)&=k\min\{x,y\}+\max\{x,y\},\end{aligned} \tag{5}$$

$\forall x,y\in[0,1]$, $0<k<1$, puede hallarse a la hora de responder a la pregunta de si el segmento $LC(\mathrm{Min},\mathrm{Max})=$

$\{(1-a)\,\mathrm{Min}+a\,\mathrm{Max}; a\in[0,1]\}$ tiene intersección no vacía con $LC(T,S)$ para alguna t-norma T y alguna t-conorma S. Este problema puede plantearse de forma general al considerar en lugar del par $(\mathrm{Min},\mathrm{Max})$ un par cualquiera (T_0,S_0). Bajo esta interpretación, la ecuación (1) responde a la cuestión de hallar un par (T,S) de tal manera que el punto medio del segmento $LC(T,S)$ coincida con el punto medio del segmento $LC(\mathrm{Min},\mathrm{Max})$, o dicho de otro modo, la media aritmética de T y S coincida con la media aritmética de Min y Max.

El trabajo comienza con un repaso de conceptos básicos de t-normas, t-conormas y cópulas (ver, por ejemplo, [4, 6]) en la Sección 2 y de la noción de funciones de agregación con frontera lineal en la 3. En la Sección 4 se demuestra que las funciones (4) no son de la forma (3) y, finalmente, en la Sección 5 se resuelve la generalización de la ecuación de Frank propuesta.

2 PRELIMINARES

Definición 1. *Una t-norma es una función de dos variables $T:[0,1]^2\to[0,1]$ que es asociativa, conmutativa, monótona creciente en cada variable y con elemento neutro* 1. *Una t-conorma es una función de dos variables $S:[0,1]^2\to[0,1]$ que es asociativa, conmutativa, monótona creciente en cada variable y con elemento neutro* 0.

Proposición 1. *Una función $T:[0,1]^2\to[0,1]$ es una t-norma si, y solo si, la función T^* definida por $T^*(x,y)=1-T(1-x,1-y)$ es una t-conorma. Una función $S:[0,1]^2\to[0,1]$ es una t-conorma si, y solo si, la función S^* definida por $S^*(x,y)=1-S(1-x,1-y)$ es una t-norma. Decimos que T^* es la t-conorma dual de T (S^* es la t-norma dual de S).*

Definición 2. *Sea $g:[0,1]\to[0,\infty]$ una función continua y estrictamente decreciente con $g(1)=0$, y sea $g^{(-1)}$ la pseudo-inversa de g, definida por*

$$g^{(-1)}(t)=\inf\{x\in[0,1];g(x)<t\}.$$

La función

$$T(x,y)=g^{(-1)}(g(x)+g(y)),\ \ x,y\in[0,1] \qquad (6)$$

es una t-norma continua. Decimos que g es un generador aditivo de T.

Las t-normas representables en la forma anterior (6) vienen caracterizadas por:

Proposición 2. *Una t-norma continua T admite una representación de la forma (6) si, y solo si, $T(x,x)<x\ \forall x\in(0,1)$ (los únicos elementos idempotentes son los triviales* 0 *y* 1*).*

Las t-normas de la forma (6) reciben el nombre de arquimedianas. El generador de una t-norma arquimediana es único salvo constantes multiplicativas positivas. Si $g(0)=\infty$, la t-norma se llama estricta, en caso contrario se llama no estricta o nilpotente. Las t-normas básicas $P(x,y)=xy$ (producto), $W(x,y)=\max\{x+y-1,0\}$ (Łukasiewicz) son arquimedianas con generadores aditivos $g(t)=-\log(t)$ y $g(t)=1-t$ respectivamente. La primera es estricta y la otra nilpotente.

Definición 3. *Una cópula (2-dimensional) es una función de dos variables $C:[0,1]^2\to[0,1]$ que satisface las condiciones*

(i) $C(x,0)=C(0,x)=0$ *y* $C(x,1)=C(1,x)=x\ \forall x$ *(*0 *es absorbente y* 1 *es neutro).*

(ii) $C(x,y)+C(x',y')\geqslant C(x,y')+C(x',y)$ *siempre que* $x\leqslant x'$ *e* $y\leqslant y'$ *(C es 2-creciente).*

Proposición 3.

(i) Una t-norma T es una cópula si, y solo si, es lipschitziana con constante 1: $T(x',y')-T(x,y)\leqslant x'-x+y'-y$ *siempre que* $x\leqslant x', y\leqslant y'$.

(ii) Una cópula C es una t-norma si, y solo si, es asociativa.

Definición 4. *Sea C una cópula.*

(i) La función $\hat{C}:[0,1]^2\to[0,1]$ definida por $\hat{C}(x,y)=x+y-C(x,y)$ se llama la dual de C. $\hat{C}$ no es, en general, una cópula.

(ii) La función $C^*:[0,1]^2\to[0,1]$ definida por $C^*(x,y)=1-C(1-x,1-y)$ se llama la co-cópula de C. C^* no es, en general, una cópula.

Proposición 4. *Una cópula C admite una representación de la forma (6) si, y solo si, g es una función convexa.*

Así, las cópulas arquimedianas forman una subclase de la clase de las t-normas arquimedianas, determinada por aquellas que tienen generador aditivo convexo.

3 FUNCIONES DE AGREGACIÓN CON FRONTERA LINEAL

Definición 5. *Una función $F:[0,1]^2\to[0,1]$ decimos que es una función de agregación si es monótona creciente en cada variable y tal que $F(0,0)=0$, $F(1,1)=1$.*

Definición 6. *Una función de agregación $F:[0,1]^2\to[0,1]$ que satisface, para todos x,y:*

(i) $F(x,0)=F(1,0)x$, $F(0,y)=F(0,1)y$,

(ii) $F(x,1)=(1-F(0,1))x+F(0,1)$, $F(1,y)=(1-F(1,0))y+F(1,0)$,

decimos que es una función de agregación con frontera lineal.

En [5] se introduce el concepto de operación de agregación como el de una función $F:[0,1]^2\to[0,1]$ que es monótona creciente, conmutativa y tal que:

$$F(x,0)=F(1,0)x,$$
$$F(x,1)=(1-F(1,0))x+F(1,0),\ \ \forall x.$$

En este sentido, una operación de agregación es una función de agregación con frontera lineal y conmutativa. Puede demostrarse que las t-normas y t-conormas son las operaciones de agregación asociativas. También son operaciones de agregación las combinaciones lineales convexas de una t-norma y una t-conorma cualesquiera $(1-a)T+aS$, $a\in[0,1]$. Más aún, si indicamos por $\mathscr{F}$ el conjunto de las operaciones de agregación, puede comprobarse que $\mathscr{F}$ es una parte convexa del espacio vectorial real de las funciones reales definidas en $[0,1]^2$. En este trabajo nos interesamos por los segmentos de extremos una t-norma y una t-conorma.

4 LAS FUNCIONES Φ_1 Y Φ_2

Veamos que las funciones Φ_1 y Φ_2, definidas por las expresiones (4), no son de la forma (3).

En efecto, supongamos que $\min\{x,y\}+k\max\{x,y\}=f(g(x)+g(y))$, con $g(0)=0$ y $g(1)=1$. Para $0\leqslant x\leqslant y\leqslant 1$ se tiene $x+ky=f(g(x)+g(y))$, y tomando $x=0$ se obtiene $ky=f(g(y))\ \forall y$, por tanto, $f(t)=kg^{-1}(t)$, $t\in[0,1]$. Podemos pues escribir $x+ky=f(g(x)+g(y))=kg^{-1}(g(x)+g(y))$ siempre que $0\leqslant x\leqslant y\leqslant 1$ y $g(x)+g(y)\leqslant 1$.

Si hacemos $g(x)=u, g(y)=v$, entonces para $0\leqslant u\leqslant v\leqslant 1, u+v\leqslant 1$, podemos escribir $g^{-1}(u)+kg^{-1}(v)=kg^{-1}(u+v)$ y, también, con $h=g^{-1}:[0,1]\to[0,1]$ y $\frac{1}{k}=r$,

$$h(u+v)=rh(u)+h(v), 0\leqslant u\leqslant v\leqslant 1, u+v\leqslant 1. \quad (7)$$

Si h fuera solución de (7):

(i) $h(\frac{1}{2^n})=(\frac{1}{1+r})^n$, $\forall n\geqslant 1$.

(ii) $h(\frac{m}{n})=((m-1)r+1)h(\frac{1}{n})$, $\forall n\geqslant 1$, $\forall m: 1\leqslant m\leqslant n$.

(iii) $h(\frac{1}{n})=\frac{1}{(n-1)r+1}$, $\forall n\geqslant 1$.

En efecto, para $n\geqslant 2$,

$$\begin{aligned}1 &= h(1)\\ &= h(\tfrac{1}{n}+\tfrac{n-1}{n})\\ &= rh(\tfrac{1}{n})+h(\tfrac{n-1}{n})\\ &= rh(\tfrac{1}{n})+((n-2)r+1)h(\tfrac{1}{n})\\ &= ((n-1)r+1)h(\tfrac{1}{n}).\end{aligned}$$

(iv) Tendría que ser $\frac{1}{(1+r)^n}=\frac{1}{(2^n-1)r+1}$ $\forall n\geqslant 1$, que solo es cierto si $r=1$.

Así queda probado que Φ_1 y Φ_2 no son de la forma (3).

Observaciones 1. *No puede ser $g(1)=\infty$, ya que si lo fuera $f(t)=kg^{-1}(t)$ $\forall t\in[0,\infty]$, y para $0\leqslant x\leqslant y\leqslant 1$ podríamos escribir $x+ky=f(g(x)+g(y))=kg^{-1}(g(x)+g(y))$, o sea $g^{-1}(g(x)+g(y))=\frac{1}{k}x+y$, lo que es una contradición.*

5 LA GENERALIZACIÓN DE LA ECUACIÓN DE FRANK

Dado un par (T_0,S_0), ¿existe un par (T,S), $T\neq T_0$, $S\neq S_0$, tal que los segmentos $LC(T,S)$ y $L(T_0,S_0)$ se cortan en algún punto? O, en otras palabras, ¿existen (T,S) y $a,b\in(0,1)$ tales que $(1-a)T+aS=(1-b)T_0+bS_0$? En estas condiciones, podemos enunciar el siguiente resultado.

Proposición 5. *La ecuación $(1-a)T+aS=(1-b)T_0+bS_0$, $0<a,b<1$, tiene solución si, y solo si, existe $k>0$ tal que $T+kS=T_0+kS_0$.*

Demostración. Supongamos

$$(1-a)T(x,y)+aS(x,y)=(1-b)T_0(x,y)+bS_0(x,y),$$

$\forall x,y\in[0,1]$, con $0<a,b<1$, lo que implica primeramente, haciendo $(x,y)=(1,0)$, que $a=b$.

Dividiendo por $1-a$ y haciendo $k=\frac{a}{1-a}$ se obtiene la expresión deseada $T+kS=T_0+kS_0$, con $k>0$.

Recíprocamente, si $T+kS=T_0+kS_0$ con $k>0$, podemos hallar a, $0<a<1$, (único) tal que $k=\frac{a}{1-a}$.

Por tanto, $T+\frac{a}{1-a}S=T_0+\frac{a}{1-a}S_0$ y, finalmente, $(1-a)T(x,y)+aS(x,y)=(1-a)T_0(x,y)+aS_0(x,y)$, con $0<a<1$. ∎

En el caso de fijar el par $(\mathrm{Min},\mathrm{Max})$, el estudio de la ecuación $T(x,y)+kS(x,y)=\min\{x,y\}+k\max\{x,y\}$, $k>0$, $k\neq 1$, es equivalente al estudio de las ecuaciones $kT+S=k\,\mathrm{Min}+\mathrm{Max}$ y $T+kS=\mathrm{Min}+k\,\mathrm{Max}$, $0<k<1$.

5.1 La ecuación $T+kS=\mathrm{Min}+k\,\mathrm{Max}$

Consideremos la ecuación funcional:

$$T(x,y)+kS(x,y)=\min\{x,y\}+k\max\{x,y\},\ 0<k<1. \quad (8)$$

Empecemos observando que el par $(\mathrm{Min},\mathrm{Max})$ es solución para todo k y, si $k\neq k'$, entonces este par es la única solución común de las ecuaciones correspondientes a los parámetros k y k'.

Unos primeros resultados sobre cómo tienen que ser las soluciones de la ecuación (8) están expuestos a continuación.

Proposición 6. *Si (T,S) es una solución de (8), entonces:*

(i) T es continua si, y solo si, S es continua.

(ii) El conjunto de los elementos idempotentes de T, $\{x \in [0,1]; T(x,x) = x\}$, coincide con el conjunto de los elementos idempotentes de S, $\{x \in [0,1]; S(x,x) = x\}$.

(iii) Si T y S son continuas, entonces pueden ser:

(iii.1) $T = \mathrm{Min}$, $S = \mathrm{Max}$.

(iii.2) T y S arquimedianas (las dos estrictas o las dos no estrictas).

(iii.3) $T = (\langle a_i, b_i, T_i\rangle)_{i\in I}$, $S = (\langle a_i, b_i, S_i\rangle)_{i\in I}$, sumas ordinales con $(T_i, S_i)_{i\in I}$ soluciones (arquimedianas) de (8).

(iv) $\max\{\min\{x,y\} + k\max\{x,y\} - k, 0\} \leqslant T(x,y) \leqslant S(x,y) \leqslant \min\{\frac{1}{k}\min\{x,y\} + \max\{x,y\}, 1\}$.

Observaciones 2.

(i) $F_k(x,y) = \max\{\min\{x,y\} + k\max\{x,y\} - k, 0\}$ es una cópula pero no una t-norma. En los casos límite $F_0 = \mathrm{Min}$ y $F_1 = W$.

(ii) $G_k(x,y) = \min\{\frac{1}{k}\min\{x,y\} + \max\{x,y\} - k, 1\}$ no es una t-norma. En los casos límite $G_0 = \mathrm{Max}$ y $G_1 = W^*$.

(iii) Si C es una cópula tal que $C \geqslant F_k$, consideremos la función:

$$\tilde{C}(x,y) = \frac{1}{k}\left(\min\{x,y\} + k\max\{x,y\} - C(x,y)\right).$$

Entonces:

(iii.1) $0 \leqslant \tilde{C}(x,y) \leqslant 1$.

(iii.2) $\tilde{C}$ es commutativa si, y solo si, lo es C.

(iii.3) $\tilde{C}(x,0) = \tilde{C}(0,x) = x$; $\tilde{C}(x,1) = \tilde{C}(1,x) = 1\ \forall x$.

(iii.4) $\tilde{C}$ es creciente en cada variable.

Diremos que $\tilde{C}$ es la k-dual de C. Está claro que en el caso límite $k = 1$, $\tilde{C}$ no es más que la ya conocida dual $\hat{C}$.

(iv) En particular, la k-dual de F_k: $\tilde{F}_k(x,y) = \frac{1}{k}(\min\{x,y\} + k\max\{x,y\} - F_k(x,y)) = \min\{\frac{1}{k}\min\{x,y\} + \max\{x,y\}, 1\} = G_k$.

Un hecho importante a destacar es el siguiente.

Proposición 7. *Si (T,S) es una solución de (8), entonces:*

(i) T es una cópula y $S = \frac{1}{k}(\mathrm{Min} + k\mathrm{Max} - T)$ es la k-dual de T. S es lipschitziana con constante $\frac{1}{k}$.

(ii) T y S son continuas.

Demostración. Veamos que T es de Lipschitz con constante 1. En efecto, si $x' \leqslant x$ entonces para todo $y \in [0,1]$ se tiene $T(x,y) - T(x',y) = \min\{x,y\} + k\max\{x,y\} - kS(x,y) - \min\{x',y\} - k\max\{x',y\} + kS(x',y) = \min\{x,y\} - \min\{x',y\} + k(\max\{x,y\} - \max\{x',y\}) - k(S(x,y) - kS(x',y)) \leqslant \min\{x,y\} - \min\{x',y\} + k(\max\{x,y\} - \max\{x',y\}) \leqslant x - x'$. ∎

A partir de ahora, de acuerdo con el apartado (iii) de la Proposición 6, basta reducir el estudio a pares de soluciones arquimedianas. Por este camino, podemos establecer el siguiente resultado.

Proposición 8. *Si (T,S) es una solución arquimediana de (8), entonces:*

(i) T y S tienen que ser estrictas

(ii) Los generadores aditivos de T y S son funciones convexas.

Demostración. Si T tuviera región cero, existiría $a \in (0,1)$ tal que $T(a,a) = 0$. En $[0,a]^2$, $T(x,y) = 0$ y debería ser $S(x,y) = \frac{1}{k}\min\{x,y\} + \max\{x,y\}$. Tomamos $0 < b < \frac{a}{\frac{1}{k}+2}$. Esta acotación nos asegura $2b, S(b,b), S(b,2b) < a$. Veamos que S no es asociativa.

- $S(S(b,b),2b) = S\left(\frac{b}{k} + b, 2b\right) = \frac{2b}{k} + \frac{b}{k} + b = \frac{3b}{k} + b$,
- $S(b,S(b,2b)) = S\left(b, \frac{b}{k} + 2b\right) = \frac{b}{k} + \frac{b}{k} + 2b = \frac{2b}{k} + 2b$.

Pero la igualdad se da solo si $k = 1$.

Por otra parte S es lipschitziana de constante $\frac{1}{k}$ (mayor que 1) y si $z \leqslant x \leqslant y$ entonces $S(y,z) - S(x,z) \leqslant y - x$. Si S tiene generador aditivo g, entonces $g^{-1}\left(\frac{x+y}{2}\right) \geqslant \frac{g^{-1}(x) + g^{-1}(y)}{2}$ siempre que $0 \leqslant x \leqslant y \leqslant 3x$, lo que implica que g es convexo. ∎

Siendo (T,S) una solución arquimediana estricta de (8), esta ecuación puede escribirse ahora a partir de los generadores aditivos f y g de T y S, respectivamente:

$$\begin{aligned} f^{-1}(f(x) + f(y)) + kg^{-1}(g(x) + g(y)) \\ = \min\{x,y\} + k\max\{x,y\},\ \ 0 < k < 1, \end{aligned} \tag{9}$$

siendo $f, g\colon [0,1] \to [0,\infty]$ funciones estrictamente monótonas (f decreciente y g creciente), continuas y convexas, con $f(1) = g(0) = 0$, $f(0) = g(1) = \infty$.

En la proposición que sigue se enuncian propiedades de los generadores aditivos f y g. La demostración se obtiene de forma semejante a la conseguida por Frank en [3] para resolver la ecuación $T + S = \mathrm{Min} + \mathrm{Max}$.

Proposición 9. *Si (f,g) es una solución de (9), entonces:*

(i) f y g son derivables en el intervalo abierto $(0,1)$.

(ii) f' y g' son crecientes y continuas en $(0,1)$.

(iii) $f'_+(0) = -\infty$, $f'(x) < 0\ \forall x \in (0,1)$; $f'_-(1) \leqslant 0$.

(iv) $f'_+(0) \geqslant 0$, $g'(x) > 0\ \forall x \in (0,1)$; $g'_-(1) = +\infty$.

(v) Para todos $x, y \in (0,1)$ se verifica:

$$\frac{f'(x)}{f'T(x,y)} + k\frac{g'(x)}{g'S(x,y)} = 1 \text{ si } x < y,$$

$$\frac{f'(x)}{f'T(x,y)}+k\frac{g'(x)}{g'S(x,y)}=k \text{ si } x>y.$$

(vi) f' *y* g' *son estrictamente crecientes en* $(0,1)$.

(vii) $f'_-(1)<0$ *y* $g'_+(0)>0$.

Proposición 10. *Si* (f,g) *es una solución de (9), consideremos la función* $\varphi(x)=1-\frac{f'_-(1)}{f'(x)}$, $x\in[0,1]$. *Entonces:*

(i) φ *es estrictamente decreciente y continua, con* $\varphi(0)=1$, $\varphi(1)=0$.

(ii) $\sigma(x)\sigma(y)=k\sigma S(x,y)\ \forall x,y$, *siendo* $\sigma(t)=\varphi(t)-1+k$, $t\in[0,1]$.

Demostración. Veamos el apartado (ii). Se tiene:

$$\begin{aligned}
\varphi(x) &= 1-\tfrac{f'_-(1)}{f'(x)}\\
&= 1-\lim_{z\to1^-}\tfrac{f'(z)}{f'T(x,z)}\\
&= 1-\lim_{z\to1^-}\left(k-k\tfrac{g'(z)}{g'S(x,z)}\right)\\
&= 1-k+k\lim_{z\to1^-}\tfrac{g'(z)}{g'S(x,z)}\\
&= 1-k+k\lim_{z\to1^-}\tfrac{g'S(y,z)}{g'S(S(x,y),z)},\ \ 0<y<1\\
\varphi(y) &= 1-\tfrac{f'_-(1)}{f'(y)}\\
&= 1-\lim_{z\to1^-}\tfrac{f'(z)}{f'T(y,z)}\\
&= 1-k+k\lim_{z\to1^-}\tfrac{g'(z)}{g'S(y,z)}
\end{aligned}$$

Así,

$$\begin{aligned}
&\tfrac{1}{k^2}(\varphi(x)-1+k)(\varphi(y)-1+k)\\
&\quad=\lim_{z\to1^-}\tfrac{g'(z)}{g'S(S(x,y),z)}\\
&\quad=\tfrac{1}{k}(\varphi S(x,y)-1+k).
\end{aligned}$$

Luego $(\varphi(x)-1+k)(\varphi(y)-1+k)=k(\varphi S(x,y)-1+k)$ y, finalmente, $\sigma(x)\sigma(y)=k\sigma S(x,y)$. ∎

Proposición 11. *La única solución de la ecuación (8) es el par* $(\mathrm{Min},\mathrm{Max})$.

Demostración. Sabemos por la Proposición 10 que se verifica la igualdad $\sigma(x)\sigma(y)=k\sigma S(x,y)\ \forall x,y\in[0,1]$, siendo $\sigma(t)=\varphi(t)-1+k$, $t\in[0,1]$. Si α es el único valor $0<\alpha<1$ tal que $\sigma(\alpha)=0$, entonces tenemos $k\sigma S(\alpha,\alpha)=0$, lo que implica $S(\alpha,\alpha)=\alpha$, que es una contradicción con el hecho de que S sea arquimediana. Al haber demostrado que la ecuación (8) no puede tener soluciones arquimedianas y, de acuerdo con el apartado (iii) de la Proposición 6, podemos asegurar que la única solución de la ecuación planteada es el par $(\mathrm{Min},\mathrm{Max})$. ∎

5.2 La ecuación $kT+S=k\,\mathrm{Min}+\mathrm{Max}$

Consideremos ahora la ecuación:

$$kT(x,y)+S(x,y)=k\min\{x,y\}+\max\{x,y\},\ \ 0<k<1. \tag{10}$$

El estudio y resolución de esta ecuación se lleva a cabo de forma análoga al caso anterior. Destacamos a continuación solo aquellos resultados que marcan las diferencias.

Proposición 12. *Si* (T,S) *es una solución de (10), entonces* $\max\{\min\{x,y\}+\frac{1}{k}\max\{x,y\}-\frac{1}{k},0\}\leqslant T(x,y)\leqslant S(x,y)\leqslant\min\{\min\{x,y\}+\max\{x,y\},1\}$, $\forall x,y$.

Observaciones 3.

(i) $F_k(x,y)=\max\{\min\{x,y\}+\frac{1}{k}\max\{x,y\}-\frac{1}{k},0\}$ *no es ni cópula ni t-norma. En los casos límite,* $F_0=\mathrm{Min}$ *y* $F_1=W$.

(ii) $G_k(x,y)=\min\{k\min\{x,y\}+\max\{x,y\},1\}$, *no es una t-conorma. En los casos límite,* $G_0=\mathrm{Max}$ *y* $G_1=W^*$.

Proposición 13. *Si* (T,S) *es una solución de (10), entonces:*

(i) S es lipschitziana con constante 1. T es lipschitziana con constante $\frac{1}{k}$.

(ii) T y S son continuas.

Proposición 14. *Si* (T,S) *es una solución arquimediana de (10), entonces:*

(i) T y S tienen que ser estrictas.

(ii) Los generadores aditivos de T y S son funciones convexas.

(ii) T es una cópula y $S=k\,\mathrm{Min}+\mathrm{Max}-kT$.

Demostración. T es lipschtziana con constante $\frac{1}{k}$ (mayor que 1), y si $x\leqslant y\leqslant z$ entonces $T(y,z)-T(x,z)\leqslant y-x$. Por tanto, si T tiene generador aditivo f, entonces $f^{-1}\left(\frac{x+y}{2}\right)\leqslant\frac{f^{-1}(x)+f^{-1}(x)}{2}$ siempre que $0\leqslant x\leqslant y\leqslant 3x$. Esto implica que f es convexo. ∎

Siendo (T,S) una solución arquimediana estricta de (10), esta ecuación puede escribirse ahora a partir de los generadores aditivos f y g de T y S, respectivamente, así:

$$\begin{aligned}
&kf^{-1}(f(x)+f(y))+g^{-1}(g(x)+g(y))\\
&\quad=k\min\{x,y\}+\max\{x,y\},\ \ 0<k<1,
\end{aligned} \tag{11}$$

siendo $f,g\colon[0,1]\to[0,\infty]$ funciones estrictamente monótonas (f decreciente y g creciente), continuas y convexas, con $f(1)=g(0)=0$, $f(0)=g(1)=\infty$.

La proposición que sigue es análoga a la Proposición 9 y solo señalamos el apartado donde aparecen diferencias respecto de lo dicho en la citada proposición.

Proposición 15. *Si* (f,g) *es una solución de (11), entonces:*

$$k\frac{f'(x)}{f'T(x,y)}+\frac{g'(x)}{g'S(x,y)}=k \text{ si } x<y,$$

$$k\frac{f'(x)}{f'T(x,y)}+\frac{g'(x)}{g'S(x,y)}=1 \text{ si } x>y.$$

Proposición 16. *Si (f,g) es una solución de (11), consideremos la función $\psi(x) = 1 - \frac{g'_+(0)}{g'(x)}$, $x \in [0,1]$. Entonces:*

(i) es estrictamente creciente y continua, con $\psi(0) = 0$, $\psi(1) = 1$.

(ii) $\gamma(x)\gamma(y) = k\gamma T(x,y)$ $\forall x,y$, siendo $\gamma(t) = \psi(t) - 1 + k$, $t \in [0,1]$.

Demostración. Veamos el apartado (ii). Se tiene:

$$\begin{aligned}
\psi(x) &= 1 - \tfrac{g'_+(0)}{g'(x)} \\
&= 1 - \lim_{z\to 0^+} \tfrac{g'(z)}{g'S(x,z)} \\
&= 1 - \lim_{z\to 0^+} \left(k - k\tfrac{f'(z)}{f'T(x,z)}\right) \\
&= 1 - k + k\lim_{z\to 0^+} \tfrac{f'(z)}{f'T(x,z)} \\
&= 1 - k + k\lim_{z\to 0^+} \tfrac{f'T(y,z)}{f'T(T(x,y),z)}, \quad 0 < y < 1 \\
\psi(y) &= 1 - \tfrac{g'_+(0)}{g'(y)} \\
&= 1 - \lim_{z\to 0^+} \tfrac{g'(z)}{g'S(y,z)} \\
&= 1 - k + k\lim_{z\to 0^+} \tfrac{f'(z)}{f'T(y,z)}
\end{aligned}$$

Así,

$$\begin{aligned}
&\tfrac{1}{k^2}(\psi(x) - 1 + k)(\psi(y) - 1 + k) \\
&\quad = \lim_{z\to 0^+} \tfrac{f'(z)}{f'T(T(x,y),z)} \\
&\quad = \tfrac{1}{k}(\psi T(x,y) - 1 + k).
\end{aligned}$$

Luego $(\psi(x) - 1 + k)(\psi(y) - 1 + k) = k(\psi T(x,y) - 1 + k)$.

Pero si $\gamma(t) = \psi(t) - 1 + k$, tenemos $\gamma(x)\gamma(y) = k\gamma T(x,y)$.

Proposición 17. *La única solución de la ecuación (10) es el par* $(\mathrm{Min}, \mathrm{Max})$.

Demostración. Sabemos por la Proposición 17 que se verifica la igualdad $\gamma(x)\gamma(y) = k\gamma T(x,y)$, siendo $\gamma(t) = \psi(t) - 1 + k$, $t \in [0,1]$. Si β es el único valor $0 < \beta < 1$ tal que $\gamma(\beta) = 0$, entonces tenemos $k\gamma T(\beta,\beta) = 0$, lo que implica $T(\beta,\beta) = 0$, que es una contradicción con el hecho de que T sea arquimediana. Al haber demostrado que (10) no puede tener soluciones arquimedianas y, de acuerdo con el apartado (iii) de la Proposición 6, podemos asegurar que la única solución de la ecuación planteada es el par $(\mathrm{Min}, \mathrm{Max})$. ∎

Combinando las Proposiciones 11 y 17 obtenemos el siguiente resultado final.

Proposición 18. *Los pares (T,S) solución de la ecuación funcional*

$$T(x,y) + \lambda S(x,y) = \min\{x,y\} + \lambda \max\{x,y\}, \quad \lambda \geqslant 0$$

son:

(i) (Min, S), con S cualquier t-conorma, si $\lambda = 0$,

(ii) (T,S) con T y S soluciones de la ecuación de Frank, si $\lambda = 1$,

(iii) $(\mathrm{Min}, \mathrm{Max})$, si $\lambda \neq 0, 1$.

Observaciones 4.

(i) Podría entenderse que para $\lambda = \infty$ la ecuación tiene por soluciones los pares (T, Max), con T cualquier t-norma.

(ii) Otra forma de escribir la ecuación es, con $0 \leqslant a \leqslant 1$,

$$\begin{aligned}
&(1-a)T(x,y) + aS(x,y) \\
&\quad = (1-a)\min\{x,y\} + a\max\{x,y\},
\end{aligned}$$

o, de forma resumida, $W_a(T,S) = W_a$, siendo W_a el operador OWA de pesos $(1-a, a)$: $W_a(x,y) = (1-a)\min\{x,y\} + a\max\{x,y\}$. Ahora las soluciones vienen dadas por:

1) Si $a = 0$, (Min, S), con S t-conorma cualquiera.

2) Si $a = 1$, (T, Max), con T t-norma cualquiera.

3) Si $a = \frac{1}{2}$, las soluciones de la ecuación de Frank.

4) Si $a \neq 0, \frac{1}{2}, 1$, $(\mathrm{Min}, \mathrm{Max})$.

Agradecimientos. Los autores agradecen el apoyo del Govern Balear, con el proyecto PCTIB2005GC1–07, y el del Ministerio de Educación, con los proyectos MTM2009–10962 y TIN2009-07901.

Referencias

[1] C. Alsina, M.J. Frank, B. Schweizer: *Associative Functions*. World Scientific, Singapore, 2006.

[2] M.J. Frank: Associativity in a class of operations on spaces of distribution functions. *Aequationes Math.* 12, pp. 181-194, 1975.

[3] M.J. Frank: On the simultaneous associativity of $F(x,y)$ and $x + y - F(x,y)$. *Aequationes Math.* 19, pp. 194-226, 1979.

[4] E.P. Klement, R. Mesiar, E. Pap: *Triangular Norms*. Kluwer, Dordrecht, 2000.

[5] G. Mayor: Sobre una classe d'operacions entre conjunts difusos: Operacions d'Agregació. En *Actes III Congrés Català de Lògica Matemàtica*. Barcelona, 1984.

[6] R.B. Nelsen: *An Introduction to Copulas* (2ª ed.). Springer, New York, 2006.

New results on metrics aggregation

Sebastià Massanet, Oscar Valero

Departament de Ciències Matemàtiques i Informàtica, Universitat de les Illes Balears, {s.massanet, o.valero}@uib.es

Abstract

In 1981, J. Borsik and J. Doboš studied the problem of how to combine, by means of a function, a collection (not necessarily finite) of metrics in order to obtain a single one as output ([1]). To this end, they introduced the notion of metric aggregation function and gave a characterization of such functions. Recently, in [12] and [13] G. Mayor and O. Valero have extended, inspired in part by the utility of quasi-metrics in Artificial Intelligence and Computer Science, the original work of Borsik and Doboš to the context of quasi-metrics in such a way that a general description of how to merge through a function a collection of quasi-metrics into a single one has been given. In this paper our aim is threefold. On one hand, we continue the study of the quasi-metric aggregation problem. In particular we characterize the functions that allow to merge a collection of weightable quasi-metrics into a single weightable quasi-metric. On the other hand, we focus our attention on the partial metric aggregation problem. Thus we introduce the notion of partial metric aggregation function, which generalizes the well-known one for metric spaces, and provide a characterization of such a functions. Finally, a few connections between the problem of merging quasi-metrics, metrics and partial metrics are given.

Keywords: Partial metric, quasi-metric, metric, monotone function, (sub)additive function, superadditive function, aggregation function.

1 Introduction and preliminaries

Throughout this paper we shall use the letters $\mathbb{R}$, $\mathbb{R}^+$, and $\mathbb{N}$ to denote the set of real numbers, the set of nonnegative real numbers and the set of positive integer numbers, respectively.

From now on, $\mathbb{R}_I^+$ denotes the set of all nonnegative functions defined on a (nonempty) set I of indexes. Given $\mathbf{x} \in \mathbb{R}_I^+$ we will write x_i instead of $\mathbf{x}(i)$. Moreover, we will denote by $\mathbf{0}$ the element of $\mathbb{R}_I^+$ with $0_i = 0$ for all $i \in I$.

As usual, we will consider the set $\mathbb{R}_I^+$ ordered by the pointwise order relation $\preceq$, i.e. $\mathbf{x} \preceq \mathbf{y} \Leftrightarrow x_i \leq y_i$ for all $i \in I$.

We will say that a function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is monotone provided that $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}_I^+$ with $\mathbf{x} \preceq \mathbf{y}$. Furthermore, a function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ wil be said to be subadditive (superadditive) if $\Phi(\mathbf{x}+\mathbf{y}) \leq \Phi(\mathbf{x}) + \Phi(\mathbf{y})$ ($\Phi(\mathbf{x}) + \Phi(\mathbf{y}) \leq \Phi(\mathbf{x}+\mathbf{y})$) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}_I^+$.

In 1981, J. Borsik and J. Doboš studied in depth the problem of how to combine by means of a function a collection of metrics in order to obtain a single one as a result ([1]). Such functions were called metric preserving functions (metric aggregation functions in [12] and [13]). Concretely, a function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a metric aggregation function if the function $M_{d,\Phi} : X \times X \to \mathbb{R}^+$ is a metric for every indexed family of metric spaces $\{(X_i, d_i)\}_{i \in I}$, where $X = \prod_{i \in I} X_i$ and $M_{d,\Phi}(\mathbf{x}, \mathbf{y}) = \Phi((d_i(x_i, y_i))_{i \in I})$ for all $\mathbf{x}, \mathbf{y} \in X$.

According to [1], we will denote by $\mathscr{O}_I$ the set of all functions $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ satisfying: $\Phi(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

Borsik and Doboš gave a connection between metric aggregation functions and subadditive and monotone ones. In particular they proved the below result.

Proposition 1. *Let $\Phi \in \mathscr{O}_I$. If Φ is monotone and subadditive, then Φ is a metric aggregation function.*

Unfortunately, the preceding proposition does not characterizes metric aggregation functions because there are metric aggregation functions which are not monotone (see Example 14 in Section 4). Of course Borsik and Doboš proved a characterization of metric aggregation functions in terms of the so-called triangle triplets. In order to present such a characterization let us recall that, given $a, b, c \in \mathbb{R}^+$, the

triplet (a,b,c) forms a triangle triplet whenever $a \leq b+c$, $b \leq a+c$ and $c \leq b+a$. The aforementioned characterization can be enunciated as follows:

Theorem 2. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$. Then the below assertions are equivalent:*

(1) Φ is a metric aggregation function.

(2) Φ holds the following properties:

(2.1) $\Phi \in \mathscr{O}_I$.

(2.2) Let $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbb{R}_I^+$. If (a_i,b_i,c_i) is a triangle triplet for all $i \in I$, then so is $(\Phi(\mathbf{a}),\Phi(\mathbf{b}),\Phi(\mathbf{c}))$.

Since Borsik and Doboš solved the problem of merging a collection of metrics, several authors have provided new advances in the study of the aggregation problem of a few kinds of generalized metrics. Specifically E. Castiñeira, A. Pradera and E. Trillas have provided a solution to the aggregation of C-generalized metrics, S-generalized distances and pseudometrics in [17], [18] and [16], and several general techniques for merging a finite number of metrics into another one have been explored by J. Casasnovas and F. Rosselló in [2].

Inspired by the original work of Borsik and Doboš and motivated by the utility of quasi-metrics in several fields of Artificial Intelligence and Computer Science (see, for instance, [19], [20], [21], [5], [4], [22], [14], [15], [23], [6]), G. Mayor and O. Valero studied the aggregation problem in the quasi-metric context. Thus they introduced the notion of quasi-metric aggregation function in [13] (asymmetric distance function in the quoted reference) and gave a characterization of such functions in the spirit of Theorem 2. With the aim to present such a characterization let us recall a few concepts about quasi-metrics.

Following [7], a quasi-metric (asymmetric distance in [12] and [13]) on a (nonempty) set X is a function $d : X \times X \to \mathbb{R}^+$ such that for all $x,y,z \in X$:

(i) $d(x,y) = d(y,x) = 0 \Leftrightarrow x = y$.
(ii) $d(x,z) \leq d(x,y) + d(y,z)$.

Note that a metric on a set X is a quasi-metric d on X satisfying, in addition, the following condition for all $x,y \in X$:

(iii) $d(x,y) = d(y,x)$.

A quasi-metric space (asymmetric distance space in [12] and [13]) is a pair (X,d) such that X is a (nonempty) set and d is a quasi-metric on X.

If d is quasi-metric on X, then the function $d^s : X \times X \to \mathbb{R}^+$, defined by $d^s(x,y) = \max\{d(x,y),d(y,x)\}$ for all $x,y \in X$, is a metric.

According to [12] and [13], a function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a quasi-metric aggregation function if the function $Q_{d,\Phi} : X \times X \to \mathbb{R}^+$ is a quasi-metric for every indexed family of quasi-metric spaces $\{(X_i,d_i)\}_{i\in I}$, where $X = \prod_{i\in I} X_i$ and $Q_{d,\Phi}(\mathbf{x},\mathbf{y}) = \Phi((d_i(x_i,y_i))_{i\in I})$ for all $\mathbf{x},\mathbf{y} \in X$.

The quasi-metric formulation of Theorem 2 given in [13] can be stated in the following way:

Theorem 3. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$. Then the below assertions are equivalent:*

(1) Φ is a quasi-metric aggregation function.

(2) Φ holds the following properties:

(2.1) $\Phi \in \mathscr{O}_I$.

(2.2) Let $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbb{R}_I^+$. If $\mathbf{a} \preceq \mathbf{b}+\mathbf{c}$, then $\Phi(\mathbf{a}) \leq \Phi(\mathbf{b}) + \Phi(\mathbf{c})$.

(3) $\Phi \in \mathscr{O}_I$, and Φ is subadditive and monotone.

Notice that, contrary to the metric case, the preceding characterization states that every quasi-metric aggregation function is always a monotone function.

Since every quasi-metric d on a (nonempty) set X induces on X the metric d^s, Mayor and Valero gave a result which establishes a relationship between the metric induced by the quasi-metric obtained via the aggregation of an indexed family of quasi-metrics and the aggregation of the induced metrics by the aforesaid quasi-metrics. In particular it was proved the following.

Proposition 4. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a quasi-metric aggregation function and let $\{(X_i,d_i)\}_{i\in I}$ be an indexed family of quasi-metric spaces. Then $(Q_{d,\Phi})^s \leq M_{d^s,\Phi} \leq 2(Q_{d,\Phi})^s$.*

In [13], [9] and [8], Theorem 3 played a fundamental role in order to establish the aggregation theory of quasi-metrics as a possible basis for the complexity analysis of programs in Computer Science. Moreover, several connections between the aggregation of quasi-metrics and the theory of computational complexity have been obtained by Casasnovas and Valero in [3] and by P. Tirado and Valero in [24].

In 1994, S.G. Mathews introduced new metric notions, which are known as weightable quasi-metrics and partial metrics, in order to provide a mathematical framework to model several computational processes that arise in a natural way in denotational semantics for programming languages and in parallel processing (see [10] and [11] and, also, see [6] for additional applications to logic programming).

According to Matthews a quasi-metric d on a (nonempty) set X is weightable provided that there exists a function $w : X \to \mathbb{R}^+$, called the weight associated to d, such that

$$d(x,y) + w(x) = d(y,x) + w(y)$$

for all $x,y \in X$. Of course a weightable quasi-metric space is a triple (X,d,w) such that (X,d) is a weightable quasi-metric on X and w is the weight associated to d.

A partial metric on a (nonempty) set X is a function $p : X \times X \to \mathbb{R}^+$ such that for all $x, y, z \in X$:

(i) $p(x,x) = p(x,y) = p(y,y) \Leftrightarrow x = y.$
(ii) $p(x,x) \leq p(x,y).$
(iii) $p(x,y) = p(y,x).$
(iv) $p(x,z) \leq p(x,y) + p(y,z) - p(y,y).$

Of course, a partial metric space is a pair (X,p) such that X is a (nonempty) set and p is a partial metric on X. Clearly, and similarly to the quasi-metric case, the notion of metric can be also retrieved as a particular case of the partial metric one. In particular, a metric on a set X is a partial metric p on X such that $p(x,x) = 0$ for all $x \in X$.

Inspired, on one hand, by the fact that the information fusion methods based on the the use of numerical aggregation functions play a central role in applied sciences and that the numerical values to be merged come in many cases from distances between several pieces of information and, on the other hand, by the fact that weightable quasi-metrics and partial metrics have been shown to be appropriate tools to model many processes that arise in a natural way in Artificial intelligence and Computer Science (as we have pointed out above), in this paper we study the aggregation problem for the case of the aforenamed generalized metrics. Thus we introduce the notion of weightable quasi-metric aggregation function and partial metric aggregation function, and provide characterizations of such functions. Moreover, we give several connections between the problem of merging quasi-metrics, metrics and partial metrics.

2 Aggregation of weightable quasi-metrics

Although the quasi-metric aggregation problem was solved completely via Theorem 3 in [13] and [12] , it remains open the same problem in the particular case of weightable quasi-metrics. For this reason, this section is devoted to analyze what additional properties a quasi-metric aggregation function must hold in order to guarantee that a collection of weightable quasi-metrics is merged into a new single quasi-metric which is weightable too. With this aim, we need to introduce formally this new sort of aggregation functions.

Definition 5. A quasi-metric aggregation function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a weightable quasi-metric aggregation function provided that for every indexed family of weightable quasi-metric spaces $\{(X_i, d_i, w_i)\}_{i\in I}$, the function $Q_{d,\Phi} : X \times X \to \mathbb{R}^+$ is a weightable quasi-metric with associated weight function $W_{w,\Phi} : X \to \mathbb{R}^+$ where

$$W_{w,\Phi}(\mathbf{x}) = \Phi((w_i(x_i))_{i\in I})$$

for all $x \in X = \prod_{i\in I} X_i$.

The next result determines when the aggregation of an indexed family of weightable quasi-metrics is, in fact, a weightable quasi-metric.

Theorem 6. *Let $\{(X_i, d_i, w_i)\}_{i\in I}$ be an indexed family of weightable quasi-metric spaces. Then a quasi-metric aggregation function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a weightable quasi-metric aggregation function if and only if it is additive.*

In the light of Theorem 3 and 6 we obtain the next nice characterization of weightable quasi-metric aggregation functions.

Corollary 7. *A function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a weightable quasi-metric aggregation function if and only if it is additive, monotone and $\Phi \in \mathscr{O}_I$.*

3 Aggregation of partial metrics

In this section we tackle the problem of how to combine a collection (not necessarily finite) of partial metrics in order to obtain a single one as a result. For our below dissertation we introduce the notion of partial metric aggregation function.

Definition 8. A function $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a partial metric aggregation function provided that the function $P_{p,\Phi} : X \times X \to \mathbb{R}^+$ is a partial metric for every indexed family of partial metric spaces $\{(X_i, p_i)\}_{i\in I}$, where $X = \prod_{i\in I} X_i$ and for all $x, y \in X$

$$P_{p,\Phi}(x,y) = \Phi((p_i(x_i, y_i))_{i\in I}).$$

The next result turns out to be very useful to obtain the subsequent characterization of partial metric aggregation functions which solves the stated problem.

Proposition 9. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$. If Φ is a partial metric aggregation function, then the following statements hold:*

(1) Φ is monotone.

(2) If there exists $\mathbf{x} \in \mathbb{R}_I^+$ such that $\Phi(\mathbf{x}) = 0$, then $\mathbf{x} = \mathbf{0}$.

Theorem 10. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$. Then Φ is a partial metric aggregation function if and only if it satisfies the following two properties for all $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d} \in \mathbb{R}_I^+$:*

(1) $\Phi(\boldsymbol{a}) + \Phi(\boldsymbol{b}) \leq \Phi(\boldsymbol{c}) + \Phi(\boldsymbol{d})$ whenever $\boldsymbol{a} + \boldsymbol{b} \preceq \boldsymbol{c} + \boldsymbol{d}$, $\boldsymbol{b} \preceq \boldsymbol{c}$ and $\boldsymbol{b} \preceq \boldsymbol{d}$.

(2) If $\boldsymbol{b} \preceq \boldsymbol{a}$, $\boldsymbol{c} \preceq \boldsymbol{a}$ and $\Phi(\boldsymbol{a}) = \Phi(\boldsymbol{b}) = \Phi(\boldsymbol{c})$, then $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{c}$.

As a consequence of statement (1) in Theorem 10 and statement (2) in Proposition 9, the following property of partial metric aggregation functions holds.

Corollary 11. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$. If Φ is a partial metric aggregation function, then Φ is subadditive.*

The next is an example of partial metric aggregation function which will be useful later on.

Example 12. Let $I = \{1\}$. Define $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ by $\Phi(x) = x + 1$ for all $x \in \mathbb{R}^+$. It is clear that the function Φ holds conditions (1) and (2) in Theorem 10 and, thus, it is a partial metric aggregation function.

4 Some relationships between the problem of merging quasi-metrics, metrics and partial metrics

In this section we present a few connections between metrics, partial metrics and quasi-metrics by means of aggregation functions.

It seems natural to wonder what is the relationship between partial metric aggregation functions and metric and quasi-metric aggregation functions.

From Theorem 2, Theorem 3, Proposition 9 and Corollary 11 we obtain the below result which provides some information on the preceding question.

Proposition 13. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a partial metric aggregation function such that $\Phi_I(\mathbf{0}) = 0$. Then Φ is a quasi-metric aggregation function and, thus, a metric aggregation function.*

Example 12 shows that the condition $\Phi_I(\mathbf{0}) = 0$ can not be deleted in the statement of Proposition 13. Concretely it shows that there are partial metric aggregation functions which are not either metric aggregation functions or quasi-metric aggregation functions. Indeed, the function Φ defined in the aforenamed example is a partial aggregation function. However $\Phi(0) = 1$ and, thus, $\Phi \notin \mathscr{O}_I$. Then Theorem 2 guarantees that Φ is not a metric aggregation function and Theorem 3 provides that Φ is not a quasi-metric aggregation function.

In the next examples we provide a negative answer to the question about whether every metric aggregation function and every quasi-metric aggregation function is a partial metric aggregation function.

Example 14. Let $I = \mathbb{N}$. Consider the function $\Phi : \mathbb{R}_\mathbb{N}^+ \to \mathbb{R}^+$ given by $\Phi(\mathbf{0}) = 0$ and

$$\Phi(\mathbf{x}) = \begin{cases} 2 & \alpha(\mathbf{x}) \in]0,1[\\ 1 & \alpha(\mathbf{x}) \geq 1 \end{cases},$$

where $\mathbf{x} \neq \mathbf{0}$ and $\alpha(\mathbf{x})$ denotes the value of the first term of the sequence $\mathbf{x}$ different to $\mathbf{0}$. Clearly $\Phi \in \mathscr{O}_\mathbb{N}$. Furthermore, it is easily seen that Φ turns triangle triplets into triangle triplets. Thus, by Theorem 2, Φ is a distance aggregation function. However, by Proposition 9, Φ is not a partial metric aggregation function because Φ is not monotone. Indeed, let $\mathbf{x}, \mathbf{y} \in \mathbb{R}_\mathbb{N}$ such that $x_i = \frac{1}{2}$ and $y_i = 1$ for all $i \in \mathbb{N}$. Then $\Phi(\mathbf{x}) = 2$, $\Phi(\mathbf{y}) = 1$ and $\mathbf{x} \preceq \mathbf{y}$.

Example 15. Let $I = \mathbb{N}$. Consider the function $\Phi : \mathbb{R}_\mathbb{N}^+ \to \mathbb{R}^+$ defined by

$$\Phi(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} = \mathbf{0} \\ 1 & \text{otherwise.} \end{cases}$$

Since Φ is monotone, subadditive and $\Phi \in \mathscr{O}_\mathbb{N}$ we have, by Theorem 3, that it is a quasi-metric aggregation function but it is not a partial metric aggregation function. Indeed, let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}_\mathbb{N}^+$ such that $\mathbf{a}_i = 2$ for all $i \in \mathbb{N}$, $\mathbf{b}_1 = 0$, $\mathbf{c}_1 = 1$ and $\mathbf{b}_i = 1$ and $\mathbf{c}_i = 0$ for all $i \in \mathbb{N}$ with $i \geq 2$. Then $\mathbf{b} \preceq \mathbf{a}$, $\mathbf{c} \preceq \mathbf{a}$ and $\Phi(\mathbf{a}) = \Phi(\mathbf{b}) = \Phi(\mathbf{c}) = 1$ but $\mathbf{b} \neq \mathbf{a}$. So, by statement (2) in Theorem 10, Φ is not a partial metric aggregation function.

Of course every weightable quasi-metric aggregation function is a quasi-metric, and thus a metric, aggregation function. By Corollary 7, Example 14 and Example 15 show that the converse is not true.

Next we provide the existing relationship between weightable quasi-metric aggregation functions and the partial metric ones.

Proposition 16. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a weightable quasi-metric aggregation function. Then it is a partial metric aggregation function.*

Example 12 shows that the converse of the preceding result is not true. Indeed, the function Φ introduced in Example 12 is a partial metric aggregation function and, however, it is not additive. So by Corollary 7, Φ is not a weightable quasi-metric aggregation function.

In [10] Matthews proved the following interesting correspondences between metrics, quasi-metrics and partial metrics.

Proposition 17. *Let (X, p) be a partial metric space. Then the function $d_p : X \times X \to \mathbb{R}^+$, defined by $d_p(x,y) = p(x,y) - p(x,x)$ for all $x, y \in X$, is a weightable quasi-metric on X with weight $w : X \to \mathbb{R}^+$ given by $w(x) = p(x,x)$ for all $x \in X$.*

Proposition 18. *Let (X, d, w) be a weightable quasi-metric space. Then the function $p_d : X \times X \to \mathbb{R}^+$, defined by $p_d(x,y) = d(x,y) + w(x)$ for all $x, y \in X$, is a partial metric on X.*

Proposition 19. *Let (X, p) be a partial metric space. Then the function $p^s : X \times X \to \mathbb{R}^+$, defined by $p^s(x,y) = 2p(x,y) - p(x,x) - p(y,y)$ for all $x, y \in X$, is a metric on X.*

Since every partial metric induces a weightable quasi-metric (Proposition 17) it seems natural to wonder what is the relationship between the aggregation of the weightable quasi-metrics induced by an indexed family of partial metrics and the weightable quasi-metric induced by the partial metric obtained via the aggregation of the aforesaid partial metrics. The next result provides the answer to the formulated question.

Proposition 20. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a weightable quasi-metric aggregation function and let $\{(X_i, p_i)\}_{i\in I}$ be an indexed family of partial metric spaces. Then $d_{P_{p,\Phi}} = Q_{d_p,\Phi}$.*

Since every weightable quasi-metric induces a partial metric (Proposition 18) it seems natural to wonder what is the relationship between the aggregation of the partial metrics induced by an indexed family of weightable quasi-metrics and the partial metric induced by the weightable quasi-metric obtained via the aggregation of the aforesaid weightable quasi-metrics. The next result provides the answer to this new question.

Proposition 21. *Let $\{(X_i, d_i, w_i)\}_{i\in I}$ be an indexed family of weightable quasi-metric spaces. If $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ is a weightable quasi-metric aggregation function, then $P_{p_d,\Phi} = p_{Q_{d,\phi}}$.*

Motivated by the fact that every partial metric induces a metric (Proposition 19), we discuss in the below result the relationship between the metric induced by the aggregation of a indexed family of partial metrics and the aggregation of the metrics generated by the aforementioned partial metrics.

Proposition 22. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a partial metric aggregation function such that $\Phi_I(\mathbf{0}) = 0$ and let $\{(X_i, p_i)\}_{i\in I}$ be an indexed family of partial metric spaces. Then the following statements hold:*

1) *$M_{p^s,\Phi} \leq (P_{p,\Phi})^s$ whenever $\Phi(\boldsymbol{a}-\boldsymbol{b}) \leq \Phi(\boldsymbol{a}) - \Phi(\boldsymbol{b})$ for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}_I^+$ with $\boldsymbol{b} \preceq \boldsymbol{a}$, where for all $\boldsymbol{x},\boldsymbol{y} \in X$.*
2) *If Φ is a superadditive function, then $(P_{p,\Phi})^s \leq M_{p^s,\Phi}$.*
3) *$(P_{p,\Phi})^s = M_{p^s,\Phi} \Leftrightarrow \Phi(2\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}) = 2\Phi(\boldsymbol{a}) - \Phi(\boldsymbol{b}) - \Phi(\boldsymbol{c})$ for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c} \in \mathbb{R}_I^+$ with $\boldsymbol{b} \preceq \boldsymbol{a}$ and $\boldsymbol{c} \preceq \boldsymbol{a}$.*

Finally, mixing the results presented in Propositions 4, 20 and 22, we present some connections, via aggregation functions, between the metrics induced by both, an indexed family of partial metrics and their associated weightable quasi-metrics.

Proposition 23. *Let $\Phi : \mathbb{R}_I^+ \to \mathbb{R}^+$ be a partial metric aggregation function such that $\Phi_I(\mathbf{0}) = 0$ and let $\{(X_i, p_i)\}_{i\in I}$ be an indexed family of partial metric spaces. The below statements hold:*

1) *$(d_{P_{p,\Phi}})^s \leq (P_{p,\Phi})^s \leq 2(d_{P_{p,\Phi}})^s$.*
2) *$(Q_{d_p,\Phi})^s \leq M_{d_p^s,\Phi} \leq 2(Q_{d_p,\Phi})^s$.*
3) *$(Q_{d_p,\Phi})^s \leq M_{d_p^s,\Phi} \leq M_{p^s,\Phi} \leq 2(Q_{d_p,\Phi})^s \leq 2M_{d_p^s,\Phi} \leq 2M_{p^s,\Phi}$.*
4) *$(Q_{d_p,\Phi})^s \leq M_{d_p^s,\Phi} \leq M_{p^s,\Phi} \leq (P_{p,\Phi})^s \leq 2(d_{P_{p,\Phi}})^s$, whenever $\Phi(\boldsymbol{a}-\boldsymbol{b}) \leq \Phi(\boldsymbol{a}) - \Phi(\boldsymbol{b})$ for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}_I^+$ with $\boldsymbol{b} \preceq \boldsymbol{a}$.*
5) *$(d_{P_\Phi})^s \leq M_{p^s,\Phi} \leq (P_{p,\Phi})^s \leq 2(Q_{d_p,\Phi})^s \leq 2M_{d_p^s,\Phi} \leq 2M_{p^s,\Phi}$, whenever Φ is superadditive.*

Acknowledgements

The authors acknowledge the support of the Spanish Ministry of Science and Innovation grants MTM2009-10962 and MTM2009-10320.

References

[1] J. Borsik, J. Doboš: On a product of metric spaces. *Math. Slovaca* 31, pp. 193-205, 1981.

[2] J. Casasnovas, F. Roselló: Averaging fuzzy biopolymers. *Fuzzy Set. Syst.* 152, pp. 139-158, 2005.

[3] J. Casasnovas, O. Valero: A connection between computer science and fuzzy theory: midpoints and and running time o f computing. *Mathware Soft Comput.* 15, pp. 251-261, 2008.

[4] L.M. García-Raffi, S. Romaguera, E.A. Sánchez-Pérez: The supremum asymmetric norm on sequence algebras: a general framework to measure complexity spaces. *Electron. Notes Theor. Comput. Sci.* 74 , pp.39-50, 2003.

[5] L.M. García-Raffi, S. Romaguera, E.A. Sánchez-Pérez: Sequence spaces and asymmetric norms in the theory of computational complexity. *Math. Comput. Model.* 36, pp. 1-11, 2002.

[6] P. Hitzler, A.K. Seda: Generalized distance functions in the theory of computation. *The Computer Journal* 53, pp 443-464, 2010.

[7] H.P.A. Künzi: Nonsymmetric distances and their associated topologies: About the origins of basic ideas in the area of asymmetric topology. In C.E. Aull and R. Lowen (Eds.), *Handbook of the History of General Topology*, vol. 3, Kluwer Acad. Publ., Dordrecht (The Netherlands), pp. 853-968, 2001.

[8] J. Martin, G. Mayor, O. Valero: A fixed point theorem for asymmetric distances via aggregation functions. In *Proc. of the 6th International Summer School on Aggregation Operators*, Benevento (Italy), pp. 217-222, 2011.

[9] J. Martin, G. Mayor, O. Valero: On aggregation of normed structures. *Math. Comput. Model.* 54, pp. 815-827, 2011.

[10] S.G. Matthews: Partial metric topology. *Ann. New York Acad. Sci.* 728, pp. 183-197, 1994.

[11] S.G. Matthews: An extensional treatment of lazy data flow dedlock. *Theoret. Comput. Sci.* 151, pp. 195-205, 1995.

[12] G. Mayor, O.Valero: Aggregating asymmetric distances in Computer Sciences. In *Proc. of the 8th International FLINS Conference on Computational Intelligence in Decision and Control*, Madrid (Spain), pp. 477-482, 2008.

[13] G. Mayor, O. Valero: Aggregation of asymmetric distances in computer science. *Inform. Sci.* 180, pp. 803-812, 2010.

[14] V. Pestov, A. Stojmirović: Indexing schemes for similarity search: an illustrated paradigm. *Fund. Inform.* 70, pp. 367-385, 2006.

[15] V. Pestov, A. Stojmirović: Indexing schemes for similarity search in datasets of short protein fragments *Inform. Syst.* 32, pp. 1145-1165, 2007.

[16] A. Pradera, E. Trillas, E. Castiñeira: On the aggregation of some classes of fuzzy relations. In: B. Bouchon-Meunier, J. Gutierrez, L. Magdalena and R. Yager (Eds.), *Technologies for constructing intelligent systems*, Springer-Verlag, Heidelberg, pp. 125-147, 2002.

[17] A. Pradera, E. Trillas, E. Castiñeira: On distances Aggregation. In *Proc. Information Processing and Management of Uncertainty in Knowledge-based Systems international Conference*, Dortmund (Germany), pp. 693-700, 2000.

[18] A. Pradera, E. Trillas: A note on pseudometrics aggregation. *Int. J. Gen. Syst.* 31, pp. 41-51, 2002.

[19] M. Schellekens: The Smyth completion: a common foundation for denonational semantics and complexity analysis. *Electron. Notes Theor. Comput. Sci.* 1, pp 211-232, 1995.

[20] A. K. Seda: Quasi-metrics and the semantics of logic programs. *Fund. Inform.* 29, pp. 97-117, 1997.

[21] A. K. Seda: Some Issues Concerning Fixed Points in Computational Logic: Quasi-Metrics, Multivalued Mappings and the Knaster-Tarski Theorem. *Topology Proc.* 24, pp. 223-250, 1999.

[22] A. Stojmirović: Quasi-metrics, similarities and searches: aspects of geometry of protein datasets. *Ph.D. Thesis*, Victoria University of Wellington, Wellington, 2005.

[23] A. Stojmirović, Y.-K. Yu: Geometric aspects of biological sequence comparison. *J. Comput Biol.* 16, pp. 579-610, 2009.

[24] P. Tirado, O. Valero: The average running time of an algorithm as a midpoint between fuzzy sets. *Math. Comput. Model.* 49, pp. 1852-1868, 2009.

Algunas Propiedades de Consistencia de las Familias de Operadores de Agregación

Karina Rojas[1], **Daniel Gómez**[2], **J.Tinguaro Rodríguez**[3], **Javier Montero**[4]

[1]*krpatuelli@yahoo.com*,
[2]*dagomez@estad.ucm.es*,
[3,4]*jtrodrig, monty@mat.ucm.es*,

Resumen

Diversos estudios se han centrado en analizar amplia y profundamente distintas propiedades de los operadores de agregación. Sin embargo, no ha ocurrido lo mismo en el estudio de propiedades relativas a las *Familias de Operadores de Agregación*. ¿Debiera haber una relación entre los miembros de una familia de operadores de agregación?, ¿es posible agregar n elementos de una base de datos con operadores de agregación de menor dimensión?, ¿debería haber algún tipo de consistencia interna en una familia de operadores de agregación?. En este trabajo se analiza la estabilidad para una familia de operadores de agregación, a fin de asegurar la consistencia entre operadores de una misma familia, y a través de esto, la robustez del proceso de agregación. La propiedad de *Estabilidad* que aquí se propone, intenta forzar a que una familia de operadores de agregación tenga una definición estable/contínua en términos de robustez de los resultados, para lo cual se condiciona que la agregación de n elementos no debe ser muy distinta a la de $n-1$ elementos, cuando un ítem es el valor agregado de los otros. Se presentan algunas definiciones y resultados siguiendo esta idea para distintos niveles de estabilidad y dos tipos de estructuras en un proceso de agregación lineal.

Palabras Clave: Agregación, Operadores, Familia de Funciones, Funciones Extendidas, Conjuntos Borrosos.

1 INTRODUCCIÓN

La agregación de la información es un camino natural en las áreas del conocimiento basadas en sistemas (véase [2, 4, 6, 11]). Usualmente, el proceso de agregación/fusión produce una reducción de la cardinalidad de la base de datos original, en tanto el principal objetivo de la agregación es simplificar la información a fin de facilitar la toma de desiciones.

Un operador de agregación es una función real de n valores que es utilizada para reducir la dimensión de la base de datos original, y la definición estándar de una familia de operadores de agregación bajo un contexto borroso es $\{A_n : [0,1]^n \to [0,1], n\varepsilon\mathbb{N}\}$, para cualquier tamaño n de la base de datos.

Hasta la fecha, se han estudiado diversas propiedades relativas a las funciones de agregación, tales como continuidad, conmutatividad, monoticidad, asociatividad, entre otras. Estas propiedades evidencian algunas características deseables que están relacionadas con cada función de agregación, pero no entregan ninguna información respecto de la consistencia de las familias de operadores de agregación, en términos de la relación que debiera existir entre sus miembros, tal como ha sido señalado por algunos autores (véase [1, 5, 7, 8, 9, 12]).

En la practica, es frecuente que se produzcan cambios de cardinalidad en las bases de datos a lo largo de un proceso de agregación, y cada vez que esto ocurre, se requiere un nuevo operador A_m que agregue la nueva colección de m elementos, no existiendo necesariamente una relación entre los operadores de agregación A_n y A_m. Es en este contexto, bajo cambios de cardinalidad, cuando parece ser natural incorporar alguna propiedad que mantenga la consistencia lógica entre estos operadores, para lo cual se requiere una definición de familia de operadores de agregación en términos de su consistencia lógica, independientemente de la cardinalidad de la base de datos.

Teniendo en cuenta que una secuencia del tipo $A(2)$, $A(3)$, $A(4), \ldots$, está presente a lo largo de cualquier proceso de agregación, parece lógico estudiar propiedades que otorguen sentido a esta secuencia a través de un concepto unificador que esté a la base de cada una de dichas funciones de agregación, ya que de otra forma, podrían ser funciones

desconectadas, como por ejemplo, una definición de agregación que entregue el valor Mínimo para $n = 2$, la Media Aritmética para $n = 3$, la Media Geométrica para $n = 4$ y la Mediana para $n = 5$. Si bien, pareciera que una formulación formal podría resolver este problema al exigir una unidad conceptual mediante una fórmula matemática, este último ejemplo actúa como contra ejemplo, en tanto podría formalizarse matemáticamente para estos primeros cinco casos.

Por lo tanto, tal como es mostrado en [12], una vez asumida la importancia de considerar operadores de agregación como parte de una misma famlia, se requiere indagar respecto de alguna idea de consistencia interna. Es posible encontrar algunas referencias en esta línea, por ejemplo las reglas recursivas (véase en [4]), que recogen la idea de que es posible construir un operador de agregación A_n de manera secuencial. La clave de la recursividad es que para que una regla de agregación sea consistente, debe basarse en una aplicación iterativa de operadores binarios en base a agregaciones anteriores. Esta idea también es estudiada en [1, 12], en la cual las reglas recursivas son generalizadas de un modo más flexible. No obstante lo anterior, la recursividad no es la única propiedad de consistencia de una familia de operadores. En [17, 18], es analizado el concepto de operatividad desde un punto de vista computacional, en el que se intenta capturar la consistencia en la implementación de una función de agregación.

En este trabajo, se estudia una propiedad relacionada con la consistencia de una familia de operadores de agregación; la *Estabilidad*. La estabilidad de una familia de operadores de agregación intenta forzar a que una familia tenga una definición estable/continua, en el sentido que la agregación de n ítems debería ser similar a la agregación de $n-1$ ítems cuando un ítem es la agregación de los otros. La *Estabilidad* puede ser considerada como una generalización de la propiedad *Self-Identity* de *Yager* (1997), en tanto esta última es definida sólo para una función de agregación bajo un proceso lineal por la derecha, mientras que en este trabajo se presentan distintos niveles de estabilidad para una familia de operadores de agregación con dos tipos de estructuras lineales, por la derecha y por la izquierda.

En nuestra opinión, la idea y definición de consistencia de una familia de operadores de agregación debe tener en cuenta la estructura de la base de datos. Se pueden encontrar algunas ideas preliminares en [21], en la cual se considera la estructura de un conjunto de clases borrosas para distinguir entre algunas situaciones aparentemente similares pero que tienen distintas estructuras subyacentes (como los intervalo-valorados borrosos y los cojuntos borrosos intuicionistas). En otras áreas, como por ejemplo en la teoría de la probabilidad, es evidente que el proceso de estimación es dependiente de la estructura de la base de datos que va a ser agregada. Por ejemplo, a fin de pronosticar el clima de mañana, es más relevante la información del clima de hoy que la de ayer. Estas situaciones son usualmente representadas con una estructura lineal (como en series temporales) o por medio de gráficos (como en cadenas de Markov).

2 ESTABILIDAD DE UNA FAMILIA DE OPERADORES DE AGREGACIÓN

Como se menciona en [4], la estabilidad de cualquier modelo matemático para problemas de ingeniería hace alusión a que "pequeños errores de entrada" no dan lugar a "grandes errores de salida". La estabilidad de una función de agregación es definida en el mismo sentido que la continuidad, en tanto pequeños cambios en el vector x no deberían producir grandes cambios en $A(x)$. La estabilidad es definida como: Si máx $|x-y|$ es pequeño, entonces $|A(x)-A(y)|$ también lo es.

Tal como ha sido mencionado en la introducción de este documento, la estabilidad sólo ha sido definida para funciones de agregación, no se ha encontrado en la literatura una definición que persiga esta idea de estabilidad para toda una familia de operadores de agregación. Y como esto está relacionado con la robustez del proceso de agregación, la noción de estabilidad para una familia de operadores de agregación que se propone, está inspirado en la continuidad de funciones, pero se centra en la cardinalidad de la base de datos más que en los datos en si mismos. En términos generales, se entiende por *Estabilidad* de una familia de operadores de agregación a pequeños cambios en el valor agregado cuando se produce un cambio de cardinalidad en la base de datos.

DEFINICION 1: Sea $\{A_n : [0,1]^n \to [0,1], n \in \mathbb{N}\}$ una familia de operadores de agregación. Entonces:

1. A_n cumple con la propiedad de *Estabilidad Estricta por la Derecha* si $\forall x_1, \ldots x_{n-1} \in [0,1]$, se satisface lo siguiente:

   $$A_n(x_1, \ldots x_{n-1}, A_{n-1}(x_1, \ldots, x_{n-1})) = A_{n-1}(x_1, \ldots, x_{n-1})$$

2. A_n cumple con la propiedad de *Estabilidad Estricta por la Izquierda* si $\forall x_1, \ldots x_{n-1} \in [0,1]$ se satisface lo siguiente:

   $$A_n(A_{n-1}(x_1, \ldots, x_{n-1}), x_1, \ldots x_{n-1}) = A_{n-1}(x_1, \ldots, x_{n-1})$$

3. A_n cumple con la propiedad de *Estabilidad Estricta* si se satisfacen los dos puntos anteriores.

Observemos que el primer punto de la Definición 1 fue definida en [24] para una función de agregación específica llamada función de agregación *Self-Identity*. No obstante, nótese que la propiedad descrita en Definición 1.1 es definida para una familia completa y no para una función en particular. En este sentido, la *Estabilidad Estricta por la*

Derecha puede ser considerada como una generalización de la propiedad dada en [24], a la que se le añade la *Estabilidad Estricta por la Izquierda* y la *Estabilidad Estricta*, estableciendo por tanto una diferencia entre la estabilidad de las familias de operadores de agregación en términos de la estructura del proceso de agregación.

Las definiciones anteriores por tanto, intentan asegurar que cualquier cambio de cardinalidad que se produzca en la base de datos, no impactará significativamente en los resultados de la agregación.

Es sencillo ver que las familias de operadores Mínimo (*Min*) y Máximo (*Max*) satisfacen trivialmente esta propiedad tanto por la derecha como por la izquierda. Lo mismo ocurre con la Media Aritmética $M_n(x_1,\ldots,x_n)=\sum_{i=1}^{n}\frac{x_i}{n}$, la Media Geométrica $G_n(x_1,\ldots,x_n)=(\prod_{i=1}^{n}x_i)^{1/n}$, la Media Armónica $H_n(x_1,\ldots,x_n)=\frac{n}{\sum_{i=1}^{n}1/x_i}$ y la Mediana (*Med*). Además, cualquier familia de operadores binarios idempotentes satisfacen la propiedad por la derecha si tienen una extensión inductiva hacia adelante (A_*), y por la izquierda si tienen una extensión indictiva hacia atrás (A^*). Sin embargo, otras familias de operadores satisfacen débilmente la propiedad de estabilidad, como por ejemplo es el caso de la Media Ponderada bajo esta definición: $W_n(x_1,\ldots,x_n)=\frac{\sum_{i=1}^{n}x_i*w_i}{\sum_{i=1}^{n}w_i}$. En esta situación, se entiende que W_n converge en ley a W_{n-1} y por lo tanto, la propiedad de estabilidad se satisface en el límite.

DEFINICION 2: Sea $\{A_n:[0,1]^n\to[0,1],n\in\mathbb{N}\}$ una familia de operadores de agregación. Entonces:

1. A_n cumple con la propiedad de *Estabilidad por la Derecha* si $\forall x_1,\ldots x_{n-1}\in[0,1]$, se satisface lo siguiente:

$$\lim_{n\to+\infty}|A_n(x_1,x_2,...x_{n-1},A_{n-1}(x_1,x_2,...,x_{n-1}))-A_{n-1}(x_1,x_2,...,x_{n-1})|=0.$$

2. A_n cumple con la propiedad de *Estabilidad por la Izquierda* si $\forall x_1,\ldots x_{n-1}\in[0,1]$, se satisface lo siguiente:

$$\lim_{n\to+\infty}[\ |A_n(A_{n-1}(x_1,x_2,...,x_{n-1}),x_1,x_2,...x_{n-1})-A_{n-1}(x_1,x_2...,x_{n-1})|=0.$$

3. A_n cumple con la propiedad de *Estabilidad* si se satisfacen los dos puntos anteriores.

Por último, otras familias de operadores de agregación satisfacen la propiedad de estabilidad pero en un sentido más débil, por ejemplo el Producto $P_n(x_1,\ldots,x_n)=\prod_{i=1}^{n}x_i$. En esta situación, se observa que P_n tiene una convergencia casi segura a P_{n-1}, y por lo tanto, la propiedad de estabilidad se cumple bajo una convergencia asintótica.

DEFINICION 3: Sea $\{A_n:[0,1]^n\to[0,1],n\in\mathbb{N}\}$ una familia de operadores de agregación. Entonces:

1. A_n cumple con la propiedad de *Estabilidad Débil por la Derecha* si $\forall x_i\sim U(0,1)$, se satisface lo siguiente:

$$\mathbb{P}[\lim_{n\to+\infty}|A_n(x_1,x_2,...x_{n-1},A_{n-1}(x_1,x_2,...,x_{n-1}))-A_{n-1}(x_1,x_2,...,x_{n-1})|=0\ ]=1.$$

2. A_n cumple con la propiedad de *Estabilidad Débil por la Izquierda* si $\forall x_i\sim U(0,1)$, se satisface lo siguiente:

$$\mathbb{P}[\lim_{n\to+\infty}[\ |A_n(A_{n-1}(x_1,x_2,...,x_{n-1}),x_1,x_2,...x_{n-1})-A_{n-1}(x_1,x_2...,x_{n-1})|=0\ ]=1.$$

3. A_n cumple con la propiedad de *Estabilidad Débil* si se satisfacen los dos puntos anteriores.

PROPOSICION 1: Sea $\{A_n:[0,1]^n\to[0,1],n\in\mathbb{N}\}$ una familia de operadores de agregación.

1. Si la familia $\{A_n\}$ satisface la propiedad de *Estabilidad Estricta*, entonces también satisface la propiedad de *Estabilidad*.

2. Si la familia $\{A_n\}$ satisface la propiedad de *Estabilidad*, entonces también satisface la propiedad de *Estabilidad Débil*.

Finalmente, con la siguiente definición se completa la definición de niveles de estabilidad de una familia de operadores de agregación. Y bajo este último escenario, una familia de operadores de agregación que otorgue como resultado el valor máximo si n es impar y el mínimo si n es par por ejemplo, es definida como una familia *Inestable*.

DEFINICION 4: Sea $\{A_n:[0,1]^n\to[0,1],n\in\mathbb{N}\}$ una familia de operadores de agregación. Entonces:

1. A_n cumplirá la propiedad de *Inestabilidad por la Derecha*, si esta familia no satisface la *Estabilidad Débil por la Derecha*.

2. A_n cumplirá la propiedad de *Inestabilidad por la Izquierda*, si esta familia no satisface la *Estabilidad Débil por la Izquierda*.

3. A_n cumplirá la propiedad de *Inestabilidad* si esta familia no satisface la *Estabilidad Débil*.

3 MAPEO DEL NIVEL DE ESTABILIDAD DE ALGUNAS FAMILIAS DE OPERADORES DE AGREGACIÓN

El nivel de estabilidad de una familia de operadores de agregación otorga información respecto de la robustez del proceso de agregación en el que está involucrada. A continuación, se analiza el nivel de estabilidad de las familias de operadores de agregación comúnmente utilizadas, como un acercamiento a la consistencia de los operadores de una misma familia, independientemente de la cardinalidad de la base de datos a agregar.

Tabla 1: Nivel de Estabilidad de algunas Familias de Operadores de Agregación.

F. de Operadores de Agregación $\{A_n\}_{n\varepsilon\mathbb{N}}$	Estab. Estr.	Estab.	Estab. Débil	Inest.
$Min_n = Min(x_1,..,x_n)$	*R, L*	–	–	–
$Max_n = Max(x_1,..,x_n)$	*R, L*	–	–	–
$Md_n = Md(x_1,..,x_n)$	*R, L*	–	–	–
$M_n = \sum_{i=1}^{n} \frac{x_i}{n}$	*R, L*	–	–	–
$G_n = (\prod_{i=1}^{n} x_i)^{1/n}$	*R, L*	–	–	–
$Hn = \frac{n}{\sum_{i=1}^{n} 1/x_i}$	*R, L*	–	–	–
$Qn = \prod_{i=1}^{n} x_i^i$	–	–	*R, L*	–
$P_n = \prod_{i=1}^{n} x_i$	–	–	*R, L*	–
$A_n^* = A^*(x_1,\ldots,x_n)$	*R*	–	–	*L*
$A_{*n} = A_*(x_1,\ldots,x_n)$	*L*	–	–	*R*
$W_n = \sum_{i=1}^{n} x_i \cdot \frac{v_i}{\sum_{j=1}^{n} v_j}$	*R*	–	*L*	–
$O_n = \sum_{i=1}^{n} x_{(i)} \cdot \frac{v_i}{\sum_{j=1}^{n} v_j}$	–	–	*R, L*	–

Nota: *R* y *L* indican estabilidad por la derecha o por la izquierda.
Nota: Los pesos de W_n y O_n no tienen restricciones.

En la Tabla 1 se observa que las seis primeras familias satisfacen la propiedad de *Estabilidad Estricta* para ambos tipos de estructura lineal, por la derecha y por la izquierda, mientras que las otras familias no. Por ejemplo, $\{Q_n\}$ y $\{P_n\}$ satisfacen la propiedad de *Estabilidad Débil*, y los operadores idempotentes binarios son *Estrictamente Estables por la Derecha* si su extensión es inductiva hacia adelante, y por la izquierda si es hacia atrás. Además, el nivel de estabilidad de las familias $\{W_n\}$ y $\{O_n\}$ dependen de la definición de sus pesos, y particularmente para la definición expuesta, la Media Ponderada es una familia *Estrictamente Estable por la Derecha* y *Débilmente Estable por la Izquierda*, y el operador OWA es una familia *Débilmente Estable* bajo la misma definición de pesos.

4 CONCLUSIONES

En nuestra opinión, una familia de operadores de agregación no se debe entender sólo como una familia de operadores $n-ary$, en tanto todos estos operadores deben estar íntimamente relacionados a lo largo del proceso de agregación en el que están involucrados. En consideración con esto, se presentan algunas propiedades de consistencia a fin de asegurar la robustez de un proceso de agregación.

Es evidente que no debería ser posible definir una familia de operadores de agregación $\{A_n\}$ en la cual $\{A_2\}$ corresponde a la Media Aritmética, $\{A_3\}$ a la Media Geométrica y $\{A_4\}$ al Mínimo por ejemplo. Este simple ejemplo pone en evidencia que el proceso de agregación demanda una unidad conceptual más que sólo una fórmula matemática.

La noción de *Estabilidad* propuesta en este trabajo enfatiza la idea de *Robustez/Continuidad/Estabilidad* de una familia de operadores de agregación, en el sentido que un operador definido para n elementos no debiera diferir demasiado de uno definido para $n-1$ elementos.

Referencias

[1] A. Amo, J. Montero, E. Molina. Representation of consistent recursive rules. European Journal of Operational Research 130, 29-53, 2001.

[2] G. Beliakov, A. Pradera, T. Calvo. Aggregation Functions, a Guide to Practitioners (Springer-Verlag, Berlin, 2007).

[3] H. Bustince, E. Barrenechea, J. Fern´andez, M. Pagola, J. Montero and C. Guerra: Aggregation of neighbourhood information by means of interval type 2 fuzzy relations.

[4] T. Calvo, A. Kolesarova, M. Komornikova, R. Mesiar. Aggregation operators, properties, classes and construction methods. In T. Calvo et al. (Eds.): Aggregation Operators New trends ans Aplications (Physica-Verlag, Heidelberg, 3-104, 2002.

[5] T. Calvo, G. Mayor, J. Torrens, J. Suñer, M. Mas and M. Carbonell. Generation of weighting triangles associated with aggregation fuctions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 8(4) (2000) 417-451.

[6] C.H. Carlsson, R. Fuller. Fuzzy reaasoning in decision making and optimization. Heidelberg, Springfield-Verlag, 2002.

[7] V. Cutello, J. Montero. Hierarchical aggregation of OWA operators: basic measures and related computational problems. Uncertainty, Fuzzinesss and Knowledge-Based Systems 3, 17-26,1995.

[8] V. Cutello, J. Montero. Recursive families of OWA operators. Proceedings FUZZ-IEEE Conference (IEEE Press, Piscataway, 1994), 1137-1141.

[9] V. Cutello, J. Montero. Recursive connective rules. International Journal of Intelligent Systems 14, 3-20, 1999.

[10] D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk, H. Prade. Terminological difficulties in fuzzy set theory - the case of intuitionistic fuzzy sets. Fuzzy Sets and Systems 156, 485-491, 2005.

[11] L.W. Fung, K.S. Fu. An axiomatic approach to rational decision making in a fuzzy environment. In: L.A. Zadeh et at. (Eds.): Fuzzy Sets and their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975), 227-256.

[12] D. Gómez, J. Montero. A discussion of aggregation functions. Kybernetika 40, 107-120, 2004.

[13] D. Gómez, J. Montero, J. Yáñez, C. Poidomani (2007): A graph coloring algorithm approach for image segmentation. Omega 35:173–183.

[14] D. Gómez, J. Montero and J. Yánez (2006): A coloring algorithm for image classification. Information Sciences 176, 3645-3657.

[15] M. Grabisch, J. Marichal, R. Mesiar and E. Pap. Aggregation Functions (Encyclopedia of Mathematics and its Applications, 2009).

[16] A. Kolesárová. Sequential aggregation. In: M. González et al. (Eds.): Proceedings of the Fifth International Summer School on Aggregation Operators, AGOP (Universitat de les Illes Balears, Palma de Mallorca, 2009), 183-187.

[17] V. López, L. Garmendia, J. Montero, G. Resconi. Specification and computing states in fuzzy algorithms. Uncertainty, Fuzziness and Knowledge-Based Systems 16, 301-336, 2008.

[18] V. López, J. Montero. Software engineering specification under fuzziness. Multiple-Valued Logic and Soft Computing 15, 209-228, 2009.

[19] V. López, J. Montero and J. Tinguaro Rodriguez. Formal Specification and implementation of computational aggregation Functions. Computational Intelligence, Foundations and Applications Proceedings of the 9th International FLINS Conference 2010, 523-528

[20] J. Montero. A note on Fung-Fu‘s theorem. Fuzzy Sets and Systems 13, 259-269, 1985.

[21] J. Montero, D. Gómez, H. Bustince. On the relevance of some families of fuzzy sets. Fuzzy Sets and Systems 158, 2429-2442, 2007.

[22] J. Montero, V. López, D. Gómez. The role of fuzziness in decision making. Studies in Fuzziness and Soft Computing 215, 337-349, 2007.

[23] R.R. Yager. On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man and Cybernetics 18, 183-190, 1988.

[24] R.R. Yager and A.Rybalov. Nonconmutative self-identity aggregation. Fuzzy Sets and Systems, 85:73-82, 1997.

Sobre t-conormas y uninormas migrativas

M. Mas, **M. Monserrat**, **D. Ruiz-Aguilera**, **J. Torrens**

Departament de Ciències Matemàtiques i Informàtica,
Universitat de les Illes Balears
Campus UIB, Crta. Valldemossa Km. 7.5,
07122 Palma de Mallorca, {dmimmg0, dmimma0, daniel.ruiz, dmijts0}@uib.es

Resumen

En este trabajo se introducen las nociones de t-conormas α-migrativas respecto a otra t-conorma fija S_0 y uninormas con elemento neutro $e \in (0,1)$ α-migrativas respecto a otra uninorma fija U_0 con el mismo elemento neutro. Se caracterizan entonces todas las t-conormas continuas que son α-migrativas respecto del máximo, la suma probabilística y la t-conorma de Łukasiewicz. Asimismo se caracterizan las uninormas que son α-migrativas respecto de otra uninorma U_0 cuando éstas pertenecen a alguna de las clases $\mathscr{U}_{\text{mín}}$, $\mathscr{U}_{\text{mín}}$, idempotentes y/o representables.

Palabras Clave: Función de agregación, migratividad, t-norma, t-conorma, uninorma.

1 INTRODUCCIÓN

Las funciones de agregación han experimentado un gran desarrollo en los últimos años debido a su gran cantidad de aplicaciones en diversos campos, tanto científicos como sociales (ver [3, 6, 16]). En particular, diversas clases de funciones de agregación resultan especialmente importantes por su papel fundamental en la teoría de los conjuntos borrosos y sus aplicaciones (ver [18]). Es el caso de las t-normas, t-conormas y, más en general, de los conjuntores y disyuntores que pretenden generalizar respectivamente los conectivos lógicos "AND" y "OR" de la lógica clásica. A menudo, el estudio de estos conectivos se dirige hacia la caracterización de aquellos que cumplen determinadas propiedades que puedan ser útiles en cada contexto. El estudio de estas propiedades para determinadas funciones de agregación se traduce habitualmente en la resolución de sendas ecuaciones funcionales. Ejemplos de ello, tanto para t-normas y t-conormas como para uninormas, pueden hallarse en [5, 11, 12, 13, 14, 18, 19, 20].

Una de estas propiedades es la denominada propiedad α-*migrativa* introducida en [10] y que, para $\alpha \in [0,1]$ y una aplicación $F : [0,1]^2 \to [0,1]$, se escribe como

$$F(\alpha x, y) = F(x, \alpha y) \quad \text{para todo } x, y \in [0,1]. \tag{1}$$

La propiedad migrativa (y generalizaciones de la misma) ha sido estudiada para t-normas en [11, 12], para semicópulas, quasi-cópulas y cópulas en [21, 9] y para funciones de agregación en general en [4, 15]. El interés de esta propiedad radica en sus aplicaciones, por ejemplo, en procesos de toma de decisiones ([4]), cuando una repetida información parcial necesita ser fusionada en un resultado global, o en procesamiento de imágenes ([21]).

Notemos que en la ecuación (1) el producto αx puede ser reemplazado por una t-norma cualquiera T_0 obteniendo la propiedad para t-normas, llamada (α, T_0)-migrativa, que puede escribirse como

$$T(T_0(\alpha, x), y) = T(x, T_0(\alpha, y)) \quad \text{para todo } x, y \in [0,1] \tag{2}$$

siendo T_0 una t-norma cualquiera y $\alpha \in [0,1]$. Esta generalización de la migratividad para t-normas ha sido estudiada recientemente en [12].

En esta comunicación queremos realizar un estudio similar para otras generalizaciones, basadas en la misma idea de la ecuación (2), pero para t-conormas y también para uninormas. De esta manera, definiremos los conceptos análogos de t-conorma (α, S_0)-migrativa, donde S_0 es una t-conorma, y de uninorma (α, U_0)-migrativa, donde U_0 es una uninorma cualquiera. Haremos un estudio similar al realizado en [12] para estas nuevas propiedades, caracterizando las t-conormas (α, S_0)-migrativas y, en el caso de uninormas, iniciaremos el estudio con algunos de los casos posibles a considerar dependiendo del tipo de uninorma que fijemos previamente.

Hemos organizado este artículo en diversas secciones. Tras esta introducción incluimos unos preliminares para establecer la notación necesaria y recordar algunas definiciones

básicas. Le siguen las dos secciones principales dedicadas la primera al caso de t-conormas continuas y la segunda al caso de uninormas con las operaciones subyacentes continuas. Terminamos con un apartado de conclusiones y trabajo futuro.

2 PRELIMINARES

Se supondrá conocida la teoría básica de t-normas y t-conormas (ver [1, 18]). Solo daremos en esta sección las definiciones iniciales, todas las notaciones y resultados necesarios pueden encontrarse en [18].

Definición 1. Una aplicación binaria $T : [0,1]^2 \to [0,1]$ se dice que es una t-*norma* si es asociativa, conmutativa, creciente en cada variable y tiene 1 como neutro.

Definición 2. Una aplicación binaria $S : [0,1]^2 \to [0,1]$ se dice que es una t-*conorma* si es asociativa, conmutativa, creciente en cada variable y tiene 0 como neutro.

Definición 3. Una aplicación $N : [0,1] \to [0,1]$ se dice que es una *negación* si es decreciente con $N(0) = 1$ y $N(1) = 0$. Si además N es involutiva, i.e. $N(N(x)) = x$ para todo $x \in [0,1]$, se dice que N es una *negación fuerte*.

Definición 4. Dada una aplicación binaria $F : [0,1]^2 \to [0,1]$ se define su aplicación N-*dual* como la aplicación $F_N^* : [0,1]^2 \to [0,1]$ dada por

$$F_N^*(x,y) = N(F(N(x),N(y))) \quad \text{para todos } x,y \in [0,1].$$

En el caso $N(x) = 1-x$, denotaremos simplemente la N-dual por F^*.

Es bien sabido que, cuando N es una negación fuerte, la operación de dualización es involutiva, esto es, para cualquier aplicación binaria F se verifica $(F_N^*)_N^* = F$. Además, se verifica la siguiente proposición.

Proposición 1. *Sea N una negación fuerte. Una aplicación binaria $F : [0,1]^2 \to [0,1]$ es una t-norma si y solo si su aplicación dual F_N^* es una t-conorma.*

Definición 5. Dada una t-norma T_0 y un $\alpha \in [0,1]$, una t-norma T se dice que es (α, T_0)-*migrativa* o que es α-*migrativa respecto de* T_0 si se verifica

$$T(T_0(\alpha,x),y) = T(x,T_0(\alpha,y)) \quad \text{para todo } x,y \in [0,1] \tag{3}$$

Un estudio detallado de las t-normas continuas que son α-migrativas respecto de las t-normas mínimo, producto y Łukasiewicz puede hallarse en [12]. Una extensión de este trabajo se ha realizado en [23] donde se caracterizan las t-normas continuas que son α-migrativas respecto de cualquier t-norma continua T_0.

2.1 UNINORMAS

Definición 6. Una aplicación binaria $U : [0,1]^2 \to [0,1]$ se dice que es una *uninorma* si es asociativa, conmutativa, creciente en cada variable y tiene un elemento neutro $e \in [0,1]$.

Evidentemente, una uninorma con elemento neutro $e = 1$ es una t-norma y una uninorma con elemento neutro $e = 0$ es una t-conorma. Para cualquier otro valor $e \in (0,1)$ se obtiene un operador que funciona como una t-norma en el cuadrado $[0,e]^2$, como una t-conorma en el cuadrado $[e,1]^2$ y cuyos valores oscilan entre el mínimo y el máximo en el conjunto $A(e)$ dado por

$$A(e) = [0,e) \times (e,1] \cup (e,1] \times [0,e).$$

De forma usual denotaremos una uninorma con elemento neutro e y t-norma y t-conorma subyacentes, T y S, por $U = (e,T,S)$. Para cualquier uninorma se tiene $U(0,1) \in \{0,1\}$ y una uninorma U se dice que es conjuntiva cuando $U(1,0) = 0$ y disyuntiva cuando $U(1,0) = 1$. Por otra parte, las tres clases de uninormas más estudiadas son:

- Las uninormas en $\mathscr{U}_{\text{mín}}$ (respectivamente $\mathscr{U}_{\text{máx}}$), aquellas que vienen dadas por el mínimo (respectivamente máximo) en $A(e)$.
- Las uninormas *idempotentes*, aquellas que verifican $U(x,x) = x$ para todo $x \in [0,1]$.
- Las *representables*, aquellas que son continuas en $[0,1]^2 \setminus \{(1,0),(0,1)\}$.

En lo que sigue damos una caracterización de cada una de las clases de uninormas anteriormente descritas.

Teorema 2. ([13]) *Sea $U : [0,1]^2 \to [0,1]$ una uninorma con elemento neutro $e \in]0,1[$. Entonces, las secciones $x \mapsto U(x,1)$ y $x \mapsto U(x,0)$ son continuas excepto quizás en el punto $x = e$ si y solo si U viene dada por una de las siguientes fórmulas.*

(a) Si $U(0,1) = 0$, entonces $U(x,y) =$

$$\begin{cases} eT_U\left(\frac{x}{e},\frac{y}{e}\right) & si\ (x,y) \in [0,e]^2 \\ e+(1-e)S_U\left(\frac{x-e}{1-e},\frac{y-e}{1-e}\right) & si\ (x,y) \in [e,1]^2 \\ \text{mín}(x,y) & en\ cualquier\ otro\ caso. \end{cases} \tag{4}$$

(b) Si $U(0,1) = 1$, entonces $U(x,y) =$

$$\begin{cases} eT_U\left(\frac{x}{e},\frac{y}{e}\right) & si\ (x,y) \in [0,e]^2 \\ e+(1-e)S_U\left(\frac{x-e}{1-e},\frac{y-e}{1-e}\right) & si\ (x,y) \in [e,1]^2 \\ \text{máx}(x,y) & en\ cualquier\ otro\ caso. \end{cases} \tag{5}$$

En ambas fórmulas T_U es una t-norma y S_U es una t-conorma. Denotaremos por $\mathscr{U}_{\text{mín}}$ la clase de uninormas de la forma (4) y por $\mathscr{U}_{\text{máx}}$ la clase de uninormas de la forma (5).

Teorema 3. ([22]) *U es una uninorma idempotente con elemento neutro $e \in [0,1]$ si y solo si existe una función decreciente $g : [0,1] \to [0,1]$, simétrica respecto de la identidad con $g(e) = e$ tal que*

$$U(x,y) = \begin{cases} \min(x,y) & \text{si } y < g(x) \text{ o} \\ & y = g(x) \text{ y } x < g^2(x) \\ \max(x,y) & \text{si } y > g(x) \text{ o} \\ & y = g(x) \text{ y } x > g^2(x) \\ \min(x,y) \text{ o } \max(x,y) & \text{si } y = g(x) \text{ y } x = g^2(x) \end{cases}$$

siendo conmutativa en los puntos (x,y) tales que $y = g(x)$ con $x = g^2(x)$.

Teorema 4. ([13]) *Sea $e \in]0,1[$. Una aplicación $U : [0,1]^2 \to [0,1]$ es una uninorma representable si y solo si existe una función continua estrictamente creciente $h : [0,1] \to [-\infty,+\infty]$ con $h(0) = -\infty$, $h(e) = 0$ y $h(1) = +\infty$ de forma que*

$$U(x,y) = h^{-1}(h(x) + h(y))$$

para todos $(x,y) \in [0,1]^2 \setminus \{(0,1),(1,0)\}$ y $U(0,1) = U(1,0) \in \{0,1\}$. La función h se conoce habitualmente como generador aditivo *de U.*

3 t-CONORMAS (α, S_0)-MIGRATIVAS

En esta sección estudiaremos el caso de las t-conormas continuas. Notemos para empezar que si requerimos la ecuación (1) a una t-conorma y tomamos el valor $x = 0$ en la ecuación obtenemos $y = \alpha y$ para todo $y \in [0,1]$. Es decir no existen t-conormas α-migrativas para $\alpha \neq 1$, mientras que para $\alpha = 1$ evidentemente todas lo son.

Lo mismo ocurre si buscamos las t-conormas que verifican (3) para una t-norma cualquiera T_0. Resulta evidente de todo ello que necesitamos adaptar la ecuación al caso de t-conormas requiriendo la migratividad para otra t-conorma S_0 fija.

Definición 7. Sea $\alpha \in [0,1]$ y sea S_0 una t-conorma cualquiera. Diremos que una t-conorma S es (α, S_0)-*migrativa* o que S es α-*migrativa respecto a* S_0, si se verifica

$$S(S_0(\alpha,x),y) = S(x,S_0(\alpha,y)) \tag{6}$$

para todos $x, y \in [0,1]$.

Notemos de la definición anterior que toda t-conorma S es siempre $(0,S_0)$-migrativa y también $(1,S_0)$-migrativa para cualquier t-conorma S_0. De esta forma, en lo que sigue nos centraremos solo en el caso $\alpha \in (0,1)$.

Otra reflexión inmediata es que, gracias a la asociatividad, se tiene que la propia S_0 es siempre (α, S_0)-migrativa para todo α. Asimismo, por dualidad obtenemos fácilmente el siguiente resultado.

Proposición 5. *Sea $\alpha \in (0,1)$ y sea S_0 una t-conorma cualquiera. Una t-conorma S es (α, S_0)-migrativa si y solo si su t-norma dual S^* es $(1-\alpha, T_0)$-migrativa, donde T_0 es la t-norma dual de S_0.*

De esta forma, a partir de los resultados conocidos para t-normas (ver [11] y [12]) se pueden deducir los siguientes para t-conormas.

Proposición 6. *Sea $\alpha \in (0,1)$ y sea S_0 una t-conorma cualquiera. Las siguientes afirmaciones son equivalentes para una t-conorma S cualquiera:*

i) S es (α, S_0)-migrativa

ii) S_0 es (α, S)-migrativa

iii) Las secciones α de S y S_0 coinciden. Es decir, $S(\alpha,x) = S_0(\alpha,x)$ para todo $x \in [0,1]$.

Los tres teoremas siguientes caracterizan las t-conormas continuas que son (α, S_0)-migrativas para los tres casos básicos en que la t-conorma S_0 es, respectivamente, el máximo, la suma probabilística y la t-conorma de Łukasiewicz.

Teorema 7. *Sea $\alpha \in (0,1)$ y sea $S_0 = \max$ la t-conorma máximo. Una t-conorma cualquiera S es $(\alpha, \max)$-migrativa si y solo si S es una suma ordinal de dos t-conormas S_1, S_2 de la forma*

$$S = \langle (0,\alpha,S_1), (\alpha,1,S_2) \rangle. \tag{7}$$

Notar que la caracterización anterior es válida, según el teorema, para t-conormas cualesquiera no necesariamente continuas. En los siguientes casos sí nos restringimos al caso continuo.

Teorema 8. *Sea $\alpha \in (0,1)$ y sea $S_0 = S_{\mathbf{P}}$ la t-conorma suma probabilística. Dada una t-conorma continua S las siguientes afirmaciones son equivalentes.*

i) S es $(\alpha, S_{\mathbf{P}})$-migrativa.

ii) S es una t-conorma estricta con generador aditivo s verificando

$$s(\alpha + x - \alpha x) = s(\alpha) + s(x) \quad \text{para todo } x \in [0,1].$$

iii) S es una t-conorma estricta con generador aditivo s y existe una función continua estrictamente creciente $s_0 : [0,\alpha] \to [0, s_0(\alpha)]$ con $s_0(0) = 0$ y $s_0(\alpha) < +\infty$ de manera que s viene dado por

$$s(x) = k s_0(\alpha) + s_0\left(1 - \frac{1-x}{(1-\alpha)^k}\right)$$

si $1-(1-\alpha)^k \leq x < 1-(1-\alpha)^{k+1}$, para todo $k \geq 0$.

Dado $a \in [0,1]$ definimos recursivamente $a_{\mathbf{L}}^{(n)}$ por $a_{\mathbf{L}}^{(0)} = 1$, $a_{\mathbf{L}}^{(1)} = a$ y

$$a_{\mathbf{L}}^{(k)} = T_{\mathbf{L}}(a_{\mathbf{L}}^{(k)}, a) \quad \text{para todo} \quad k \geq 2.$$

Con esta notación podemos dar el resultado correspondiente a las t-conormas $(\alpha, S_{\mathbf{L}})$-migrativas.

Teorema 9. *Sea $\alpha \in (0,1)$ y sea $S_0 = S_{\mathbf{L}}$ la t-conorma de Łukasiewicz. Dada una t-conorma continua S las siguientes afirmaciones son equivalentes.*

i) S es $(\alpha, S_{\mathbf{L}})$-migrativa.

ii) S es una t-conorma nilpotente con generador aditivo normalizado s (tal que $s(1) = 1$) verificando

$$s(\alpha + x) = s(\alpha) + s(x) \quad \textit{para todo } x \in [0, 1-\alpha].$$

iii) S es una t-conorma nilpotente con generador aditivo normalizado s y existe un automorfismo φ_0 del intervalo $[0,1]$ y un número real a con $(n-1)/n \leq a \leq n/(n+1)$ donde $n = \max\{k \mid 1-\alpha \geq (k-1)/k\}$, de manera que

$$\varphi_0\left(\frac{\alpha - (1-\alpha)_{\mathbf{L}}^{(n)}}{\alpha}\right) = \frac{1-a-a_{\mathbf{L}}^{(n)}}{1-a}$$

y s viene dado por

$$s(x) = 1 - a_{\mathbf{L}}^{(k)} - (1-a)\varphi_0\left(\frac{1-x-(1-\alpha)_{\mathbf{L}}^{(k)}}{\alpha}\right)$$

si $1-(1-\alpha)_{\mathbf{L}}^{(k-1)} \leq x < 1-(1-\alpha)_{\mathbf{L}}^{(k)}$, cuando $1 \leq k \leq n$ y, en cambio, $s(x) =$

$$1 - (a_{\mathbf{L}}^{(n)} + a - 1) - (1-a)\varphi_0\left(\frac{1-x-(1-\alpha)_{\mathbf{L}}^{(n)} + \alpha}{\alpha}\right)$$

si $1-(1-\alpha)_{\mathbf{L}}^{(n)} \leq x \leq 1$.

4 UNINORMAS (α, U_0)-MIGRATIVAS

En esta sección queremos iniciar el estudio para el caso de uninormas. De forma similar a los casos de t-normas y t-conormas podemos dar la siguiente definición.

Definición 8. Sea $\alpha \in [0,1]$ y U_0 una uninorma con elemento neutro $e \in (0,1)$. Diremos que una uninorma U con elemento neutro e es (α, U_0)-migrativa o que U es *α-migrativa respecto a U_0*, si se verifica

$$U(U_0(\alpha, x), y) = U(x, U_0(\alpha, y)) \tag{8}$$

para todos $x, y \in [0,1]$.

De ahora en adelante supondremos que $e \in (0,1)$ es fijo y todas las uninormas U_0, U utilizadas tienen elemento neutro e. Notemos que entonces un primer resultado evidente es que toda uninorma U es (e, U_0)-migrativa para cualquier uninorma U_0. Por otra parte, es fácil ver también que

Proposición 10. *Sea $\alpha \in [0,1]$ y sea U_0 una uninorma cualquiera con elemento neutro e. Las siguientes afirmaciones son equivalentes para una uninorma U cualquiera con elemento neutro e:*

i) U es (α, U_0)-migrativa

ii) U_0 es (α, U)-migrativa

iii) $U(\alpha, x) = U_0(\alpha, x)$ para todo $x \in [0,1]$.

El proceso a seguir ahora será distinguir casos según que la uninorma U_0 sea de una clase o de otra de entre las más habituales (ver los preliminares). En cualquier caso supondremos también a partir de ahora que $\alpha \neq e$ y que las t-normas y t-conormas subyacentes a las uninormas consideradas son continuas.

4.1 CASO EN QUE U_0 ES DE $\mathcal{U}_{\text{mín}}$

Supongamos en esta sección que U_0 es una uninorma de $\mathcal{U}_{\text{mín}}$. Según la Proposición 10, para encontrar las uninormas (α, U_0)-migrativas, se trata de hallar las uninormas cuya α-sección coincida con la α-sección de U_0. Así, podemos dar los siguientes resultados.

Proposición 11. *Sea $\alpha \in [0,1]$ y $U_0 = (e, T_0, S_0), U = (e, T, S)$ uninormas de $\mathcal{U}_{\text{mín}}$. Entonces,*

i) Si $\alpha < e$, U es (α, U_0)-migrativa si y solo si T es $(\frac{\alpha}{e}, T_0)$-migrativa

ii) Si $\alpha > e$, U es (α, U_0)-migrativa si y solo si S es $(\frac{1-\alpha}{1-e}, S_0)$-migrativa

Proposición 12. *Sea $\alpha \in [0,1]$ y $U_0 = (e, T_0, S_0)$ una uninorma de $\mathcal{U}_{\text{mín}}$. Si U es una uninorma representable o una uninorma en $\mathcal{U}_{\text{máx}}$, entonces U nunca es (α, U_0)-migrativa*

Proposición 13. *Sea $\alpha \in [0,1]$ y $U_0 = (e, T_0, S_0)$ una uninorma de $\mathcal{U}_{\text{mín}}$. Si U es una uninorma idempotente con función asociada g, entonces*

i) Si $\alpha < e$, U es (α, U_0)-migrativa si y solo si $g(\alpha) = 1$, $U(\alpha, 1) = \alpha$ y T_0 es una suma ordinal de la forma

$$T_0 = \left\langle \left(0, \frac{\alpha}{e}, T_1\right), \left(\frac{\alpha}{e}, 1, T_2\right) \right\rangle.$$

ii) Si $\alpha > e$, U es (α, U_0)-migrativa si y solo si $g(\alpha) = e$ y S_0 es una suma ordinal de la forma

$$S_0 = \left\langle \left(0, \frac{\alpha - e}{1-e}, S_1\right), \left(\frac{\alpha - e}{1-e}, 1, S_2\right) \right\rangle.$$

4.2 CASO EN QUE U_0 ES DE $\mathscr{U}_{\text{máx}}$

Los resultados de esta sección pueden obtenerse de forma similar a los correspondientes en la sección anterior.

Proposición 14. *Sea $\alpha \in [0,1]$ y $U_0 = (e,T_0,S_0), U = (e,T,S)$ uninormas de $\mathscr{U}_{\text{máx}}$. Entonces,*

i) Si $\alpha < e$, U es (α,U_0)-migrativa si y solo si T es $(\frac{\alpha}{e},T_0)$-migrativa

ii) Si $\alpha > e$, U es (α,U_0)-migrativa si y solo si S es $(\frac{1-\alpha}{1-e},S_0)$-migrativa

Proposición 15. *Sea $\alpha \in [0,1]$ y $U_0 = (e,T_0,S_0)$ una uninorma de $\mathscr{U}_{\text{máx}}$. Si U es una uninorma representable o una uninorma en $\mathscr{U}_{\text{mín}}$, entonces U nunca es (α,U_0)-migrativa*

Proposición 16. *Sea $\alpha \in [0,1]$ y $U_0 = (e,T_0,S_0)$ una uninorma de $\mathscr{U}_{\text{máx}}$. Si U es una uninorma idempotente con función asociada g, entonces*

i) Si $\alpha < e$, U es (α,U_0)-migrativa si y solo si $g(\alpha) = e$ y T_0 es una suma ordinal de la forma

$$T_0 = \left\langle \left(0, \frac{\alpha}{e}, T_1\right), \left(\frac{\alpha}{e}, 1, T_2\right) \right\rangle.$$

ii) Si $\alpha > e$, U es (α,U_0)-migrativa si y solo si $g(\alpha) = 0$, $U(\alpha,0) = \alpha$ y S_0 es una suma ordinal de la forma

$$S_0 = \left\langle \left(0, \frac{\alpha-e}{1-e}, S_1\right), \left(\frac{\alpha-e}{1-e}, 1, S_2\right) \right\rangle.$$

4.3 CASO EN QUE U_0 ES IDEMPOTENTE

Suponemos en esta sección que U_0 es una uninorma idempotente con elemento neutro e y función asociada g_0. En este caso la mayoría de resultados son deducibles de forma inmediata a partir de la Proposición 10 y de los resultados en la sección anterior. Así, tenemos

- Si $U = (e,T,S)$ es una uninorma de $\mathscr{U}_{\text{mín}}$, entonces U es (α,U_0)-migrativa si y solo si U_0 es (α,U)-migrativa y este caso viene caracterizado por la Proposición 13.
- Si $U = (e,T,S)$ es una uninorma de $\mathscr{U}_{\text{máx}}$, entonces U es (α,U_0)-migrativa si y solo si U_0 es (α,U)-migrativa y este caso viene caracterizado por la Proposición 16.

Por otra parte los demás casos se detallan a continuación.

Proposición 17. *Sea $\alpha \in [0,1]$ y U_0 una uninorma idempotente con elemento neutro e y función asociada g_0. Entonces, si U es una uninorma representable nunca es (α,U_0)-migrativa.*

Proposición 18. *Sea $\alpha \in [0,1]$ y U_0 una uninorma idempotente con elemento neutro e y función asociada g_0. Si U es una uninorma idempotente con función asociada g, entonces U es (α,U_0)-migrativa si y solo si $g(\alpha) = g_0(\alpha)$ y $U(\alpha,g(\alpha)) = U_0(\alpha,g_0(\alpha))$.*

4.4 CASO EN QUE U_0 ES REPRESENTABLE

Suponemos en esta sección que U_0 es una uninorma representable con elemento neutro e y generador aditivo h_0. Esta situación la denotaremos por $U_0 = (e,h_0)$. En esta ocasión, teniendo en cuenta de nuevo la Proposición 10, la siguiente proposición se deriva de forma inmediata de los resultados en las secciones anteriores.

Proposición 19. *Sea $\alpha \in [0,1]$ y U_0 una uninorma representable con elemento neutro e. Entonces, si U es una uninorma de $\mathscr{U}_{\text{mín}}$, de $\mathscr{U}_{\text{máx}}$ o idempotente, nunca es (α,U_0)-migrativa.*

De esta manera queda únicamente el caso en que las dos uninormas U_0 i U son ambas representables.

Proposición 20. *Sea $\alpha \in [0,1]$ y U_0, U uninormas representables con elemento neutro e y generadores aditivos h_0 y h respectivamente. Entonces, U es (α,U_0)-migrativa si y solo si la función $\varphi = h \circ h_0^{-1}$ verifica la ecuación*

$$\varphi(\beta + x) = \varphi(\beta) + \varphi(x) \quad \textit{para todo} \quad x \in [0,1] \qquad (9)$$

donde $\beta = h_0(\alpha)$.

5 CONCLUSIONES Y TRABAJO FUTURO

En este artículo hemos introducido y estudiado una extensión de la propiedad migrativa para t-conormas y para uninormas. En el primer caso hemos caracterizado todas las t-conormas continuas que son α-migrativas respecto a las t-conormas máximo, suma probabilística y Łukasiewicz. Con respecto a las uninormas hemos realizado una caracterización similar para los casos de uninormas de $\mathscr{U}_{\text{mín}}$, $\mathscr{U}_{\text{máx}}$, idempotentes y representables.

Queda pendiente como trabajo futuro, las siguientes cuestiones:

- Extender el estudio a t-conormas α-migrativas respecto de cualquier t-conorma continua (para el caso de t-normas se ha hecho en el trabajo pendiente de publicación [23]).
- Resolver la ecuación (9) para caracterizar completamente el caso de uninormas representables.
- Extender el estudio a otros tipos de uninormas como las continuas en $(0,1)^2$ (ver [17]) o las localmente internas ($U(x,y) \in \{x,y\}$) en la región $A(e)$ (ver [7, 8]).

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto nacional MTM2009-10320 con el apoyo de fondos FEDER.

Referencias

[1] Alsina, C., Frank, M.J., Schweizer, B.; Associative Functions. Triangular Norms and Copulas, World Scientific, New Jersey (2006).

[2] De Baets, B.; Idempotent uninorms. European Journal of Operational Research **118**, 631–642 (1999).

[3] Beliakov, G., Pradera, A., Calvo, T.; Aggregation Functions: A Guide for Practicioners, Springer, Berlin Heidelberg (2007).

[4] Bustince, H., Montero, J., Mesiar, R.; Migrativity of aggregation functions. Fuzzy Sets and Systems **160**, 766–777 (2009).

[5] Calvo, T., De Baets, B., Fodor, J.; The functional equations of Frank and Alsina for uninorms and nullnorms. Fuzzy Sets and Systems **120**, 385–394 (2001).

[6] Calvo, T.; Mayor, G., Mesiar, R. (eds.) Aggregation operators. New trends and applications, Studies in Fuzziness and Soft Computing, **97**, Physica-Verlag, Heidelberg, (2002).

[7] Drygaś, P.; Discussion of the structure of uninorms. Kybernetika **41**, 213–226 (2005).

[8] Drygaś, P.; On properties of uninorms with underlying t-norm and t-conorms given as ordinal sums. Fuzzy Sets and Systems **161**, 149–157 (2010).

[9] Durante, F., Ghiselli Ricci, R.; Supermigrative semicopulas and triangular norms. Information Sciences **179**, 2689–2694 (2009).

[10] Durante, F., Sarkoci, P.; A note on the convex combination of triangular norms. Fuzzy Sets and Systems **159**, 77–80 (2008).

[11] Fodor, J., Rudas, I.J.; On continuous triangular norms that are migrative. Fuzzy Sets and Systems **158**, 1692–1697 (2007).

[12] Fodor, J., Rudas, I.J.; An extension of the migrative property for triangular norms. Fuzzy Sets and Systems **168**, 70–80 (2011).

[13] Fodor, J., Yager, R.R., Rybalov, A.: Structure of uninorms. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **5**, 411–427 (1997).

[14] Frank, M.J.; On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Mathematicae **19**, 194–226 (1979).

[15] Ghiselli Ricci, R.; Supermigrative aggregation functions. In the Proceedings of the 6th International Summer School on Aggregation Operators, AGOP-2011. Benevento (Italy), July, 2011. pp. 145–150.

[16] Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation functions. In: Encyclopedia of Mathematics and its Applications **127**, Cambridge University Press (2009).

[17] Hu, S.K., Li, Z.F.; The structure of continuous uninorms. Fuzzy Sets and Systems **124**, 43–52 (2001).

[18] Klement, E.P., Mesiar, R., Pap, E.: Triangular norms, Kluwer Academic Publishers, Dordrecht (2000).

[19] Mas, M., Mayor, G., Torrens, J.; The modularity condition for uninorms and t-operators. Fuzzy Sets and Systems **126**, 207–218 (2002).

[20] Mas, M., Mayor, G., Torrens, J.; The distributivity condition for uninorms and t-operators. Fuzzy Sets and Systems **128**, 209–225 (2002).

[21] Mesiar, R., Bustince, H., Fernandez, J.; On the α-migrativity of semicopulas, quasi-copulas and copulas. Information Sciences **180**, 1967–1976 (2010).

[22] Ruiz-Aguilera, D., Torrens, J., De Baets, B., Fodor, J.: Some remarks on the characterization of idempotent uninorms. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) Lecture Notes in Artificial Intelligence **6178**, pp. 425–434 (2010).

[23] Xie, A., Liu, H.; Solutions to the extended migrative functional equation based on continuous triangular norms. Preprint (2011).

A new proof of a theorem of Ling

Jesús Sols[1]

Abstract

We give a simpler proof of Ling's theorem on additive generation of continuous Archimedean t-norms, and provide some families of such norms.

Keywords: fuzzy relations, t-norms, Ling's theorem.

1 Introduction

A map $T:[0,1]\times[0,1]\longrightarrow[0,1]$ is called a t-norm if fulfills properties

(T1) $T(a,b)=T(b,a)$ (commutativity)

(T2) $T(a,T(b,c))=T(T(a,b),c)$ (associativity)

(T3) $T(a,c)\leq T(b,c)$ whenever $a\leq b$ (monotonicity)

(T4) $T(a,1)=a$ (contour condition)

The t-norms are of particular interest in the theory of fuzzy equivalence relations in a set X, i.e. maps $R:X\times X\longrightarrow[0,1]$ satisfying an obvious property of symmetry ($R(x,y)=R(y,x)$), and of reflexivity ($R(x,x)=1$), but with a property of transitivity which is relative to a t-norm T: If $R(x,y)=a$, $R(y,z)=b$, then $R(x,z)\geq T(a,b)$. Properties T1-T4 are demanded from T in order to give sense to the T-transitivity in some particular contexts. It is also sometimes demanded that for all $a\in(0,1)$, to be $T(a,a)<a$ (T5); And to be continuous (T6). When T satisfies T1-T6, we say T is an Archimedean and continuous t-norm. It is an immediate and well known result that any continuous decreasing function $t:[0,1]\longrightarrow[0,\infty]$ such that $t(1)=0$, additively generates an Archimedean and continuous t-norm (in the sense made precise in the statement of this lemma 1). A theorem of Ling in 1965 [5] asserts the reciprocal of the lemma: any continuous Archimedean T- norm is additively generated by a continuous decreasing function t, with $t(1)=0$, unique up to multiplication by a constant. We present in section 3 a proof of Ling's theorem 1 which looks simpler (actually, it is proved in the expanded version [8] of this note that the algorithmic velocity of our construction is faster).

While recalling without proof the easy direct lemma, we deal, as an application with strictly decreasing continuous functions $h:[0,1]\longrightarrow[0,1]$ which assign to each probability $s=s_{x,y}$ of two elements x,y of a set X resulting distinct while observing only one random factor, the probability $h(s)$ of these two elements resulting distinct after different instances of experimentation. The inverses t of the different instances of h additively generate continuous Archimedean T-norms, each suited to a different situation. In particular, we deal with a new family of t-norms $T_{n,\alpha}$ suited for the transitivity of the fuzzy equivalence relation assigning to two elements $x,y\in X$ of a set, their degree of indistinguishability after observation of a number n of random aspects of it, when the probability α_r of distinguishing them by each $r\leq n$ disagreements has been given in advance. (These include as particular cases the Lukasiewicz t-norm and some of the t-norms of Schweizer-Sklar, and of Yager, so having as limit cases the product t-norm and the drastic t-norm); We describe also a new family suited to Poisson observation procedures (Section 2).

Finally, and as preparatory material, let us recall that the quasi-inverse $t^{[-1]}$ of a function $t:[x_1,x_2]\twoheadrightarrow[y_1,y_2]\subseteq[z_1,z_2]$ with x_i,y_i,z_i real numbers possibly $+\infty$ or $-\infty$, in the case t is continuous and strictly decreasing, is the extension of the continuous strictly decreasing function $t^{-1}:[y_1,y_2]\twoheadrightarrow[x_1,x_2]$ to a continuous nonincreasing function $t^{[-1]}:[z_1,z_2]\twoheadrightarrow[x_1,x_2]$ by taking $t^{[-1]}(z)=x_2$ if $z<y_1$; and $t^{[-1]}(z)=x_1$ if $z>y_2$.

2 Generating families of t-norms

Let Ω be a universe of factors, so that each $x\in X$ is observable through observation $x(\omega)$ of each of the factors $\omega\in\Omega$ in the element x, with the result that two elements $x,y\in X$ appear in the observation undistinguished, $x(\omega)\equiv y(\omega)$, or distinguished.

Let n be a positive integer and r an integer $0 \leq r < n$. Take as degree of indistinguishability $R_{n,r}(x,y)$ between elements $x,y \in X$ the probability of observing no more than r differences (say "tolerance" r) in n observations. This is a reflexive, symmetric, and $T_{n,r}$-transitive fuzzy relation, for the Archimedean and continuous t-norm $T_{n,r}$ generated by the continuous and strictly decreasing function

$$h_{n,r}(s) = (1-s)^n + \Sigma_{k=1}^{r} \binom{n}{k} s^k (1-s)^{n-k}$$

for $s \in [0,1]$, and $h_{n,r}(s) = 0$ for $s > 1$, by applying the following well known

Lemma 1. *Let $h:\ [0,\infty] \longrightarrow [0,1]$ be a continuous function, strictly decreasing in $[0,m]$ for some $m \in (0,\infty]$, with $h(0) = 1$ and $h(x) = 0$ for all $x \in [m,\infty]$ The function $T : [0,1] \times [0,1] \longrightarrow [0,1]$ defined by*

$$T(a,b) = h(h^{[-1]}(a)) + h^{[-1]}(b))$$

is an Archimedean and continuous t-norm.

Observe that $T_{0,n}$ is the t-norm $T_{1/n}^{SS}$ of the Schweizer and Sklar's family [7]

$$T_{1/n}^{SS}(a,b) = (\max\{0, a^{1/n} + b^{1/n} - 1\})^n$$

which for $n = 1$ is the Lukasiewicz's t-norm, and has limit $T_0^{SS}(a,b) = ab$ the product t-norm. Observe also that our norm $T_{n-1,n}$ is the t-norm $T_{1/n}^{Y}$ of the Yager's family defined by

$$T_{1/n}^{Y}(a,b) = \max\ \{0, 1 - ((1-a)^{1/n} + (1-b)^{1/n})^n\}$$

Observe that the limit of the t-norms $T_{n-1,n}$, as n goes to infinity, is the drastic t-norm.

More in general, for a sequence $\alpha = (\alpha_r)$ of real numbers $0 = \alpha_0 \leq ... \leq \alpha_r \leq ... \leq \alpha_{n-1} \leq \alpha_n = 1$ (α_r being the probability of x, y being distinguished by observation of r disagreements) it is natural to define the criticality degree $\alpha^r \in [0,1]$ of r as $\alpha^r = \alpha_{r+1} - \alpha_r$ (a number r of disagreements would be critical, i.e. $\alpha^r = 1$, when $\alpha_r = 0$ and $\alpha_1 = 1$, i.e. when $r+1$ disagreements distinguish but only r disagreements do not distinguish. So we are dealing with this predicate "criticality of r " as a fuzzy predicate). Let us associate then, as indistinguishability relation, the fuzzy relation $R_{n,\alpha}$ that assigns to $x,y \in X$ the probability

$$R_{n,\alpha}(x,y) = \Sigma_{r=0}^{n} \alpha_r\ P(\{\omega_1,...,\omega_n) \in \Omega^n \mid \sharp \{i \mid x(\omega_i) \ncong y(\omega_i)\} = r)\})$$

of indistinguishability of x,y while observing n factors $\omega \in \Omega$, when the probability of distinguishing x,y by r disagreements is α_r. This fuzzy relation is reflexive, symmetric and $T_{n,\alpha}$-transitive for the Archimedean and continuous t-norm $T_{n,\alpha}$ generated by the continuous and strictly decreasing function $h_{n,\alpha} : [0,\infty] \longrightarrow [0,1]$ defined by

$$h_{n,\alpha}(s) = 1 - \Sigma_{r=0}^{n} \binom{n}{r} s^r (1-s)^{n-r} \alpha_r$$

if $s \in [0,1]$; and $h_{n,\alpha}(s) = 0$, if $s > 1$. If $\alpha_r = 0$ and $\alpha_{r+1} = 1$, i.e. if r is "the" critical value, then $R_{n,\alpha} = R_{n,r}$, and $T_{n,\alpha} = T_{n,r}$.

Another application goes to Poisson observational procedures. Assume that the observational factors ω appear at random following a Poisson distribution of parameter λ_0. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a nondecreasing sequence in $[0,1]$ with $\alpha_0 = 0$ and $\lim\ \alpha_n = 1$, with α_n meaning the probability of two elements x,y of a set X being distinguished when they disagree in the observation of n factors ω. We define then a continuous and strictly decreasing function $h_{\lambda_0,\alpha} : [0,1] \longrightarrow [0,1]$ assigning to each probability $s = s_{x,y}$ of x,y being distinguished by observation of only one random factor, a degree $h_{\lambda_0,\alpha}(s) = 1 - \frac{g(s\ \lambda_0)}{g(\ \lambda_0)}$ of indistinguishability of x,y in the time interval $\tau = 1$, where $g(\lambda) = e^{-\lambda} \Sigma_{k=0}^{\infty} \alpha_k\ \frac{\lambda^k}{k!}$ is the probability of two elements being distinguished at time $\tau = 1$ when disagreements ocurr at random following a Poisson of parameter λ. Its inverse $t_{\lambda_0,\alpha} = h_{\lambda_0,\alpha}^{-1}$ generates a t-norm $T_{\lambda_0,\alpha}^{Poiss}$ controlling the transitivity of the natural fuzzy equivalence relation $R_{\lambda_0,\alpha}^{Poiss}\ (x,y) = h_{\lambda_0,\alpha}(s_{x,y})$.

3 The theorem of Ling

In 1965, Ling [5] proved the reciprocal of lemma 1.

Theorem 1. *Given a continuous and Archimedean t-norm $T : [0,1] \times [0,1] \longrightarrow [0,1]$, there exists a continuous additive generator t of T, i.e. a continuous strictly decreasing function $t : [0,1] \longrightarrow [0,\infty]$ such that $t(1) = 0$ and $T(x,y) = t^{[-1]}(t(x) + t(y))$ for all $x,y \in [0,1]$. This function t is unique up to multiplication by a positive constant. (Ling[5])*

The value of the quasi-inverse $h = t^{[-1]}$ of t on any non-negative real number $x \in \mathbb{R}_{\geq 0}$, written uniquely as series $x = \Sigma_{n\in\mathbb{Z}} x_n\ 2^n \geq 0$, with $x_n = 0,1$ and $x_n = 0$ for $n > \log_2 x$, is

$$h(x) = T_{n\in\mathbb{Z}}(x_n a_n + 1 - x_n) = \lim\ T_{n\geq k}(x_n a_n + 1 - x_n)$$

where $... < a_2 < a_1 < 1/2 = a_0 < a_{-1} < a_{-2} < ...$ satisfies $a_{n+1} = T(a_n, a_n)$ for all $n \in Z$.

Here we have seen the binary commutative operator $T(a,b) = a * b$ can be considered as an operator of a finite arbitrary number of arguments, by using the associativity of T. The equality $a_{n+1} = T(a_n, a_n)$ enables us to construct the sequence (a_n): For $n \geq 0$, it is defined recursively by the last equality, but for the negative values of n, a choice of a_n must be made, for instance $a_n = \sup\{a \mid T(a,a) < a_{n+1}\}$.

To prove this theorem, we can assume, without loss of generality, that the wanted function t satisfies $t(1/2) = 1$. Indeed, if t generates the t-norm T, then, for any $k \in \mathbb{R}^+$, also kt generates it, because $t(T(x,y)) = t(x)+t(y)$ implies $kt(T(x,y) = kt(x)+kt(y)$.

As preparation for the proof, denote $b_0 = 1/2$, so that $t(b_0) = 1$. Since $T(0,0) = 0$, $T(1,1) = 1$, and the function $T(x,x)$ is continuous, there is a real number $b_1 \in (0,1)$ such that $T(b_1,b_1) = b_0 < b_1$ (take, for instance, the minimum satisfying this). We have then $t(b_1) = b_0$, since $1 = t(b_0) = t(T(b_1,b_1)) = t(b_1)+t(b_1)$. Take any $b_2 > b_1$ with $T(b_2,b_2) = b_1$ (for instance, the minimum) We then have $t(b_2) = 1/4$, because $\frac{1}{2} = t(b_1) = t(T(b_2,b_2)) = t(b_2)+t(b_2)$. Iterating this argument we can construct an increasing sequence $1/2 = b_0 < b_1 < b_2 < ...$ for which $t(b_n) = 1/2^n$ for $n = 0,1,2,...$

Lemma 2. $\lim b_n = 1$

Proof. The sequence (b_n) is bounded above by 1, so having a supremum $l = \sup b_n$, which, being the sequence increasing, is in fact the limit $l = \lim b_n$. By continuity of t we have $\lim t(b_n) = t(l)$, so that $t(1) = 0 = \lim 1/2^n = t(l)$. Observe l is a fixed point of $T(x,x)$. Indeed, $l = \lim b_n = \lim T(b_{n+1},b_{n+1}) = T(l,l)$, as $T(x,x)$ is continuous.

Being T Archimedean, the only fixed points of $T(x,x) : [0,1] \longrightarrow [0,1]$ are $0,1$. Since $l > 1/2$, this implies $l = 1$ □

Let us denote $a_0 = \frac{1}{2}$, so that $t(a_0) = 1$. Let $a_1 = T(a_0,a_0) < 1/2$. If $a_1 \neq 0$, then

$$t(a_1) = t(T(a_0,a_0)) = t(a_0)+t(a_0) = 2$$

Let now $a_2 = T(a_1,a_1) < a_1$. If $a_2 \neq 0$, it is

$$4 = t(a_1)+t(a_1) = t(T(a_1,a_1)) = t(a_2)$$

We consider now a nonincreasing sequence $\frac{1}{2} = a_0 > a_1 \geq a_2 \geq ...$ with all inequalities being strict inequalities until we eventually arrive to some $a_n = 0$, being then $a_{n'} = 0$ for all $n' \geq n$. This sequence is defined inductively by $a_{n+1} = T(a_n,a_n)$.

If T is strict, i.e. if for each $x \in (0,1]$ it is $T(x,x) \neq 0$, then by an obvious iteration of the former argument, $t(a_n) = 2^n$, for $n = 0,1,2,...$ In particular $t(0) = \infty$

If T is nilpotent, i. e. nonstrict, there is a positive integer n_0 such that $a_n = 0$ for all $n \geq n_0$. Therefore $t^{[-1]}(2^n) = 0$ for all $n \geq n_0$. This implies $t(0) \leq 2^n$. In general $t^{[-1]}(x) = 0$ implies $t(0) \leq x$.

Lemma 3. *lim* $a_n = 0$

Proof. The sequence (a_n) is bounded below by 0, so it has an infimum q which, being the sequence decreasing, is its limit: $q = \lim a_n$. This value q is a fixed for the function $T(x,x) : [0,1] \longrightarrow [0,1]$. Indeed,

$$q = \lim a_n = \lim T(a_{n-1},a_{n-1}) = T(q,q)$$

by being $T(x,y)$, thus $T(x,x)$, continuous. Since T is Archimedean, the only fixed points of $T(x,x)$ are $0,1$. Being $q < 1/2$, this implies $q = 0$ □

We will use in the sequel the notation $a_{-n} := b_n$ for all natural numbers $n = 1,2,...$ so $(a_n)_{n\in\mathbb{Z}}$ is decreasing.

Let us first prove the *unicity* in the theorem. For this, assume t is a continuous additive generator of T as in the statement, and let $h = t^{[-1]}$ be its continuous nonincreasing quasi-inverse. We then have $h(2^n) = a_n$, for all $n \in \mathbb{Z}$. Let $\mathbb{Z}_2$ be the ring of dyadic numbers, i.e. rational numbers with denominator a power of 2 (this is the ring of fractions $\mathbb{Z}_2 = \frac{\mathbb{Z}}{2^{\mathbb{N}}}$ of $\mathbb{Z}$ by the multiplicatively closed subset $2^{\mathbb{N}} \subseteq \mathbb{Z}$ of the powers of 2). Let $x \in \mathbb{Z}_2$ be a dyadic number, and express it as in the statement, thus with $x_n = 0$ for almost all $n \in \mathbb{Z}$. Define

$$s_x := T_{n\in\mathbb{Z}}(x_n a_n + 1 - x_n) \in [0,1]$$

(almost all arguments $x_n a_n + 1 - x_n$ are 1). This is $s_x = T_{x_n\neq 0}(a_n)$ if $x \neq 0$, and $s_0 = 1$. We then have

$$\begin{aligned} h(x) &= h(\Sigma_{n\in\mathbb{Z}} x_n 2^n) = h(\Sigma_{x_n\neq 0} 2^n) = T_{x_n\neq 0}(a_n) = \\ &= T_{n\in\mathbb{Z}}(x_n a_n + 1 - x_n) = s_x \end{aligned}$$

if $x \neq 0$, and $h(0) = 1 = s_0$. This proves the wanted unicity, since two different choices of continuous additive generator t would have two different continuous quasi-inverse functions $h = t^{[-1]}$ having same behavior $h(x) = s_x$ on $\mathbb{Z}_2 \cap [0,\infty)$, in contradiction with the fact that this set is dense in $[0,\infty)$.

Next, we prove the *existence*. To construct a continuous additive generator t as required, or equivalently its quasi-inverse $h : [0,\infty] \longrightarrow [0,1]$, we first assign values $h(x) = s_x$ to any $x \in \mathbb{Z}_2 \cap [0,\infty)$, where s_x is as in the proof of unicity. Observe that, for two $x,y \in \mathbb{Z}_2 \cap [0,\infty)$, this assignation satisfies $h(x+y) = h(x) * h(y)$. Indeed, we can assume, without loss of generality, that $x = \frac{x'}{2^k}$, $y = \frac{y'}{2^k}$ for nonnegative integers x',y',k. Since for any $n \geq -k$, it is

$$a_n = a_{n-1} * a_{n-1} = a_{-k} * {}^{2^{n+k+1})} * a_{-k} = a_{-k}^{*2^{n+k+1}}$$

we have $h(x) = a_{-k}^{*x'}$ and $h(y) = a_{-k}^{*y'}$, thus

$$h(x+y) = a_{-k}^{*(x'+y')} = a_{-k}^{*x'} * a_{-k}^{*y'} = h(x) * h(y)$$

Let us see that this assignation $h : \mathbb{Z}_2 \cap [0,\infty) \longrightarrow [0,1]$ is strictly decreasing in $\mathbb{Z}_2 \cap [0,l]$, where $l = \sup\{x \mid h(x) > 0\}$. To see this, express two dyadic numbers $x < y$ in this range with a same denominator $\frac{x'}{2^n} < \frac{y'}{2^n}$, being $x',y' \in \mathbb{Z}_{\geq 0}$.

For $k = 1$, ..., $y' - x'$, the sequence (c_k) with $c_k = \frac{x'}{2^n} + \frac{k}{2^n}$ has image $(h(c_k))$ decreasing since

$$\begin{aligned} h(c_k) &= h(\frac{x'}{2^n} + \frac{k}{2^n}) = h(\frac{x'}{2^n} + \frac{k-1}{2^n}) * h(\frac{1}{2^n}) \\ &\leq h(\frac{x'}{2^n} + \frac{k-1}{2^n}) = h(c_{k-1}) \end{aligned}$$

by axiom (T3) . Therefore, $h(x) \geq h(y)$. If one of the former inequalities were an equality, for say k_0 , then it would be the same for all $k \geq k_0$. This would be true in particular for $k = 2k_0$, arriving then to

$$h(c_{k_0}) = h(c_{k_0} + c_{k_0}) = h(c_{k_0}) * h(c_{k_0})$$

with $0 < x < c_{k_0} \leq y < 1$, which is a contradiction with T being Archimedean. We thus obtain $h(x) > h(y)$, as wanted.

For a positive real number $x = \Sigma_{n \in \mathbb{Z}} x_n \, 2^n \in \mathbb{R}^+$ (all $x_n = 0$, for $n > \log_2 x$) and positive integer $k > 0$, define

$$x_{[k]} = \Sigma_{n \geq -k} x_n \, 2^n \in \mathbb{Z}_2^+ \text{ and } x_{(k)} = x_{[k]} + \frac{1}{2^k} \in \mathbb{Z}_2^+$$

The increasing sequence $x_{[k]}$ and the decreasing sequence $x_{(k)}$ have limit x, as well as the sequences $h(x_{[k]})$ and $h(x_{(k)})$ have also the same limit, which we will denote $h(x)$, since

$$\begin{aligned} \lim \, h(x_{(k)}) &= \lim T(h(x_{[k]}), h(\frac{1}{2^k})) = \\ &= T(\lim h(x_{[k]}), \lim \, a_{-k}) = \lim \, h(x_{[k]}) \end{aligned}$$

The nonincreasing map $h : [0, \infty) \longrightarrow [0, 1]$ so defined is continuous, since we have seen that each $x \in [0, \infty)$ is limit of an increasing sequence and a decreasing sequence whose images converge to $h(x)$. Observe that h is decreasing in $[0, l]$, where $l = \sup \{y \mid h(y) > 0\}$, since we proved this on the dyadic numbers in this interval. Clearly, its extension $h : [0, \infty] \longrightarrow [0, 1]$ by taking $h(\infty) = 0$, is still continuous. The inverse $t : [0, 1] \longrightarrow [0, \infty]$ of the decreasing restricted function $h\lfloor_{[0,l]}$ is the wanted continuous additive generator of T. This proves the existence, and the last part of the theorem follows from the construction. □

The error in our approximation to $h(x)$ is estimated in [8] by using the inequalities $h(x_{[k]}) < h(x) < h(x_{(k)})$. The computational advantage of our construction stems from the possibility of storing the values $h(2^n)$ for high numbers 2^n which are expressed by low exponents $n \geq 0$, and similar remark for numbers with many decimal digits. The way we approximate $h(x)$ on any $x \in \mathbb{R}^+$ reaches easily any required precision level, by making a product $*$, i.e. with the t-norm T, of some stored values $h(2^n)$

In fact, we prove in [8] that our approximation $h(x_{[k]})$ to the value $h(x)$ is computable with algorithmic velocity $O(\log_2 x, k)$, clearly faster than the approximation arising from the original proof of Ling.

Concluding remarks. We have given a proof of the theorem of Ling, which is a nontrivial reciprocal of the trivial lemma asserting that any decreasing continuous function t with $t(0) = 1$ additively generates a continuous Archimedean t-norm T. This reciprocal asserts that any continuous Archimedean t-norm T can be generated in this way and that its generator t is unique. The way in which we construct this generator looks simpler than the one in the original proof of Ling, and in fact we show in [8] that it has computational advantages, summarized in the fact that its algorithmic velocity is faster than the one of Ling's.

We also have constructed a family of t-norms which generalizes some t-norms of Schweizer and Sklar, including the Lukasiewicz norm, and also some t-norms of Yager. It seems interesting and realistic, as future work, to look for more general norms generalizing all t-norms of Schweizer-Sklar and of Yager, and even some other norms now scattered in the bibliography.

Acknowledgements

I thank professor Jordi Recasens for having introduced me to this subject.

References

[1] D. Dubois, H.M. Prade, "Fundamentals of Fuzzy Sets", Kluwer Academic Publishers, Norwell, 2000.

[2] J. Hernandez, I. Montes, D. Martinetti, S. Montes, "Estudio de la probabilidad de Zadeh para t- normas Arquimedianas", ESTYLF 2010

[3] E.P. Klement, R. Mesiar, E. Pap, "Triangular Norms", Kluwer Academic Publishers, Dordrecht, 2000, 122.

[4] R. Kruse, J. Gebhardt, F. Klawonn, "Foundations of Fuzzy Systems", Wiley, 1994, Chichester.

[5] C.M. Ling, "Representation of associative functions", *P*ubl. Mat. Debrecen 12. 189-212, 1965.

[6] J. Recasens, "A Small course on indistinguishability operators", 2007, Preprint de la Escuela de Arquitectura de Barcelona.

[7] B. Schweizer, A. Sklar, "Associative Functions and Statistical Triangle Inequalities", Publ Math Debrecen 8: 169-186, 1961.

[8] J. Sols "Sobre t-normas y el teorema de Ling" Proyecto de fin de Master en Ingeniería Matemática. Universitat Politecnica de Catalunya, Facultat de Matematicas y Estadistica. Barcelona, 2012. (To appear)

UNA TIPOLOGÍA CONJUNTA DE LA BIPOLARIDAD Y LA INCERTIDUMBRE BORROSA

J. Tinguaro Rodríguez[1], Camilo A. Franco[1], Javier Montero[1]

[1] *Departamento de Estadística e Investigación Operativa, Facultad de Matemáticas, Universidad Complutense de Madrid*
jtrodrig@mat.ucm.es

Resumen

El objetivo de este trabajo es presentar una metodología que permita clarificar las relaciones existentes entre las nociones de incertidumbre borrosa (*fuzziness*) y bipolaridad. Esta metodología está basada en el uso de extensiones continuas 4-valoradas sobre pares de evidencia positiva y negativa (o bipolares). Estas extensiones permiten definir rigurosamente la semántica, en términos cognitivos o epistémicos, de algunos formalismos habituales, diferentes a la luz de este método y que, sin embargo, suelen considerarse equivalentes .

Palabras Clave: Bipolaridad, incertidumbre borrosa o lingüística, intuicionismo

1 INTRODUCCIÓN

El desarrollo de la inteligencia artificial durante la segunda mitad del siglo pasado ha conllevado un notable esfuerzo científico y tecnológico orientado a replicar y plasmar en procesos automáticos (i.e., informáticos) algunas de las capacidades y habilidades características de la mente humana.

En tanto que la habilidad para el manejo de lenguajes complejos es quizás la faceta esencial de la naturaleza humana, el tratamiento del lenguaje natural ha sido (y es) uno de los retos fundamentales a enfrentar dentro de este contexto.

La aparición de las nociones de conjunto y lógica *borrosos* (*fuzzy set* y *fuzzy logic*, [19]) es sin duda uno de los hitos claves en relación con el tratamiento informático del lenguaje natural, en tanto que permiten una modelización matemática rigurosa de la imprecisión y ambigüedad inherentes al lenguaje humano. Recordemos que muchos conceptos (i.e. palabras o predicados) P, como *bueno* o *joven*, no tienen una definición totalmente precisa, de manera que es habitual la aparición de incertidumbre y/o dudas en relación a si determinados objetos x del universo de discurso X considerado (e.g. *alternativas* de un problema de decisión o *edades* de posibles clientes) verifican o no esos predicados. En este marco, por medio de la asignación de grados continuos de pertenencia o verificación $\mu_P(x) \in [0,1]$, la lógica borrosa permite flexibilizar y salvar las dificultades de la lógica clásica binaria para afirmar la ocurrencia de $P(x)$ o $\neg P(x)$.

Paralelamente, la psicología cognitiva ha ido señalando [4,12] que el razonamiento humano tiende a analizar y organizar la realidad (e.g. una decisión a tomar) en términos de las relaciones positivas (de distensión, placer, beneficio, etc.) y/o negativas (de tensión, dolor, pérdida, etc.) que, a través de los aprendizajes pertinentes, guardan los objetos y la información disponible con respecto a la emotividad o racionalidad del individuo (e.g., una alternativa puede ser *buena* respecto a ciertos criterios y *mala* respecto de otros). De este modo, la realidad mental es construida en base al carácter y la relevancia afectivos de los objetos en consideración, o en otras palabras, en términos de pares de *polos de referencia P/Q*, como *bueno/malo* o *verdadero/falso*, que organizan y dan relevancia a la información disponible. Así pues, lo relevante no es saber si un objeto x es *bueno* ($P(x)$) o *no-bueno* ($\neg P(x)$), sino si es *bueno* ($P(x)$) o *malo* ($Q(x)$).

Esta coexistencia de pares de polos opuestos de referencia semántica suele ser referida en la literatura cognitiva como *bipolaridad*, aunque es preciso señalar que algunos autores [1] han usado también el término *intuicionismo* en relación a estas ideas.

Además, diferentes *tipos* de incertidumbre lingüística o borrosa (*fuzziness*) [20] y de bipolaridad [8] han sido definidos y estudiados con el objetivo de generalizar y ampliar el marco de aplicación de estas nociones seminales.

Así, mientras que el tratamiento habitual (o de tipo 1, F1) de la incertidumbre borrosa permite evaluar la imprecisión lingüística de una manera precisa (asignando valores de verdad graduables pero precisos $\mu_P(x) \in [0,1]$ a la proposición $P(x)$), la incertidumbre borrosa de tipo 2 (F2) hace referencia a la posibilidad de medir esa imprecisión de forma *imprecisa* (asignando conjuntos borrosos de tipo 1 como grados de verdad de $P(x)$).

De manera similar, en tanto que la bipolaridad de tipo 1 (B1) asume que la información negativa viene dada por la negación o complementación de la positiva, la bipolaridad de tipo 2 (B2) permite que la relación entre los polos no sea tan simple (e.g. *malo* ≠ *no-bueno*), lo que suele conllevar la necesidad de evaluar el par $(\mu_P(x), \mu_Q(x))$ para capturar toda la información relevante.

Nótese que, en tanto que tratan diferentes tipos de incertidumbre, las nociones de incertidumbre borrosa y bipolaridad parecen a primera vista independientes y complementarias, de modo que, en principio, un formalismo B2 puede ser asociado a una incertidumbre borrosa de tipo F1 o F2 (o incluso ser nítido), e igualmente un marco F2 puede estar asociado a un formalismo B1 o B2.

Sin embargo, el descubrimiento de la equivalencia formal [5,6] entre los conjuntos borrosos *intuicionistas* de Atanassov (AFS [1]), originalmente ideados como objetos de tipo F1 y B2 y los conjuntos borrosos intervalo-valorados (IVFS [11]), en principio concebidos como objetos F2 y B1, ha conducido a identificar esos marcos en cierta medida, lo cual ha acaecido no sin algunas disputas y enfrentamientos [3,7] entre algunos autores y escuelas.

El objetivo de este trabajo es arrojar algo de luz sobre las relaciones entre bipolaridad e incertidumbre borrosa desde una perspectiva diferente. Con este fin, se utilizará la noción de extensión 4-valorada (que recuerda claramente a la de estructura de preferencia [9, 10]) para definir y comparar rigurosamente la semántica y la estructura lógica subyacente de las posibles combinaciones de tipos de bipolaridad e incertidumbre borrosa. Esto permitirá separar y distinguir los AFS de los IVFS de una manera práctica, mostrando a su vez la independencia de esas nociones.

2 TIPOS DE INCERTIDUMBRE BORROSA

La teoría de conjuntos borrosos [19] permite que los objetos de un universo de discurso *X* verifiquen de forma parcial un predicado impreciso *P*. Para ello, la función característica del conjunto $P = \{x \in X \mid P(x)\}$ es generalizada mediante una *función de pertenencia* $\mu_P : X \to [0,1]$, de modo que $\mu_P(x) \in [0,1]$ indica el grado en que el objeto $x \in X$ verifica el predicado *P*. Así, la semántica (o el uso) de *P* sobre *X* es modelada a través de un conjunto borroso de tipo 1 (T1FS) $P = \{(x, \mu_P(x)) \mid x \in X\}$.

En respuesta a algunas críticas que señalaban la posible complejidad de asignar grados de pertenencia totalmente precisos, L.A. Zadeh introdujo en [20] la noción de incertidumbre borrosa de tipo *n*, permitiendo evaluar la veracidad de la proposición $P(x)$ por medio de un conjunto borroso de tipo *n*-1. Particularmente, si $F([0,1])$ denota el conjunto de todos los conjuntos borrosos de tipo 1 sobre [0,1], un conjunto borroso *P* de tipo 2 (T2FS) se asocia a una función de pertenencia $\psi_P : X \to F([0,1])$, de manera que $\psi_P(x) : [0,1] \to [0,1]$ expresa la *plausibilidad* de la proposición "*x* verifica *P* con grado de verdad μ" para cada grado de verdad $\mu \in [0,1]$. Así se introduce la imprecisión en la evaluación de los grados de verdad de tipo 1 $\mu_P(x)$ y, en general, los tipos de incertidumbre borrosa más elevados (i.e. tal que $n>1$) permiten introducir una mayor imprecisión en la evaluación de los grados de verdad.

Los conjuntos borrosos intervalo-valorados (IVFS, [11]) son quizás la clase más simple y común de conjuntos borrosos de tipo 2, en tanto que asignan a cada objeto *x* y predicado *P* un intervalo $[\mu_L(x), \mu_U(x)]_P$ como valores (igualmente y totalmente) plausibles del grado de verdad $\mu_P(x)$. Por tanto, el espacio de valuación de los IVFS viene dado por el conjunto $L^I = \{[a,b] \subseteq [0,1] \mid a \le b\} \subset F([0,1])$. Además, a mayor amplitud del intervalo $[\mu_L(x), \mu_U(x)]$, mayor es la incertidumbre asociada a él, de modo que su longitud $u(x) = \mu_U(x) - \mu_L(x)$ suele verse como el grado de *ignorancia* asociada a esa evaluación. Así, $u(x) = 0$ es cierto si, y solo si (sii), no hay incertidumbre sobre el grado $\mu_P(x)$ en que *x* verifica *P*, esto es, sii $\mu_L(x) = \mu_P(x) = \mu_U(x)$.

Los operadores lógicos clásicos, como la negación, la disyunción o la conjunción pueden ser generalizados al caso difuso mediante diversos operadores borrosos. Así, la negación habitual para conjuntos borrosos de tipo 1 está dada por $n(\mu) = 1 - \mu$, de modo que $\mu_{\neg P}(x) = n(\mu_P(x))$, con $\neg P$ =*no-P*. La negación habitual sobre los IVFS vienen dada por

$$n_{L^I}([\mu_L(x), \mu_U(x)]_P) = [1 - \mu_U(x), 1 - \mu_L(x)]_{\neg P}.$$

Similarmente, t-normas y t-conormas [17] suelen tomarse como operadores para la conjunción y la disyunción, respectivamente. Ambos tipos de conectores están relacionados a través de la negación, de manera que si *T* es una t-norma, entonces $S(\mu_P, \mu_Q) = n[T(n(\mu_P), n(\mu_Q))]$ es una t-conorma. Para T1FS, ejemplos habituales de operadores así relacionados son la t-norma y conorma de Lukasiewicz, respectivamente dados, para $a,b \in [0,1]$, por $a \odot b = \max\{a+b-1, 0\}$ y $a \oplus b = \min\{a+b, 1\}$, o también la t-norma $a \wedge b = \min\{a,b\}$ y su respectiva t-conorma $a \vee b = \max\{a,b\}$. Para IVFS, estos operadores son extendidos de la siguiente forma:

$$[a,b] \odot_I [c,d] = [a \odot c, b \odot d], [a,b] \oplus_I [c,d] = [a \oplus c, b \oplus d],$$

$[a,b] \wedge_I [c,d] = [a \wedge c, b \wedge d], [a,b] \vee_I [c,d] = [a \vee c, b \vee d]$.

Por otro lado, nótese que la lógica borrosa parece infravalorar la noción de información negativa, en tanto que asume que la falsedad de $P(x)$ ha de ser igual a la verdad de $\neg P(x)$, i.e., $\mu_{\neg P}(x)$, aunque este grado de falsedad puede ser de naturaleza diferente a la información negativa. En cualquier caso, es claro que la lógica borrosa no considera una evaluación independiente junto a los grados (precisos o no) $\mu_P(x)$ o $\psi_P(x)$.

3 TIPOS DE BIPOLARIDAD

Básicamente, los modelos de representación bipolares asumen la existencia de un par de polos de referencia semántica *P*/*Q*, tales como *verdadero*/*falso* o *bueno*/*malo*, los cuales proporcionan marcadores absolutos que confieren a la información su carácter intrínsicamente positivo o negativo. La información sin carácter positivo ni negativo es entonces considerada irrelevante o *neutral* en términos de esas referencias, del mismo modo que la información con carácter simultáneamente positivo y negativo es considerada como *conflictiva* o contradictoria. Así, los polos *P*/*Q* organizan y dan relevancia a la información disponible, de manera que la información positiva para uno de ellos es considerada como negativa para el otro.

Si la relación entre los polos está dada por la negación, i.e. si $Q = no\text{-}P = \neg P$ (e.g. *malo=no-bueno*), entonces una única evaluación $\mu_P(x) \in [0,1]$ (o $\mu_P(x) \in L^I$) es suficiente para capturar toda la información relevante en términos de la polaridad *P*/*Q*, ya que entonces $\mu_Q(x) = \mu_{\neg P}(x) = 1 - \mu_P(x)$ y por lo tanto la información negativa es solo la negación de la positiva. El término bipolaridad de tipo 1 (B1, [8]) es habitualmente usado para referirse a esta situación, y nótese que los conjuntos borrosos (tanto F1 como F2) y la lógica clásica pertenecen a esta categoría.

Sin embargo, si la relación entre *P* y *Q* es más compleja (e.g. $malo \neq no\text{-}bueno$), entonces es normalmente necesario evaluar también $\mu_Q(x)$ para capturar toda la información relevante sobre $P(x)$. En particular, es posible usar dos escalas L_P^+ y L_P^- para evaluar, respectivamente, $\mu_P(x)$ y $\mu_Q(x)$, de manera que el par $(\mu^+(x), \mu^-(x))_P = (\mu_P(x), \mu_Q(x)) \in L_P^+ \times L_P^-$ mide de manera conjunta el grado de información positiva y negativa sobre $P(x)$. Por otro lado, si se asume que la relación entre los polos es simétrica (aunque es posible eliminar esta asunción, ver [16]), entonces

$(\mu^+(x), \mu^-(x))_Q = (\mu_Q(x), \mu_P(x)) \in L_Q^+ \times L_Q^- = L_P^- \times L_P^+$.

Estos pares de evidencia son típicos de la bipolaridad de tipo 2 (B2, [8]), y nótese que, en tanto que *P* y *Q* no son necesariamente disjuntos, los modelos B2 admiten en principio 4 casos posibles respecto de la relación entre un objeto *x* y una polaridad *P*/*Q*:

- *Veracidad positiva t*: *x* verifica *P* pero no *Q*, i.e., la información es positiva y no negativa. El par $(\mu^+, \mu^-) = (1,0)$ es el representante extremo (o nítido) de esta situación.
- *Veracidad negativa* (o *falsedad positiva) f*: *x* cumple *Q* pero no verifica *P*, luego la información es negativa y no positiva. El par extremo es $(\mu^+, \mu^-) = (0,1)$.
- *Irrelevancia o neutralidad i: x* no verifica *P* ni *Q*, por lo que la información es no positiva y no negativa y en el caso extremo se tiene $(\mu^+, \mu^-) = (0,0)$;
- *Conflicto k: x* verifica *P* y *Q* simultáneamente, luego la información es positiva y negativa, como en el caso extremo $(\mu^+, \mu^-) = (1,1)$.

Es importante señalar que los casos tercero y cuarto no son posibles en un marco B1, en el que *P* y *Q* están estrechamente ligados a través de la negación. Si en un marco B2 la relación entre los polos es restringida, el tercer o cuarto caso pueden dejar de ser posibles, pero no simultáneamente, ya que entonces ese marco sería B1.

Es interesante examinar con más detalle el caso de los AFS [1]: estos asignan a $P(x)$ un grado de pertenencia $\mu_P(x)$ así como un grado de *no-pertenencia* $\mu_{\neg P}(x)$, cumpliendo que $\mu_P(x), \mu_{\neg P}(x) \in [0,1]$ y $\mu_P(x) + \mu_{\neg P}(x) \leq 1$. Por tanto, el espacio de valuación de los AFS es el conjunto $L^* = \{(x,y) \in [0,1]^2 \mid x + y \leq 1\} \subset [0,1]^2$, y nótese que la misma escala $L^+ = L^- = [0,1]$ es usada para ambas evaluaciones.

Aunque la presencia de dos evaluaciones hace pensar que los AFS pertenecen a un marco B2, lo cierto es que, implícitamente, Atanassov trata un marco B1 (en el que la polaridad relevante es $P/\neg P$) como un marco B2, asumiendo que la verdad de $\neg P(x)$ no puede ser directamente obtenida a partir de la de $P(x)$, por tanto requiriendo una evaluación independiente $\mu_{\neg P}(x) = \nu_P(x)$. En tanto que $\mu_P(x) + \nu_P(x) \leq 1$, se obtiene que la disyunción de P y su negación (en términos de la t-conorma de Lukasiewicz $\oplus$) puede no ser una tautología, violando por ello la ley del tercio excluso ($P \vee \neg P$), en consonancia con la intención de Atanassov de formular un modelo *intuicionista* [1].

Sin embargo, otras estructuras lógicas borrosas también incumplen esa ley sin considerar una negación independiente (e.g. tomando la t-conorma máximo $\vee$ junto a $\mu_{\neg P}(x) = 1 - \mu_P(x)$). Por ello, el significado

concreto de los AFS es relativamente oscuro, pues muestran la semántica de un marco B1 mientras que formalmente se asemejan a un marco B2. Además, en tanto que el modelo de Atanassov nunca proporciona una definición clara de su estructuras lógica y semántica subyacente (ver [12]), no hubo razones sólidas para separar los AFS de los IVFS cuando se demostró la equivalencia [5,6] de ambos formalismos (mediante el isomorfismo $\Phi: L^* \to L^I$ tal que

$$\Phi(\langle \mu_P(x), \mu_{\neg P}(x) \rangle) = [\mu_P(x), 1 - \mu_{\neg P}(x)]).$$

Esta equivalencia produjo una fuerte controversia (ver [3,7]) entre Atanassov y sus partidarios, por un lado, que defendieron la validez de los AFS como objetos B2 (o al menos su validez como modelo *intuicionista*) diferentes de los IVFS, y por otro lado, una importante parte de la comunidad *fuzzy*, que consideraba que el significado intuicionista de los AFS no estaba en absoluto claro, y, además, en tanto que los IVFS fueron formulados previamente a los AFS, la relevancia de estos como objetos bipolares debía ser reducida debido a tal equivalencia (a pesar de que los IVFS fueron originalmente concebidos como objetos B1).

En nuestra opinión, la semántica original de los AFS (sea cual sea) no es de tipo B2. Sin embargo, si ν es interpretada como la función de pertenencia de un polo Q tal que $Q \neq \neg P$, i.e. $\nu_P(x) = \mu_Q(x)$, y se mantiene la restricción $\mu_P(x) + \nu_P(x) \leq 1$, entonces se obtiene fácilmente una semántica B2 (con una relación particular entre los polos). Nuestra opinión es que estos *nuevos* objetos, (que denominaremos AFS *bipolares*, BAFS) no son semánticamente equivalentes a los IVFS. Para sostener esta afirmación, en la próxima sección definiremos formalmente la semántica de los objetos B2, mostrando que la semántica de los BAFS (y en general la de los formalismos F1 y B2) difiere de la de los IVFS (y en general de la semántica de los formalismos F2 y B1). Esta diferenciación semántica será introducida mediante la noción de extensión 4-valorada.

4 EXTENSIONES 4-VALORADAS

Como se señaló en la sección anterior, en un marco B2 la información positiva y negativa proporcionada por los pares $(\mu^+, \mu^-) \in \{0,1\}^2$ conduce de forma natural a cuatro casos que reflejan el significado o semántica de esos pares en términos cognitivos. Esto es, la descripción de la información contenida en los pares se traduce en 4 diferentes *estados epistémicos*, que reflejan la situación de conocimiento, en términos de una polaridad *P*/*Q*, a la que aboca la evidencia disponible sobre un objeto *x*.

Cuando la evidencia es evaluada de forma borrosa, esto es, cuando $(\mu^+, \mu^-) \in [0,1]^2$ ($o \in (L^I)^2$), la traducción de los pares en estados epistémicos es más compleja, en tanto que un par puede ser parcialmente compatible con varios de esos 4 estados. En otras palabras, en un contexto borroso esos cuatro estados son también graduales, y es entonces preciso proporcionar una extensión continua 4-valorada de esos pares que especifique la semántica de los pares no extremos. Aunque es posible definir concretamente estas extensiones de diversas maneras (dando lugar a diferentes semánticas), en este trabajo se opta por una extensión que generaliza las nociones de verdad y falsedad habituales de la lógica borrosa.

Así pues, suponiendo que $L^+ = L^- = L$, $L = [0,1]$ ó L^I, con cada par $(\mu^+, \mu^-) \in L^+ \times L^-$ se asocia una *matriz de evidencia*

$$EM = \begin{bmatrix} t(\mu^+, \mu^-) & i(\mu^+, \mu^-) \\ k(\mu^+, \mu^-) & f(\mu^+, \mu^-) \end{bmatrix},$$

obtenida mediante las fórmulas

$$t(\mu^+, \mu^-) = \mu^+ \wedge \neg\mu^-,\ f(\mu^+, \mu^-) = \neg\mu^+ \wedge \mu^-, \qquad (1)$$

$$i(\mu^+, \mu^-) = \neg\mu^+ \odot \neg\mu^-,\ k(\mu^+, \mu^-) = \mu^+ \odot \mu^-, \qquad (2)$$

donde $\neg\mu$ denota la negación (en *L*) de la evaluación μ. Nótese que $t, f, i, k \in L$, y por tanto *EM* evalúa, en la misma escala en que se valoran μ^+ y μ^-, el grado en que se verifica cada uno de los cuatro estados epistémicos.

La extensión dada en las fórmulas (1)-(2) fue propuesta para el caso *L*=[0,1] en [14] (y ha seguido siendo analizada en [18]), donde se prueba que es la *única* extensión continua basada en t-normas verificando simultáneamente que los estados *i* y *k* son mutuamente exclusivos (ya que si $i>0$ entonces $k=0$ y viceversa) y que $t + f + i + k = 1$ para todo $(\mu^+, \mu^-) \in L^+ \times L^-$. Nótese que intercambiando las posiciones de la t-normas mínimo $\wedge$ y de Lukasiewicz $\odot$ se obtiene otra extensión (utilizada en modelización de preferencias [9]), que también cumple $t + f + i + k = 1$ pero en la que ahora *t* y *f* son exclusivos. Sin embargo, como veremos, esta última extensión no generaliza las nociones clásicas de verdad y falsedad en el caso en que $Q = \neg P$, mientras que la dada por (1)-(2) sí lo hace.

Los resultados en [14] relativos a la extensión dada por (1)-(2) pueden ser generalizados al caso $L = L^I$ mediante el siguiente teorema, cuya demostración puede encontrarse en [15]:

Teorema: Sean $\mu^+ = [\mu_L^+, \mu_U^+]$ y $\mu^- = [\mu_L^-, \mu_U^-]$ ($0 \leq \mu_L^- \leq \mu_U^- \leq 1$, $0 \leq \mu_L^+ \leq \mu_U^+ \leq 1$) las evaluaciones intervalo-valoradas positiva y negativa de la proposición "$x \in P$". Sean también $t = [t_L, t_U]$, $f = [f_L, f_U]$, $i = [i_L, i_U]$ y $k = [k_L, k_U]$ la extensión 4-valorada asociada al par (μ^+, μ^-), obtenida a través de las fórmulas (1)-(2).

Entonces k_L y i_U así como i_L y k_U son mutuamente exclusivos y se cumple que

$$t_L + f_L + i_U + k_U = t_U + f_U + i_L + k_L = 1 + m,$$

donde $m = \min\{u^+, u^-, u_t, u_f, u_u, u_k\}$ y $u^+ = \mu_U^+ - \mu_L^+$, $u^- = \mu_U^- - \mu_L^-$, $u_t = t_U - t_L$, $u_f = f_U - f_L$, $u_i = i_U - i_L$, $u_k = k_U - k_L$ son los grados de ignorancia asociados con cada intervalo.

A continuación, la extensión (1)-(2) será aplicada a diferentes combinaciones de tipos de bipolaridad y de incertidumbre borrosa para ilustrar la semántica particular de cada caso:

Incertidumbre borrosa + bipolaridad de tipo 1 **(F1,B1)**

En este caso es $L=[0,1]$ y $Q = \neg P$, por lo que $\mu^+ = \mu_P$ y $\mu^- = \mu_{\neg P} = n(\mu_P) = 1 - \mu_P$. Esto implica que $t = \mu^+$, $f = \mu^-$ y $k = i = 0$. En particular, $m = 0$ en tanto que no se admite ignorancia respecto de los valores de μ^+, μ^-. Como ejemplo, supóngase un $x \in X$ tal que $\mu_P(x) = 0.6$, esto es, se tiene el par $(\mu^+, \mu^-) = (0.6, 0.4)$ y la matriz

$$EM_1 = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.4 \end{bmatrix}.$$

Por tanto, la extensión (1)-(2) generaliza las nociones de verdad y falsedad de la lógica borrosa (identificadas con μ_P y $\mu_{\neg P}$), y refleja la semántica B1 de este marco, que excluye los estados *k* e *i*.

Incertidumbre borrosa tipo 2+bipolaridad tipo 1 **(F2,B1)**

Supóngase ahora que, en el ejemplo anterior, el valor de $\mu_P(x)$ no se conoce con precisión, sino que se sabe que cae en el intervalo $[\mu_L(x), \mu_U(x)]_P$. Por tanto, se asume aún que $Q = \neg P$, pero ahora $L = L^I$ y entonces $\mu^+ = [\mu_L, \mu_U]_P$ y $\mu^- = n_{L^I}([\mu_L, \mu_U]_P) = [1 - \mu_U, 1 - \mu_L]_{\neg P}$. Así, se obtiene de nuevo que $t = \mu^+$ y $f = \mu^-$ pero ahora $k = i = [0, u]$ y $m=u$, donde $u = \mu_U - \mu_L$ es la ignorancia asociada a μ_P. Por ejemplo, si $\mu^+ = [0.5, 0.7]$ (luego $\mu^- = [0.3, 0.5]$), se obtiene la matriz

$$EM_2 = \begin{bmatrix} [0.5, 0.7] & [0, 0.2] \\ [0, 0.2] & [0.3, 0.5] \end{bmatrix},$$

mostrando que ahora las evaluaciones de la veracidad y falsedad positivas son también afectadas por ignorancia, permitiendo que *i* y *k* sean no nulas. Sin embargo, si la ignorancia se reduce, i.e. al tender *u* a 0, *k* e *i* tienden a [0,0], aproximándose al caso anterior (F1,B1).

Incertidumbre borrosa tipo 1+bipolaridad tipo 2 **(F1,B2)**

Si se permite que la información negativa sobre "$x \in P$" sea independiente de la negación *n* del primer caso (F1,B1), se obtiene un marco B2 en el que $Q \neq \neg P$, $\mu^+ = \mu_P$, $\mu^- = \mu_Q$ y $L^+ = L^- = L = [0,1]$. Luego $m=u=0$, pero *t* (*f*) ya no puede ser identificada con μ^+ (μ^-), y uno de los estados *i*,*k* puede cumplirse. Por ejemplo, si $(\mu^+, \mu^-) = (0.6, 0.7)$, se obtiene la matriz

$$EM_3 = \begin{bmatrix} 0.3 & 0 \\ 0.3 & 0.4 \end{bmatrix},$$

que muestra que $t \neq \mu^+$, $f \neq \mu^-$ e informa de que la información es parcialmente conflictiva (k=0.3), implicando que $t < \mu^+$ y $f < \mu^-$.

Nótese que $(\mu^+, \mu^-) = (0.5, 0.3) \in L^*$ es el par asociado al intervalo $[0.5, 0.7] \in L^I$ del caso previo a través del isomorfismo $\Phi : L^* \to L^I$, y que produce la matriz

$$EM_4 = \begin{bmatrix} 0.5 & 0.2 \\ 0 & 0.3 \end{bmatrix}.$$

Se tiene que $EM_2 \neq EM_4$, y no solo porque EM_2 introduce ignorancia mientras que EM_4 no lo hace. Si la ignorancia *u* se reduce, entonces podría tenerse $\mu^+_{EM_2} = [0.5, 0.5]$, de manera que $t_{EM_2} = [0.5, 0.5] = 0.5 = t_{EM_4}$, pero entonces sería $f_{EM_2} = [0.5, 0.5] \neq 0.3 = f_{EM_4}$ e $i_{EM_2} = [0,0] \neq 0.2 = i_{EM_4}$ ya que $\mu^-_{EM_2} = n_{L^I}(\mu^+_{EM_2}) = [0.5, 0.5]$ por el marco B1 de los IVFS. Sin embargo, los BAFS permiten que μ^- sea independiente de μ^+ al exhibir una bipolaridad de tipo 2, aunque la condición $\mu^+ + \mu^- \leq 1$ implique que k=0. Entonces, es claro que la semántica en términos de estados epistémicos de los BAFS es diferente de la de los IVFS, y su equivalencia debería ser entendida como una coincidencia formal entre dos escalas que pertenecen a universos diferentes, pues $[0,1] \subset L^I \subset F([0,1])$ mientras que $L^* \subset [0,1]^2 \subset L^I \times L^I \subset F([0,1]) \times F([0,1])$.

Incertidumbre borrosa tipo 2+bipolaridad tipo 2 **(F2,B2)**

Finalmente, supóngase que, en un marco B2, se permite ignorancia sobre las evaluaciones μ^+, μ^-, de modo que se obtiene un par $(\mu^+, \mu^-) \in L^I \times L^I$. Nótese que este caso generaliza todos los anteriores y coincide con las hipótesis del teorema anterior. Entonces, si $\mu^+ = [0.5, 0.7]$ y $\mu^- = [0.6, 0.9]$, se obtiene la matriz

$$EM_5 = \begin{bmatrix} [0.1, 0.4] & [0,0] \\ [0.1, 0.6] & [0.3, 0.5] \end{bmatrix},$$

y obsérvese que ahora cada uno de sus elementos tiene asociada un grado de ignorancia diferente (y m=0), y que estos resultados extienden los de EM_1, EM_2 y EM_3. Nótese también que u^+ o u^- pueden ser reducidas sin afectar a la otra, y en el caso límite en que $u^+ = u^- = 0$, se alcanza un marco F1 y B2. Un ejemplo de formalismo B2 y F2 son los IVFS de Atanassov (AIVFS, ver [2]), aunque es válido el mismo argumento de la Sección 3 para afirmar que estos no pueden ser vistos como objetos B2 a menos que se asuma la existencia de un polo negativo $Q \neq \neg P$ (dando lugar a AIVFS *bipolares* o BAIVFS).

5 CONCLUSIONES

Las nociones de incertidumbre borrosa y de bipolaridad han sido revisadas y analizadas. En tanto fueron desarrolladas para afrontar la modelización de características distintas del razonamiento humano, ambas nociones son de hecho complementarias e independientes. Un formalismo F2 puede ser B2 o no serlo, y recíprocamente un marco B2 puede permitir ignorancia (siendo entonces también una marco F2) o no. En particular, esta diferencia se ha hecho evidente en este trabajo, clarificando algunos aspectos de las relaciones entre los formalismos B1,F2 y los B2,F1. De manera implícita, este trabajo muestra que existe cierta diferencia entre las nociones de *intuicionismo* (en el sentido de usar una negación sub-aditiva $n(\mu) \leq 1-\mu$) y de bipolaridad. Sin embargo, si se asume la presencia de un polo negativo $Q \neq \neg P$, entonces los BAFS y BAIVFS constituyen formalismos relevantes dentro de un marco B2. La siguiente Tabla 1 presenta una tipología de los casos y formalismos estudiados en este trabajo, los cuales muestran semánticas diferentes en términos de los estados cognitivos o epistémicos a los que conducen.

Tabla 1. Combinaciones de tipos 1 y 2 de incertidumbre borrosa y bipolaridad.

Fuzziness / Bipolaridad	Tipo 1 (F1)	Tipo 2 (F2)
Tipo 1 (B1)	T1FS	T2FS, IVFS
Tipo 2 (B2)	BAFS, BFS	BT2FS, BAIVFS

En el futuro, este trabajo se extenderá en diversas direcciones: se introducirán retículos (*lattices*) para obtener un marco lógico más general y riguroso. También, se explorarán con más detenimiento las relaciones entre bipolaridad, intuicionismo, el uso de negaciones sub-aditivas y los operadores de disimilaridad [16]. Además, se estudiará el posible uso de los formalismos F2 para combinar la incertidumbre borrosa con otras teorías de la incertidumbre (probabilística, posibilística, etc.).

Agradecimientos

Este trabajo ha sido realizado en el marco del proyecto TIN2009-07901.

Referencias

[1] Atanassov KT (1986) Intuitionistic Fuzzy-Sets, F Set Syst, 20 (1) 87-96

[2] Atanassov KT, Gargov G (1989) Interval Valued Intuitionistic Fuzzy Sets, F Set Syst, 31(3) 343-349

[3] Atanassov KT (2005) Answer to D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk and H. Prade's paper "Terminological difficulties in fuzzy set theory - the case of "Intuitionistic fuzzy sets", Fuzzy Set Syst, 156 (3) 496-499

[4] Cacioppo JT, GardnerWL, Berntson GG (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. Pers Soc Psychol Rev 1:3–25.

[5] Cornelis C, Atanassov K, Kerre E, Intuitionistic fuzzy sets and interval-valued fuzzy sets: a critical comparison, Proc. Eusflat'03, 2003, 159–163

[6] Deschrijver G, Kerre E (2004) On the relationship between some extensions of fuzzy set theory, Fuzzy Set Syst 133 227–235.

[7] Dubois D, Gottwald S, Hajek P, Kacprzyk J, Prade H (2005) Terminological difficulties in fuzzy set theory - The case of "Intuitionistic Fuzzy Sets", Fuzzy Set Syst, 156 (3) 485-491

[8] Dubois D, Prade H (2008) An introduction to bipolar representations of information and preference, Int J Intell Syst, 23 (8) 866-877

[9] Fodor J, Roubens M (1994) Fuzzy preference modelling and multicriteria decision support, Kluwer Academic, Dordrecht ; Boston

[10] Franco C, Rodríguez T, Montero J. (2010) Information measures over intuitionistic four valued fuzzy preferences. Proceedings IEEE-WCCI Conference, Barcelona, July 18-23, pages 1971-1978

[11] Grattan-Guinness I (1975) Fuzzy membership mapped onto interval and many-valued quantities, Z. Math. Logik Grundlag. Mathe., 22 149-160

[12] Medin D, Schwanenflugel P (1981) Linear separability in classification learning, J. of Experimental Psychology: Human Learning & Memory, 7 355-368

[13] Montero J, Gómez D, Bustince H (2007) On the relevance of some families of fuzzy sets, Fuzzy Set Syst, 158 (22): 2439-2442

[14] Öztürk M, Tsoukiàs A (2007) Modelling uncertain positive and negative reasons in decision aiding, Decision Support Systems, 43 (4): 1512-1526

[15] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) Eurofuse

[16] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) An axiomatic approach to the notion of semantic antagonism, Procs. of the IFSA-AFSS'11, FT104-1/6

[17] Schweizer B, Sklar A (1983) Probabilistic Metric Spaces. North–Holland/Elsevier, New York.

[18] Turunen E, Ozturk M, Tsoukiàs A (2010) Paraconsistent semantics for Pavelka style fuzzy sentential logic Fuzzy Set Syst, 161 (14) 1926-1940

[19] Zadeh LA (1965) Fuzzy Sets, Information and Control, 8 (3) 338-353

[20] Zadeh, LA (1975) The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, Information Sciences 9 (1) 43-80

FINGRAMS: UNA NUEVA HERRAMIENTA PARA ANÁLISIS VISUAL DE SISTEMAS FUZZY

Jose M. Alonso, **David P. Pancho**, **O. Cordón**, **A. Quirin**, **L. Magdalena**

European Centre for Soft Computing, Mieres, 33600, Spain
{jose.alonso,david.perez,oscar.cordon,arnaud.quirin,luis.magdalena}@softcomputing.es

Resumen

Desde las ideas originales de Zadeh y Mamdani, la interpretabilidad se reconoce como una de las características más apreciadas y valoradas de los sistemas fuzzy, representando la habilidad de éstos para formalizar el comportamiento de un sistema real de una forma comprensible. Este trabajo propone la creación de mapas de inferencia de sistemas fuzzy, llamados inferenciogramas fuzzy (*fingrams*) por analogía con los mapas de la ciencia (*scientograms*) utilizados para visualizar la estructura de la ciencia. Los *fingrams* muestran gráficamente la interacción entre reglas a nivel de inferencia, en términos de co-disparo de reglas, es decir, reglas que se disparan simultáneamente dada una misma entrada. El análisis de *fingrams* ofrece múltiples posibilidades: detección de redundancias y/o inconsistencias entre reglas, descubrimiento de las reglas más significativas, evaluación de la comprensibilidad de sistemas fuzzy, etc.

Palabras Clave: Diseño de sistemas fuzzy, Representación y análisis visual, Co-disparo de reglas, Interpretabilidad.

1 INTRODUCCIÓN

La interpretabilidad de un sistema fuzzy depende de la habilidad del usuario final para interpretar su descripción lingüística con el objetivo de conocer el comportamiento del sistema. En consecuencia, caracterizar y evaluar la interpretabilidad es una tarea subjetiva que depende de los conocimientos (formación, experiencia, preferencias, etc.) de la persona que realiza la evaluación [3].

Ha de remarcarse que la interpretabilidad es una característica distintiva de los sistemas fuzzy muy valorada en la mayoría de las aplicaciones. Ésta se convierte en un requisito esencial en aquellas aplicaciones que involucran una interacción continua con personas. Por ejemplo, los sistemas de apoyo a la toma de decisiones en medicina [14] han de ser fácilmente entendibles, por médicos y pacientes, con el objetivo de ser de confianza, es decir, ampliamente aceptados y usados satisfactoriamente.

Desafortunadamente, los sistemas fuzzy no son interpretables *per se*. Las tareas de diseño deben realizarse cuidadosamente para obtener sistemas interpretables [8]. No obstante, aunque ha existido un gran interés durante la última década por definir, caracterizar y evaluar la interpretabilidad de sistemas fuzzy, aún queda mucho trabajo por hacer [1]. El análisis de interpretabilidad debe tener en cuenta tanto la legibilidad de la descripción del sistema (vista estática global) como la comprensión del comportamiento del sistema en ejecución (vista dinámica local). Obviamente, comprender el comportamiento de un sistema fuzzy a partir de su descripción lingüística es una tarea complicada que exige profundizar en la inferencia más allá del análisis de legibilidad.

Este trabajo se centra en el análisis relativo a la que hemos llamado vista dinámica local de un sistema fuzzy. Para ello, se propone una nueva metodología de visualización de sistemas fuzzy a nivel de inferencia, basada esencialmente en la adaptación de técnicas existentes para la visualización de redes de información científica.

El resto del documento se organiza de la siguiente forma. La sección 2 presenta un breve repaso de metodologías para representación visual y análisis de sistemas fuzzy. Además, incluye una pequeña introducción a las técnicas más extendidas para diseñar mapas visuales de la ciencia. La sección 3 introduce la nueva metodología propuesta en este trabajo. El enfoque general se particulariza para el análisis de clasificadores basados en reglas fuzzy. La sección 4 ilustra con un ejemplo sencillo cómo aplicar la nueva metodología. Finalmente, las conclusiones y trabajos futuros son esbozados en la sección 5.

2 TRABAJOS RELACIONADOS

2.1 ANÁLISIS VISUAL DE SISTEMAS BASADOS EN REGLAS FUZZY

En [15] se revisan las metodologías existentes para representación gráfica de sistemas fuzzy, tanto en 2D como en 3D. Además, se analizan los requisitos de visualización de datos, particiones y reglas fuzzy según distintas alternativas que dependen del usuario final. Dichos requisitos cambian en función de las tareas a realizar: (1) exploración interactiva; (2) exploración automática asistida por ordenador; (3) recepción de observaciones de los usuarios; (4) captura y adaptación a los perfiles de usuario.

Podemos concluir que hasta ahora los trabajos más relevantes para obtener representaciones visuales de reglas fuzzy multidimesionales son aquellos desarrollados por Berthold et al. [5, 11]. Estos autores realizan un mapeo desde un espacio de características multidimensional a un espacio de dos dimensiones que mantiene las distancias entre parejas de reglas, mostrando una aproximación de la dispersión y solapamiento entre reglas. Como resultado, es posible visualizar y explorar bases de reglas fuzzy multidimensionales en una representación gráfica 2D. Los autores exponen que esa representación permite un análisis de exploración intuitivo e intepretable. Sin embargo, la complejidad del análisis crece exponencialmente con el número de características y reglas a mostrar. Por ello, en problemas complejos con muchas reglas la interpretación de los gráficos resultantes no es sencilla.

2.2 SCIENTOGRAMS: DISEÑO Y APLICACIONES

El término *scientogram* se utiliza en la literatura especializada para hacer referencia a los mapas visuales de la ciencia, es decir, representaciones visuales de dominios científicos en forma de redes sociales. En los últimos años, el número de publicaciones científicas multidisciplinares ha crecido rápidamente, e identificar relaciones entre distintos dominios se ha convertido en una necesidad. Afortunadamente, los recientes avances en visualización por ordenador [6, 7, 18] hacen factibles representaciones visuales rápidas y de alta calidad para dominios científicos muy amplios.

Vargas-Quesada y Moya-Anegón [21] propusieron una metodología para crear *scientograms* con la finalidad de ilustrar interacciones entre autores y artículos, a través del análisis de citas y co-citas. La propuesta básica considera las referencias cruzadas de artículos, las cuales representan la frecuencia con la que pares de documentos se citan mutuamente. Es posible hablar acerca de la co-citación de autores, revistas, etc. Obviamente, dependiendo del tipo de co-citación seleccionada, la información que puede ser extraída de los mapas generados es distinta.

La medida estándar de co-citación fue propuesta por Salton y Bergmark [17] y queda definida por la siguiente ecuación:

$$MCN(ij) = \frac{Cc(ij)}{\sqrt{c(i) \cdot c(j)}} \quad (1)$$

donde Cc significa co-citación, c representa la cita, mientras que i y j representan dos entidades distintas (autores, documentos, revistas, categorías, etc).

El algoritmo Pathfinder [10] se ha usado eficientemente para el análisis de redes sociales construidas a partir de entidades de co-citación, alcanzando resultados de alta calidad. Por ejemplo, en psicología (para representar la estructura cognitiva de un sujeto [19]), en desarrollo de software (para depurar sistemas multi-agente [20]), o en cienciometría (para el análisis de grandes dominios científicos [9]). El algoritmo Pathfinder se encarga de podar la red definida a partir de la matriz original de co-citación (que puede verse como la matriz de adyacencia de un grafo) manteniendo sólo aquellos arcos más importantes. Las redes finales, después de ser podadas con Pathfinder, se denominan PFNETs y su intepretación visual es mucho más sencilla.

Existen varios métodos para la visualización automatizada de PFNETs. La familia de métodos conocida como *spring-embedder* es la más ampliamente utilizada en el área de ciencias de la información. Estos métodos asignan coordenadas a los nodos de tal forma que el grafo final sea agradable a la vista, colocando los elementos más importantes en el centro de la imagen. El algoritmo de Kamada-Kawai [13] es uno de los métodos más extendidos para llevar a cabo esta tarea. A partir de una disposición circular de los nodos, genera una red con criterios estéticos de tal manera que se maximiza el uso del espacio disponible, se minimiza el número de solapamientos entre enlaces, se fuerza la separación de nodos, se potencia la construcción de mapas balanceados, etc.

Por último, en relación con el análisis de *scientograms*, observar que como resultado de combinar la co-citación de entidades con PFNETs y con el algoritmo de Kamada-Kawai las representaciones visuales finales son fáciles de interpretar, y se caracterizan porque las entidades que más interactuan con el resto tienden a situarse en el centro.

3 PROPUESTA

Llamamos *fingrams* a mapas visuales que representan gráficamente la interacción entre reglas a nivel de inferencia fuzzy en términos de co-disparo de reglas, es decir, reglas disparadas al mismo tiempo por una entrada dada.

3.1 GENERACIÓN DE FINGRAMS

Los pasos a seguir para construir un *fingram* son:

1. Construir un conjunto de datos experimentales que sea significativo, es decir, que cubra de la forma más homogénea posible todas las situaciones bajo estudio.

2. Crear una matriz cuadrada (M) que caracterice las interacciones entre las N reglas del sistema fuzzy bajo estudio, de acuerdo a la proporción de co-disparo (a_{ij}) de las mismas:

$$M = \begin{pmatrix} 1 & a_{12} & \dots & a_{1N} \\ a_{21} & 1 & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & 1 \end{pmatrix}, a_{ij} = \frac{SFR_{ij}}{\sqrt{FR_i \cdot FR_j}} \quad (2)$$

   donde a_{ij} es una medida de co-disparo de reglas inspirada en la medida de co-citación expresada por la Ecuación (1). SFR_{ij} indica el número de ejemplos (del conjunto de datos experimentales) que disparan simultáneamente las reglas R_i y R_j, mientras que FR_i y FR_j cuentan el número total de ejemplos que disparan las reglas R_i y R_j respectivamente. Nótese que a_{ij} ha sido normalizada y la matriz M es simétrica.

3. Visualizar el *fingram* inicial. Para ello, se construye un grafo, cuya matriz de adyacencia viene dada por M, utilizando la herramienta de software libre Graphviz[1] [12]. La imagen que representa el grafo se pinta asignando coordenadas a los nodos que lo forman con el algoritmo de Kamada-Kawai [13]. Nótese que el *fingram* inicial es normalmente muy denso y complejo de analizar incluso para problemas de tamaño reducido, como se ve en el ejemplo de la Figura 1 que se explicará en detalle en la sección 4.

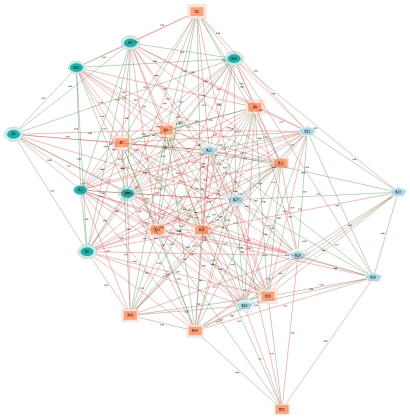


Figura 1: Ejemplo de *fingram* sin podar.

4. Generar una PFNET, es decir, un *fingram* podado con el algoritmo Pathfinder[2]. Debe tenerse en cuenta que los *fingrams* podados mantienen toda la información relevante a nivel global gracias a las propiedades de las PFNETs [10]. En el caso de problemas de clasificación, consideramos dos tipos de poda que dependen del tipo de enlaces (redundancia y/o inconsistencia) sobre los que se actúa. En primer lugar, la PFNET se puede construir a partir del *fingram* inicial completo. Como alternativa, se puede realizar un primer filtrado eliminando todos los enlaces redundantes, y a continuación construir la PFNET teniendo en cuenta sólo las potenciales inconsistencias. Las Figuras 2 y 3, correspondientes al ejemplo que se incluye en la sección 4, muestran las PFNETs resultantes de aplicar ambos tipos de poda sobre el *fingram* inicial de la Figura 1.

[1] Software libre disponible en [http://www.graphviz.org/]

[2] Se ha seleccionado MST-PathFinder [16], una variante del algoritmo Pathfinder publicada recientemente que es capaz de generar grandes mapas de la ciencia en un tiempo cúbico.

3.2 ANÁLISIS DE FINGRAMS

Los *fingrams* muestran de forma escalable la información relacionada con el proceso de inferencia de sistemas fuzzy, gracias a la combinación de co-disparo de reglas, PFNETs, y el algoritmo de Kamada-Kawai. Es importante destacar que nuestra propuesta no está afectada por el problema, bien conocido, de dimensionalidad, característico de los sistemas fuzzy, que implica que el número de reglas crece exponencialmente con el número de entradas. Primero, en los *fingrams* los nodos representan directamente reglas, y no premisas. Segundo, las PFNETs han sido aplicadas satisfactoriamente al análisis de amplios dominios científicos representando miles de entidades co-citadas [21]. Por tanto, los *fingrams* son capaces de mostrar interacciones entre miles de reglas en forma de grafos muy interpretables. Incluso cuando el número de reglas es inmenso los *fingrams* podados pueden ser aún muy útiles, analizados por un diseñador experto.

Tanto los nodos como los arcos que componen un *fingram* representan mucha información útil para el diseñador. Cada nodo representa una regla. Además, el número de líneas alrededor del mismo está directamente relacionado con la complejidad de la regla que representa (una línea indica que la regla tiene dos premisas, dos líneas representan tres premisas, y así sucesivamente). En problemas de clasificación, las reglas que apuntan a las misma clase de salida son representadas por el mismo símbolo (pentágono, rectángulo, elipse, etc.). Por su parte, los enlaces entre nodos representan la información de co-disparo entre reglas. Así, si no hay un enlace entre dos nodos es porque no existen datos experimentales que hagan que las reglas que representan se disparan simultáneamente. En el caso particular de problemas de clasificación, distinguimos dos colores para los enlaces. El color rojo sirve para destacar inconsistencias, es decir, relaciones entre reglas que se activan a la vez pero apuntan a clases de salida diferentes. A su vez, el color verde permite identificar redundancias, es decir, relaciones entre reglas que se activan simultáneamente inferiendo la misma clase de salida.

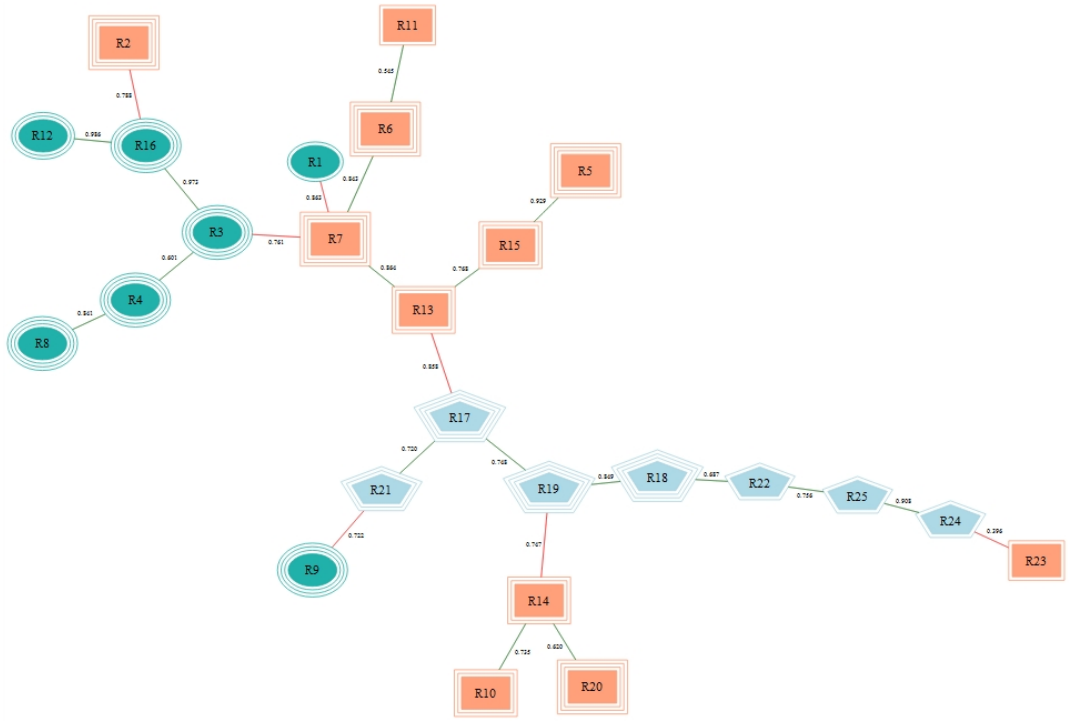


Figura 2: Ejemplo de *fingram* podado (Redundancias e inconsistencias).

Debido a la forma específica de poda y dibujado, las reglas más relevantes (aquellas que se disparan más a menudo) tienden a situarse en el centro de los *fingrams* podados, mientras que las reglas menos importantes aparecen en la periferia. Por lo tanto, la estructura global de los *fingrams* es muy informativa en si misma. Su análisis permite identificar las reglas (o grupos de reglas) más significativas en una base de conocimiento fuzzy. Pero además, en problemas de clasificación el análisis de *fingrams* sirve para detectar redundancias y/o inconsistencias entre reglas.

Con el objetivo de encontrar las reglas más significativas de un sistema fuzzy podemos utilizar las medidas de centralidad habituales en el análisis de *scientograms*. Por ejemplo, el llamado *Grado de centralidad* identifica de forma automática las reglas que concentran el mayor número de enlaces, que como dijimos previamente son aquellas reglas que tienden a localizarse en el centro de los *fingrams* podados. Así, aquellas reglas que correspondan a nodos localizados en la periferia, especialmente aquellas conectadas al resto del grafo a través de enlaces con un peso asociado muy bajo, son buenas candidatas a ser podadas. Esto tiene una interesante ventaja colateral, ya que, al eliminar estas reglas se incrementa la interpretabilidad manteniendo la precisión. Obviamente, es fácil pensar en un procedimiento de simplificación de un sistema fuzzy que consista primero en establecer un ranking de relevancia de las reglas según la información extraída del análisis del *fingram* correspondiente, para después aplicar un método de simplificación lingüística como el propuesto en [4] (o similar), empezando por eliminar las reglas menos relevantes.

En relación con el análisis de redundancias e inconsistencias a nivel de inferencia fuzzy, lo primero que hay que destacar es que la interacción entre reglas es muy difícil de apreciar si únicamente se lee la descripción lingüística del sistema fuzzy. Incluso cuando una base de reglas es completamente consistente a nivel lingüístico, puede presentar algunas redundancias y/o inconsistencias a nivel inferencial debido al proceso de agregación de reglas. Por ejemplo, en problemas de clasificación, puede suceder que varias reglas sean disparadas al mismo tiempo por un mismo vector de entrada y como resultado varias salidas (clases) sean activadas. La situación en que dos clases se activan con grados similares no es deseable, no importando si el sistema es capaz (o no) de obtener la salida correcta, porque una pequeña modificación en los datos de entrada llevará a una salida incorrecta. Un sistema fuzzy que produce muchas situaciones de este tipo es un sistema muy poco robusto que no está bien diseñado y que tiene que ser corregido.

Afortunadamente, un simple análisis visual de los *fingrams* podados sirve para descubir de forma inmediata nodos redundantes (reglas co-disparadas que producen la misma clase de salida) e inconsistentes (reglas disparadas simultáneamente que producen clases de salida distintas). Además, a mayor peso en el vínculo entre nodos, mayor interacción entre reglas, y consecuentemente mayor grado de redundancia o inconsistencia se produce. En problemas de clasificación, las redundancias e inconsistencias han de ser

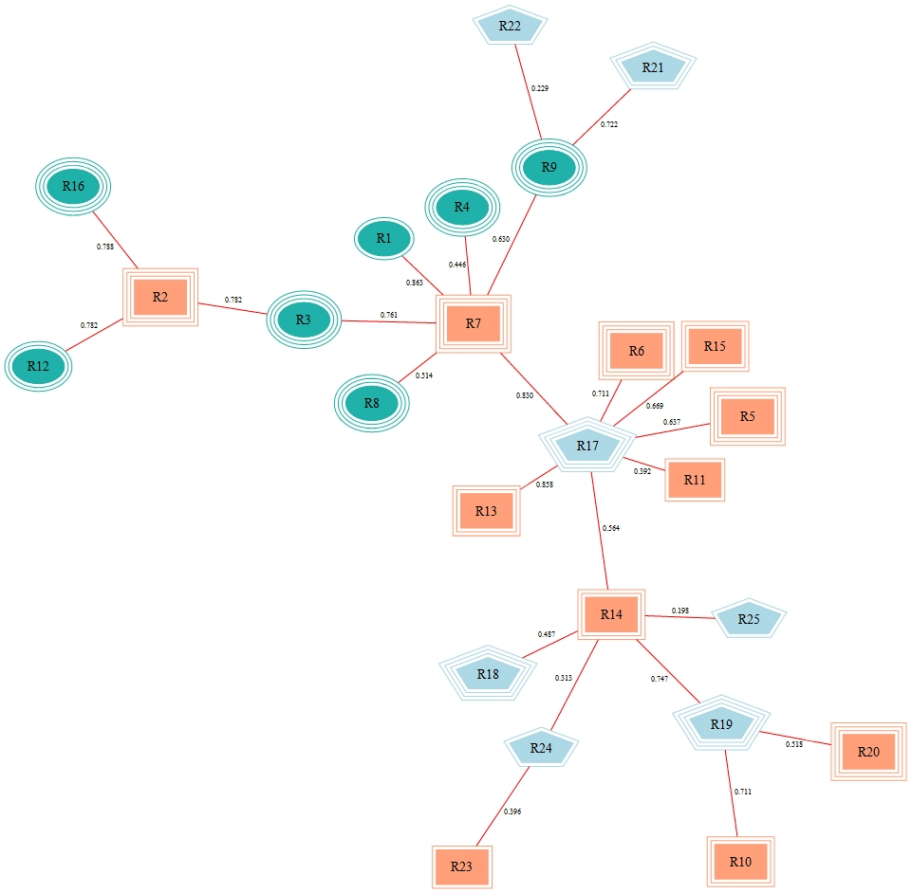


Figura 3: Ejemplo de *fingram* podado (Sólo inconsistencias).

manejadas como conflictos a resolver. En [4], se puede encontrar una explicación detallada de algunos de estos conflictos así como una metodología para detectarlos y corregirlos. Por ejemplo, solucionar redundancias implica eliminar reglas redundantes, lo que contribuye a su vez a mejorar la interpretabilidad.

4 EJEMPLO ILUSTRATIVO

Esta sección explica, paso a paso, cómo generar y analizar los *fingrams* correspondientes a un clasificador fuzzy para WINE. Hemos elegido este problema con fines ilustrativos porque es sencillo y muy conocido. Se trata de distinguir entre tres vinos de la misma región de Italia.

El clasificador fuzzy lo hemos construido siguiendo la metodología de modelado conocida como HILK[3] (Highly Interpretable Linguistic Knowledge) [2, 4]. Se ha seleccionado HILK por estar especialmente pensada para facilitar el diseño de clasificadores fuzzy interpretables. El clasificador diseñado consta de 13 variables de entrada (tantas como atributos caracterizan a WINE) y una salida categórica con 3 clases (una por cada tipo de vino a identificar). Primero, hemos generado particiones fuertes con tres conjuntos fuzzy triangulares para cada entrada. A continuación, hemos construido una base de reglas formada por 25 reglas[4] y un número total de 88 premisas.

Para construir un *fingram* que caracterice, a nivel de inferencia, el clasificador fuzzy que acabamos de describir, hemos aplicado la metodología propuesta en la sección 3:

1. El conjunto de datos experimentales de WINE contiene 178 instancias, cada una compuesta de 13 atributos de entrada y uno de salida. Es accesible desde la web de KEEL[5].
2. Construimos la matriz M que caracteriza la interacción entre las 25 reglas del clasificador, a partir del conjunto de datos descrito en el punto anterior.
3. Usamos Graphviz para generar el *fingram* inicial correspondiente a la matriz M que acabamos de construir (ver la Figura 1).

[3]La metodología HILK está implementada en el software libre GUAJE disponible en [http://www.softcomputing.es/guaje].

[4]Las reglas se han inducido de forma automática con el algoritmo "Fuzzy Decision Tree" proporcionado por GUAJE.

[5]Software libre disponible en [http://sci2s.ugr.es/keel/].

4. Usamos Pathfinder para podar el *fingram* inicial. Como resultado se obtiene la PFNET que se puede ver en la Figura 2. Así mismo, se puede considerar la creación y análisis de un *fingram* podado que sólo considere inconsistencias. Para ello, primero se anulan aquellos valores de la matriz M que relacionan reglas redundantes y posteriormente se ejecuta Pathfinder. El resultado se muestra en la Figura 3.

En la Figura 2 podemos identificar grupos de reglas periféricas que destacan principalmente por ser redundantes. Tales reglas, por ejemplo R8, R11, o R20, son las primeras que deberíamos intentar eliminar. Nótese que para tener en cuenta todas las posibles inconsistencias es necesario recurrir a la Figura 3 ya que algunas de ellas quedan parcialmente ocultas en la Figura 2, especialmente cuando las redundancias son más significativas.

5 CONCLUSIONES

Este trabajo ha introducido una nueva y potente metodología para el análisis de sistemas basados en reglas fuzzy a nivel de inferencia. Se ha implementado una primera versión (sólo para problemas de clasificación) que está integrada en la herramienta de software libre GUAJE:

http://www.softcomputing.es/guaje

En el futuro se validará la metodología con más conjuntos de datos, se buscarán otras métricas de co-disparo capaces de dar información adicional sobre consistencia, generalidad y/o especificidad de reglas, y se extenderá la implementación a problemas de regresión, control, etc.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación Español bajo los proyectos TIN2011-29824-C02-01, TIN2008-06890-C02-01 y TIN2009-07727.

Referencias

[1] J. M. Alonso and L. Magdalena. Editorial: Special issue on interpretable fuzzy systems. *Information Sciences*, 181(20):4331–4339, 2011.

[2] J. M. Alonso and L. Magdalena. HILK++: an interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers. *Soft Computing*, 15(10):1959–1980, 2011.

[3] J. M. Alonso, L. Magdalena, and G. González-Rodríguez. Looking for a good fuzzy system interpretability index: An experimental approach. *International Journal of Approximate Reasoning*, 51:115–134, 2009.

[4] J. M. Alonso, L. Magdalena, and S. Guillaume. HILK: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *International Journal of Intelligent Systems*, 23(7):761–794, 2008.

[5] M. R. Berthold and L. O. Hall. Visualizing fuzzy points in parallel coordinates. *IEEE Transactions on Fuzzy Systems*, 11(3):369–374, 2003.

[6] K. Borner, C. Chen, and K. Boyack. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37:179–255, 2003.

[7] K. Boyack, R. Klavans, and K. Borner. Mapping the backbone of science. *Scientometrics*, 64:351–374, 2005.

[8] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. *Interpretability issues in fuzzy modeling*, volume 128. Studies in Fuzziness and Soft Computing, Springer-Verlag, Heidelberg, 2003.

[9] C. Chen. *Information visualization: Beyond the horizon.* Springer, 2004.

[10] D.W. Dearholt and R.W. Schvaneveldt. Properties of pathfinder networks. In R. Schvaneveldt, editor, *Pathfinder associative networks: Studies in knowledge organization*, pages 1–30. Ablex Publishing Corporation, 1990.

[11] T. R. Gabriel, K. Thiel, and M. R. Berthold. Rule visualization based on multi-dimensional scaling. In *FUZZ-IEEE*, pages 66–71, 2006.

[12] E. R. Gansner and S. C. North. An open graph visualization system and its applications to software engineering. *Software - Practice and Experience*, 30(11):1203–1233, 1999.

[13] T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1):7–15, 1989.

[14] D. Nauck and R. Kruse. Obtaining interpretable fuzzy classification rules from medical data. *Artificial Intelligence in Medicine*, 16:149–169, 1999.

[15] B. Pham and R. Brown. Analysis of visualization requirements for fuzzy systems. In *International conference on computer graphics and interactive techniques (GRAPHITE)*, pages 181–187, 2003.

[16] A. Quirin, O. Cordón, J. Santamaría, B. Vargas-Quesada, and F. Moya-Anegón. A new variant of the pathfinder algorithm to generate large visual science maps in cubic time. *Information Processing and Management*, 44:1611–1623, 2008.

[17] G. Salton and D. Bergmark. A citation study of computer science literature. *IEEE Transactions on Professional Communication*, 22:146–158, 1979.

[18] I. Samoylenko, T.-C. Chao, W.-C. Liu, and C.-M. Chen. Visualizing the scientific world and its evolution. *Journal of the American Society for Information Science and Technology*, 57:1461–1469, 2006.

[19] R. W. Schvaneveldt, F. T. Durso, and D. W. Dearholt. Network structures in proximity data. *The psychology of learning and motivation: Advances in research and theory*, (24):249–284, 1989.

[20] E. Serrano, A. Quirin, J. Botia, and O. Cordón. Debugging complex software systems by means of pathfinder networks. *Information Sciences*, 180(5):561–583, 2010.

[21] B. Vargas-Quesada and F. Moya-Anegón. *Visualizing the structure of science*. Springer-Verlag, 2007.

CONTRIBUCIÓN A LA INTERPRETACIÓN DE SISTEMAS MAMDANI

Claudio Moraga

European Centre for Soft Computing; 33600 Mieres, Asturias, claudio.moraga@softcomputing.es

Resumen

La interpretación lingüística de la agregación de conclusiones en un sistema Mamdani presenta especiales dificultades por la presencia de conjuntos borrosos, representando términos lingüísticos, truncados a la altura del grado de satisfacción de sus premisas. Se discute un método basado en transformaciones geométricas y lingüísticas que permiten una interpretación metasemántica.

Palabras Clave: Metasemántica, Sistemás Mamdani

1 INTRODUCCIÓN

En un conjunto de reglas si-entonces de tipo Mamdani, para cada situación concreta de los datos del entorno, la conclusión representa la agregación de términos lingüísticos (mayormente triangulares o trapezoidales) truncados al nivel de satisfacción de las premisas de las reglas correspondientes (ver p. ej. [2], [5]). La interpretación lingüística de términos truncados es un problema mayormente abierto. Posiblemente un primer intento de interpretación se encuentra en [4], donde se aplica el método de la metasemántica [3] a la agregación de conclusiones, cuyos términos lingüísticos tienen representación trapezoidal, tomando en cuenta por una parte, que en la conversión de borroso a nítido [1], la abscisa del centro de gravedad del área de las conclusiones agregadas es considerada como su valor representativo, y por otra, que la aplicación de modificadores lingüísticos a términos representados por conjuntos borrosos, produce conjuntos borrosos con un *aire de familia* [8], [7] común.

El resto del trabajo está estructurado de la siguiente forma. En la próxima sección se presentarán los conceptos fundamentales y, para efectos de completitud, se hará una revisión breve del caso en el cual las alturas de truncado de las conclusiones están ordenadas de forma monótona. Seguidamente, la sección 3 estará dedicada al análisis de conclusiones en sistemás generales de Mamdani. Los Comentarios Finales estarán recogidos en la sección 4 y, por último, los Agradecimientos y las Referencias concluirán el trabajo.

2 EL CASO MÁS SIMPLE

La hipótesis de trabajo en [4] para la interpretación de términos lingüísticos trapezoidales truncados se resume como sigue:
i) Los términos lingüísticos utilizados en las conclusiones de las reglas de un sistema Mamdani tienen una interpretación expresada en etiquetas lingüísticas claras.
ii) Un trapecio truncado puede asociarse al término lingüístico de las conclusiones del cual fue generado, preservando la abscisa del centro de gravedad, por medio de transformaciones geométricas que mantienen el aire de familia.
iii) Utilizando una metasemántica se puede especificar una relación de significado entre el trapecio truncado y el trapecio de origen, modificado.

Definición 1: Metasemántica.
Dado un Predicado, es el uso del lenguaje el que determina su significado –(camino de la semántica)- y su representación como conjunto borroso en un universo apropiado. Si la representación como conjunto borroso sufre modificaciones geométricas, la metasemántica provee la especificación lingüística de una posible secuencia de transformaciones elementales, expresables como modificadores lingüísticos, que relaciona el conjunto borroso modificado, con el original y su significado (ver Figura 1).

Lemma 1 [4]:
Si se preserva la dimensión y posición relativa de las bases de un trapecio y se modifica la altura, la abscisa del centro de gravedad de superficie permanece constante.

[1] El neologismo de mal gusto "defuzzificación", suele aparecer en textos que pretenden ser de "habla hispana".

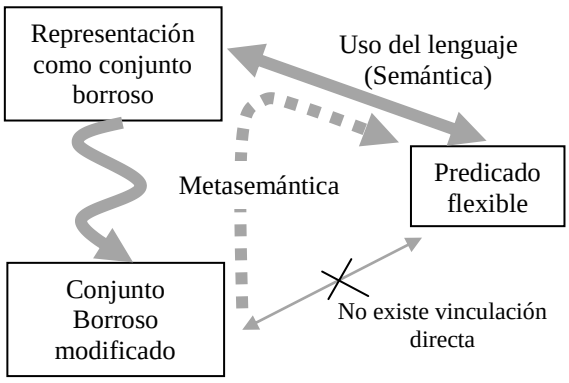


Figura 1: Papel de la semántica y de la metasemántica

Ejemplo 1:

Supóngase que las reglas para conducir el regadío de un tipo especial de cereales, de regadío crítico, fueran las siguientes:

R1. Si la lluvia caida en la noche anterior fuere prácticamente nula, entonces a la mañana siguiente el regadío deberá ser constante.

R2. Si la lluvia caida en la noche anterior fuere moderada, entonces a la mañana siguiente el regadío deberá ser mesurado.

R3. Si la lluvia caida en la noche anterior fuere abundante, entonces a la mañana siguiente el regadío deberá ser mínimo.

La representación gráfica, realizada con la asesoría de agricultores expertos se muestra en la Figura 2 y se considera la situación en la que la medición de la lluvia caída indica "*a*" litros/m^2. En la situación considerada, resulta claro que la lluvia caida, dista de ser considerada *abundante*, con lo cual la tercera regla no tiene aplicación. En cuanto a las otras dos reglas, el agua caida en la noche satisface parcialmente las categorías *moderada* y *nula.* La agregación de las conclusiones acotadas (por los grados de satisfacción de las premisas) se ilustra en la Figura 3, incluyendo la abscisa g_0 de su centro de gravedad, como también g_1, la abscisa más próxima del centro de gravedad de un término lingüístico acotado (sombreado en gris).

Dado que el término lingüístico acotado con abscisa g_1 de centro de gravedad, es el más próximo al centro de gravedad de la agregación de conclusiones acotadas, se toma como referencia para la búsqueda metasemántica.
Según el Lemma 1, si preservando las bases se aumenta la altura de este trapecio hasta alcanzar 1 (pasando a representar un conjunto borroso normal), no cambiará la posición de g_1. El nuevo trapecio con altura 1, tendrá sin embargo, un núcleo algo más ancho que el trapecio representando el término lingüístico "*constante*", el cual le dio origen. Incrementar el núcleo preservando el soporte, es una forma de expresar el modificador lingüístico "más o menos" [4], si bien es formalmente diferente a la originalmente propuesta por L.A. Zadeh [9]. Experimentos lingüísticos en el contexto apropiado serán necesarios para observar cuándo y en qué grado, el "uso del lenguaje" acredita el empleo del modificador "más o menos", y cuál sería su representación apropiada como conjunto borroso. Para el ejemplo en desarrollo se especula que al aumentar hasta 1 la altura del conjunto borroso acotado, el incremento del núcleo del término lingüístico *constante* será no superior al que modificaría este término a "*más o menos constante*". En tal caso, un coeficiente de proporcionalidad k_1 representaría el *grado* de *más o menos* inducido sobre *constante* por la modificación geométrica del trapecio truncado. El mismo procedimiento debería ser aplicado para otros modificadores lingüísticos asociados a modificaciones geométricas de la representación de los conjuntos borrosos afectados.

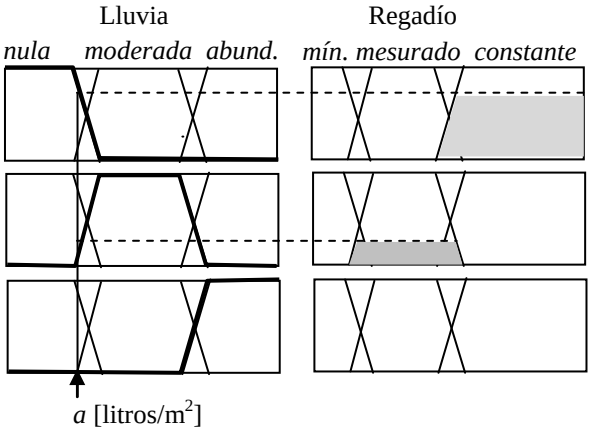


Figura 2: Representación de las reglas y el efecto de *a* litros/m^2 de lluvia caida, si las relgas son de tipo Mamdani

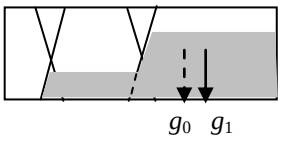


Figura 3: Conclusiones acotadas y agregadas punto a punto con la operación máximo.
Centro de gravedad de la agregación (g_0) y del término lingüístico acotado (gris) más próximo (g_1)

Hasta este punto, sobre la base de preservar la abscisa del centro de gravedad, ha sido posible asociar al trapecio acotado, la interpretación "*más o menos constante*" (con "intensidad" k_1).

Para encontrar una interpretación a la agregación completa de conclusiones acotadas, es necesario modificar "de manera interpretable" el trapecio *más o*

menos constante para que g_1 se desplace hasta coincidir con g_0. La manera posiblemente más simple de lograr esto, es desplazando el flanco izquierdo del trapecio hacia la izquierda, aumentando su área, lo cual conlleva también el desplazamiento de la abscisa del centro de gravedad hacia la izquierda. Pero ¿qué transformación lingüística tiene este efecto?

El modificador lingüístico *casi* produce el efecto de desplazar por completo hacia la izquierda (con "intensidad" k_2) el conjunto borroso representando el predicado que se va a modificar. En el caso de un trapecio, esto implica que *ambos* flancos serán igualmente desplazados. Sin embargo, tratándose de conjuntos borrosos trapezoidales la siguiente expresión produce el resultado deseado

$$k_2\ casi\ (\ k_1\ más\ o\ menos\ (constante)) \quad o\ (k_1\ más\ o\ menos\ (\ constante)) \qquad (1)$$

De manera alternativa, dado que en este ejemplo, el conjunto borroso de referencia está al extremo superior del universo de discurso, se puede utilizar el modificador "*no menos que"* (análogo al "$\geq$" de nuestra aritmética), obteniendo:

$$no\ menos\ que\ (k_2\ casi\ (k_1\ más\ o\ menos\ (constante))) \qquad (2)$$

Los paréntesis en (1) y (2) tienen por objeto mantener la distinción entre los modificadores lingüísticos utilizados.

El ejemplo 1 representa la situación posiblemente más favorable. Cabe hacer notar que a medida que la altura de truncación del término lingüístico usado como referencia disminuye, aumenta la extensión del núcleo con respecto del término lingüístico que le dió origen, cuando su altura se aumenta hasta alcanzar el valor máximo 1. Es simple entender que en el límite, el núcleo va a coincidir con el de su corte α cuando $\alpha = 0$, lo cual difícilmente se puede entender como una representación de *más o menos* (*término de origen*). En [4] se postula que 0,5 sería la altura mínima para que la interpretación metasemántica fuera aun aceptable. Por razones similares, es fácil de entender que las aproximaciones basadas en términos lingüísticos trapezoidales serán mejores que las obtenidas cuando los términos de origen tengan representación triangular (aun cuando los términos truncados de todos modos resulten trapezoidales).

3 EL CASO GENERAL

Es simple extender el método ilustrado en el ejemplo 1 a los casos en que la distribución de las alturas de truncado de las conclusiones en la agregación no es monótona; pero el término lingüístico de mayor altura relativa tiene la abscisa de su centro de gravedad próxima a la del centro de gravedad de la agregación.

Cuando el sistema modelado por un conjunto de reglas si-entonces de tipo Mamdani tiene dos o más entradas y las conclusiones están representadas por más de dos términos lingüísticos, no es difícil que la agregación de las conclusiones, para un caso dado, resulte en una distribución cóncava de las alturas de truncado. Y la situación más desfavorable (para la interpretación lingüística) será aquélla en la que el término lingüístico de menor altura, (altura < 0,5), - llámesele L - sea precisamente el que tiene la abscisa de su centro de gravedad más cercana a la de la agregación de las conclusiones. Una aproximación lingüística basada directamente en este término será necesariamente muy burda para ser útil.

Sin embargo, si se considera como situación de trabajo *el complemento* de la agregación de las conclusiones, nuevamente se tiene una situación como la analizada en la sección anterior. La solución del problema se obtendría como:

$$no\ (\ aproximación\ lingüística\ (\ complemento\ (\ L\ ))) \qquad (3)$$

Ejemplo 2:

Sea el conjunto de reglas siguientes un modelo adecuado para un problema dado, donde x_1, x_2 y x_3 representan variables físicas del entorno, z corresponde a la variable de conclusión, mientras que A_i, B_j, C_k y D_q con i,j,k,q $\in$ {1, 2, 3}, representan términos lingüísticos implementados como conjuntos borrosos.

R1: Si x_1 es A_1 y x_2 es B_2 entonces z es D_1

R2: Si x_1 es A_2 y x_3 es C_2 entonces z es D_2

R3: Si x_2 es B_2 y x_3 es C_3 entonces z es D_3

Su representación gráfica se ilustra en la Figura 4.

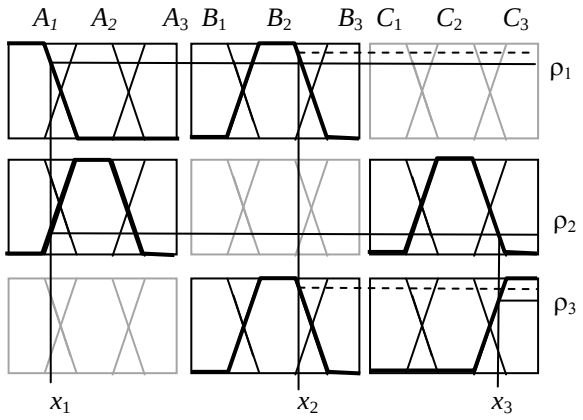


Figura 4: Representación de las premisas de las tres reglas cuando las variables de entrada toman los valores x_1, x_2 y x_3 ilustrados.

Dado que se trata de un sistema Mamdani, para los valores de x_1, x_2 y x_3 considerados, la conclusión D_1 será cortada al nivel ρ_1, la conclusión D_2 al nivel ρ_2, y la conclusión D_3 al nivel ρ_3. (En este ejemplo se usará $\rho_3 = \rho_1$.) La agregación utilizando el máximo dará la distribu-

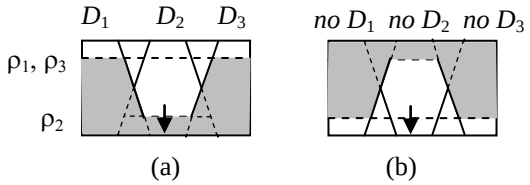


Figura 5: Agregación de conclusiones truncadas y su complemento

ción ilustrada en la Figura 5a. Su complemento es el ilustrado en la Figura 5b.

Resulta evidente que si la interpretación lingüística se basa en el término lingüístico D_2 truncado a la altura ρ_2, se obtendrá una aproximación muy burda, cuya interpretación lingüística distará de ser compresible. Si, en cambio, se procede tomando como base el complemento de la agregación de las conclusiones truncadas (Figura 5b), el objeto de una modificación geométrico-lingüística será el trapecio blanco asociado a *no* D_2. La base superior de este trapecio es ligeramente más ancha que el núcleo de *no* D_2; por lo tanto se obtiene una interpretación como

$$k_3\ más\ o\ menos\ (no\ D_2) \qquad (4)$$

donde el modificador lingüístico se aplicará con un grado k_3 pequeño. Ver una ampliación de la Figura 5(b) en la Figura 6, incluyendo la modificación antes descrita, destacando en línea de trazos largos, el trapecio correspondiente a la expresión (4).

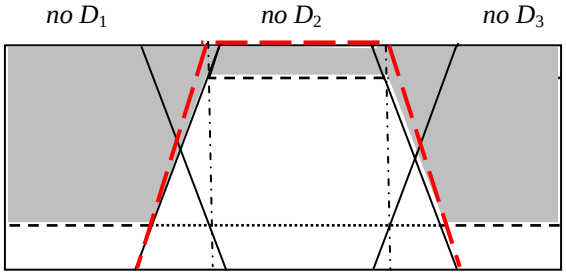


Figura 6: Análisis geométrico de la transformación conducente a la ecuación (4)

Para volver a la situación de inicio, como en la Figura 5(a), sólo hay que construir el complemento de la construcción ilustrada en la Figura 6, obteniéndose la representación mostrada en la Figura 7.

Formalmente, se trata del complemento de la expresión (4). Por lo tanto, la expresión metalingüística correspondiente, es

$$no\ (k_3\ más\ o\ menos\ (no\ D_2)\ ) \qquad (5)$$

Un análisis directo de la Figura 7, sin embargo, permite otra interpretación (la cual tendría que ser equivalente a la anterior). Obsérvese que los flancos modificados de *no* D_2 pueden verse (también) como modificaciones de los

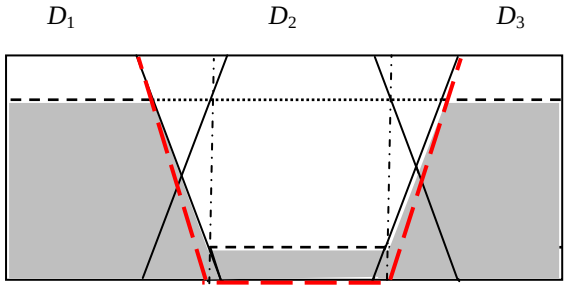


Figura 7: Efecto de la modificación (interpretable) del complemento de D_2 en el contexto de la agregación original de conclusiones truncadas.

flancos (interiores) de D_1 y D_3, respectivamente. El efecto de la modificación es comprimir el soporte respetando el núcleo de D_1 y D_2. Esta es una forma de representación del modificador lingüístico *muy* [4]. (Para el caso general, ver [1]). Cabe mencionar que, tal como en el lenguaje natural, con las representaciones elegidas, *muy* y *más o menos* no son recíprocamente inversos, pero, aparentemente "complementarios" (si sus argumentos también son complementarios).

La interpretación directa de la representación en la Figura 7 es, en general,

$$(\ k_4\ muy\ D_1\ )\ o\ (k_5\ muy\ D_3\ ) \qquad (6)$$

En primera instancia resulta sorprendente y hasta contraintuitivo, el hecho de que la interpretación (6) se ha obtenido sin tomar en cuenta el nivel ρ_1 de truncado de las conclusiones asociadas a D_1 y D_3. Sin embargo cabe pensar que en un sistema Mamdani pueden haber (puntualmente) agregaciones de conclusiones truncadas de diferente manera, pero que conduzcan a una misma abscisa del centro de gravedad. En tal caso las agregaciones tienen el mismo "significado nítido", que es el dado por tal abscisa.

Si, como en el ejemplo 2, que se está estudiando, D_2 tiene una representación simétrica, entonces será no solo $k_4 = k_5$, sino además $k_4 = k_3$. El uso del lenguaje hace plausible en tal caso la interpretación:

$$k_3\ muy\ (D_1\ o\ D_3) \qquad (7)$$

Nótese que si los términos lingüísticos de las conclusiones satisfacen la llamada partición de unidad, (como es el caso del ejemplo 2), entonces se cumple:

$$no\ D_2 = D_1\ o\ D_3 \qquad (8)$$

En estas condiciones, si se comparan las ecuaciones (5) y (7), se sostiene la conjetura anterior respecto al carácter complementario de *muy* y *más o menos* !

Para terminar el ejemplo, falta aun realizar el desplazamiento de la abscisa del centro de gravedad de D_2 hasta que coincida con la de la agregación global. Es

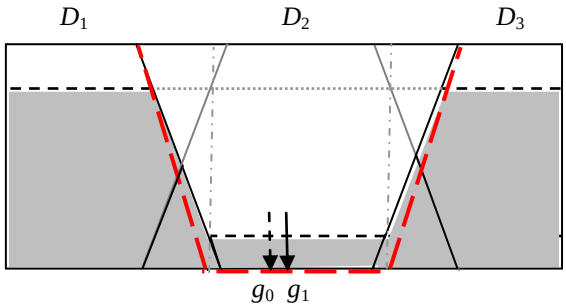


Figura 8: Posición relativa de los centros de gravedad del trapecio D_2 lingüísticamente modificado (g_1) y de la agregación de las conclusiones acotadas (g_0)

en este momento que entran a jugar un papel propio, tanto el nivel de truncado de D_1 y D_3 como sus áreas relativas. La situación se ilustra en la Figura 8. Es necesario un leve desplazamiento de la abscisa del centro de gravedad hacia la izquierda, vale decir, hay que incrementar adecuadamente el área de "k_3 *muy* D_1", a través de un modificador lingüístico para preservar la interpretación (metasemántica) en términos lingüísticos.. Dado que en este caso la corrección es "leve", puede recurrirse al modificador lingüístico "*alrededor de*", cuyo efecto sobre el conjunto borroso que representa su argumento, es preservar el núcleo y extender el soporte.

De acuerdo con esto, la interpretación lingüística final para el ejemplo 2 será:

$$(\ k_6\ \textit{alrededor de}\ (k_3\ \textit{muy}\ D_1)) \text{ o } (k_3\ \textit{muy}\ D_3) \qquad (9)$$

4 CONCLUSIONES

En este trabajo se ha considerado una forma de agregación cóncava de conclusiones, siendo más complicada que la analizada en un trabajo anterior [4]. Se ha presentado una metodología que hace posible también en este caso, una interpretación metasemántica.

Del análisis del ejemplo surge una novedosa "relación de complementariedad", expresable como

$$\neg\,(m(L)) = m'(\neg L) \qquad (10)$$

donde L representa un término lingüístico, m, un modificador lingüístico, y m' su complemento lingüístico *si es que existe*. El par (*más o menos* , *muy*) tiene una relación de tipo (m, m'), mientras que el modificador *alrededor de*, cuyo efecto es el de preservar el núcleo y aumentar el soporte del conjunto borroso que representa L, aparentemente no tiene "complementario" lingüístico, cuyo efecto sea el de preservar el soporte y disminuir el núcleo de su argumento. Cabe hacer notar que el modificador lingüístico *lejos de*, es antónimo de *alrededor de*, pero no su "complementario" lingüístico. Representa un término lingüístico nuevo; no una modificación del anterior [6].

Agradecimientos

El trabajo conducente a los resultados presentados en este artículo, fue financiado parcialmente por la Fundación para el Avance del Soft Computing, Mieres, Asturias.

Conversaciones con el Prof. Enric Trillas fueron decisivas para el desarrollo del trabajo.

El autor agradece sinceramente las correcciones y sugerencias de los evaluadores, las que contribuyeron efectivamente a mejorar el trabajo original.

Referencias

[1] P.J. Macvicar-Whelan: Fuzzy sets, the concept of height, and the hedge 'very'. *IEEE Trans. On Systems, Man, and Cybernetics*, SMC-8, pp. 507-511, 1978

[2] E.H Mamdani., S. Assilian: An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* **7**: pp. 1-13, 1975

[3] C. Moraga: A Metasemantics to refine fuzzy if-then rules. *Proc. 34th IEEE Internat. Symp. on Multiple-valued Logic*, IEEE-CS-Press, pp. 148-153, 2004

[4] C. Moraga: An essay on the interpretability of Mamdani Systems. En: E. Trillas, P. Bonissone, L. Magdalena, J. Kacprzyk (Eds.) *Theory and Experimentation. Homage to Ebrahim Mamdani.* Springer (in Press), 2011/12

[5] E. Trillas: *Fundamentals of Approximate Reasoning*, Manuscript, ECSC, 2010

[6] E. Trillas, C. Moraga, S. Guadarrama, S. Cubillo, E. Castiñeira: Computing with Antonyms. *Studies in Fuzziness and Soft Computing* **217**, pp. 133-153, Springer, 2007

[7] E. Trillas, C. Moraga, A. Sobrino: On 'Family Resemblance' with fuzzy sets. *Proc. Joint Conference IFSA/Eusflat*, pp. 306-311, 2009.

[8] L. Wittgenstein. *Philosophical Investigations*. Basil Blackwell, London, 1973.

[9] L.A. Zadeh: A fuzzy-set-theoretic interpretation of hedges. Memorandum M-335, Electronics Research Laboratory, University of California, Berkeley, 1972.

Integración de la precisión en la mejora de la interpretabilidad de modelos difusos mediante colonias de hormigas

Pablo Carmona[1], **Juan Luis Castro**[2], **José Luis Herrero**[1]

[1]*Depto. de Ingeniería de Sistemas Informáticos y Telemáticos, UEx,*
[2]*Depto. de Ciencias de la Computación e Inteligencia Artificial, UGr,*

Resumen

La ampliación de la sintaxis de las reglas difusas con nuevos predicados y el uso de algoritmos de hormigas se ha revelado efectivo en la mejora de la interpretabilidad de modelos difusos basada en la reformulación de un modelo inicial. Sin embargo, aunque esta reformulación respete las reglas iniciales, la búsqueda de una mayor interpretabilidad puede dar lugar a nuevas reglas en espacios de entrada no cubiertos inicialmente. Así, pueden surgir variaciones en la precisión del modelo mejorado respecto al original debido a la interacción entre reglas difusas iniciales y nuevas reglas adyacentes. Se hace deseable, por tanto, integrar la precisión como criterio adicional para evaluar la optimalidad del modelo reformulado y así establecer un balance entre mejora de la interpretabilidad y mantenimiento de la precisión inicial. Se presenta aquí tal mecanismo, que además permite regular el citado balance e, incluso, condicionar la mejora de la interpretabilidad al mantenimiento riguroso de la precisión inicial.

Palabras Clave: Modelado difuso, algoritmos de hormigas, interpretabilidad vs. precisión.

1 Introducción

La interpretabilidad es una característica esencial del modelado difuso. Por ello, dado que la búsqueda de un ajuste mayor del modelo suele conllevar una pérdida de interpretabilidad, a menudo es deseable alcanzar un compromiso entre ambos aspectos para que el modelo resultante sea a su vez útil como herramienta descriptiva y eficaz como modelo de cómputo. Si bien durante años pasados hubo tendencia a primar los aspectos de precisión sobre los de interpretabilidad, en la actualidad se ha devuelto la importancia al papel que la interpretabilidad ha jugado en la concepción y desarrollo del modelado difuso [2, 9, 11].

Una de las formas de mejorar la interpretabilidad de un modelo difuso consiste en tratar de identificar reglas tan generales como sea posible, de forma que cada regla cubra el mayor número de ejemplos de entrenamiento y, de este modo, el tamaño de la base de reglas disminuya [1, 6]. Además, la extensión de la sintaxis de las reglas con nuevos predicados diferentes del habitual *es-igual-a* permite generar reglas más compactas.

En trabajos previos [3, 4], los autores propusieron un método para la búsqueda de un conjunto de reglas generales que reformularan un modelo difuso inicial. Para ello, se empleó un algoritmo de hormigas, método de optimización global que se basa en el comportamiento emergente que surge de la búsqueda cooperativa realizada por un conjunto de agentes denominados hormigas artificiales [8]. Sin embargo, aunque esta reformulación respeta las reglas iniciales, la búsqueda de una mayor interpretabilidad puede dar lugar a nuevas reglas en espacios de entrada no cubiertos inicialmente, pudiendo surgir variaciones en la precisión respecto al modelo inicial debido a la interacción entre las reglas difusas iniciales y las nuevas reglas adyacentes. Se propone aquí la integración de la precisión como criterio adicional para evaluar la optimalidad del modelo reformulado y así establecer un balance entre mejora de la interpretabilidad y mantenimiento de la precisión inicial. Es más, mediante la parametrización del citado balance, la mejora de interpretabilidad podrá condicionarse al mantenimiento estricto de la precisión del modelo inicial, sin permitir pérdida alguna.

La sección siguiente introduce la sintaxis de las reglas compuestas que permitirán mejorar la interpretabilidad de los modelos difusos. La Sección 3 resume el algoritmo de hormigas original. La Sección 4 constituye la parte central del trabajo, en la que se presenta una medida heurística para evaluar el error, así como el mecanismo de integración de la precisión en el algoritmo original. En la Sección 5 se analizan experimentalmente los beneficios de la propuesta y la Sección 6 resume las conclusiones alcanzadas.

2 Reglas simples y compuestas

Se considerarán sistemas con múltiples entradas y una salida (Multiple-Input–Single-Output, MISO) con n variables de entrada $\mathbf{X} = \{X_1, \ldots, X_n\}$ definidas sobre el universo del discurso $\mathscr{X} = \mathscr{X}_1 \times \ldots \times \mathscr{X}_n$, y una variable de salida Y definida sobre el universo del discurso $\mathscr{Y}$. El dominio difuso de X_i se denota $\widetilde{\mathscr{X}_i} = \{LX_{i,1}, \ldots, LX_{i,p_i}\}$, donde $LX_{i,j}$ representa la etiqueta lingüística del j-ésimo valor difuso. $\widetilde{\mathscr{Y}} = \{LY_1, \ldots, LY_q\}$ es el dominio difuso de salida.

Habitualmente, las reglas de un modelo difuso MISO contienen en el antecedente una premisa para cada variable de entrada que la asocia a una etiqueta. A estas reglas, que denominaremos *reglas simples*, las denotamos como:

$$R^{i_1 \ldots i_n}_{LY^i} : LX_{1,i_1}, \ldots, LX_{n,i_n} \rightarrow LY^i. \quad (1)$$

Para mejorar la compactación de las reglas, se puede asociar más de una etiqueta a cada variable de entrada en el antecedente, así como utilizar otros predicados además del habitual *es-igual-a*. A estas reglas, que denominaremos *reglas compuestas*, las denotamos como:

$$R^i \ : \ [\neg]X_1 \ op_i \ SX^i_1, \ \ldots, \ [\neg]X_n \ op_n \ SX^i_n \rightarrow LY^i$$

donde $op_i \in \{=, \leq, \geq, \div, \in\}$ ($\div$ significa *entre*), $\neg$ representa al operador de negación y donde cada SX^i_j es una etiqueta si $op_i \in \{=, \leq, \geq\}$, dos si op_i es $\div$, y dos o más etiquetas si op_i es $\in$.

3 Algoritmo de hormigas para la mejora de la interpretabilidad de modelos difusos

Se resume aquí el algoritmo descrito en [3] con las mejoras introducidas en [4] (para más detalles, remitimos al lector a tales trabajos). El algoritmo parte de un conjunto de reglas iniciales (CRI) y busca su mejor transformación en un conjunto de reglas compuestas (CRC).

Para ello, una hormiga puede realizar dos tipos de pasos: de *inclusión*, que añaden una regla del CRI al CRC, y de *amplificación*, que extienden la cobertura de una regla del CRC añadiendo una etiqueta a una de sus premisas en el antecedente. Un paso de amplificación se denota $< i, j, k >$, que representa la adición de la k-ésima etiqueta, $LX_{j,k}$, a la j-ésimo premisa en el antecedente de la regla compuesta R^i. Un paso de inclusión se denota $< i, 0, 0 >$, que representa la inclusión de la regla $R^{i_1 \ldots i_n}_{LY^i}$ del CRI al CRC, dando lugar a la regla compuesta R^i.

Para construir una solución, en primer lugar la hormiga incluirá aleatoriamente una regla inicial en su CRC (cada hormiga construye un CRC propio). A continuación, la hormiga seleccionará un paso entre los válidos para cada estado de la hormiga, incluyendo en el CRC nuevas reglas del CRI o amplificando las reglas ya incluidas.

Tabla 1: Construcción de una solución

X_2 \ X_1	**S**	**M**	**L**
S	M^1	L^6	L^4
M		L^3	
L	M^2		S^5

(a) CRI

X_2 \ X_1	**S**	**M**	**L**
S	M^1	L^4	L^4
M	M^1	L^4	L^4
L	M^1	S^5	S^5

(b) CRC

Paso	1	2	3	4	5	6	7	8
Regla	4	4	5	4	1	1	5	1
Variable	0	1	0	2	0	2	1	2
Etiqueta	0	2	0	2	0	2	2	3

(c) Secuencia de pasos

Una inclusión será válida si la regla a incluir aún no ha sido cubierta por el CRC y una amplificación si la regla a amplificar no genera inconsistencias con reglas del CRI u otras reglas del CRC ya existentes (es decir, no cubra reglas del CRI o del CRC con consecuentes distintos al suyo). Con esta definición de paso válido, todo paso que se evalúe como no válido en un estado de la hormiga es añadido a una *lista tabú* y descartado directamente en futuras transiciones de dicha hormiga, pues si no es válido en un estado, no podrá serlo en estados posteriores.

A modo de ejemplo de la construcción de una solución, supóngase un sistema con dos entradas y una salida con el mismo dominio difuso $\{S, M, L\}$ para cada variable. La Tabla 1(a) muestra el CRI, donde el superíndice de cada consecuente permite relacionar cada regla compuesta con la regla inicial de la que surgió. La Tabla 1(b) muestra el modelo final descrito por el CRC

$$\begin{aligned} R^1 &: X_1 = S \rightarrow M \\ R^4 &: X_1 \geq M, \ X_2 \leq M \rightarrow L \\ R^5 &: X_1 \geq M, \ X_2 = L \rightarrow S. \end{aligned}$$

La Tabla 1(c) muestra los pasos $< i, j, k >$ que llevaron al CRC final.

El cálculo de la bondad de cada solución se basa en la simplicidad sintáctica de la base de reglas resultante, según

$$I(BR) = \frac{NRI - \sum_i C(R^i)}{NRI}, \quad (2)$$

donde NRI es el número de reglas iniciales y $\sum_i C(R^i)$ representa la complejidad sintáctica del CRC final como el número total de elementos sintácticos (variables, predicados y etiquetas) de sus reglas.

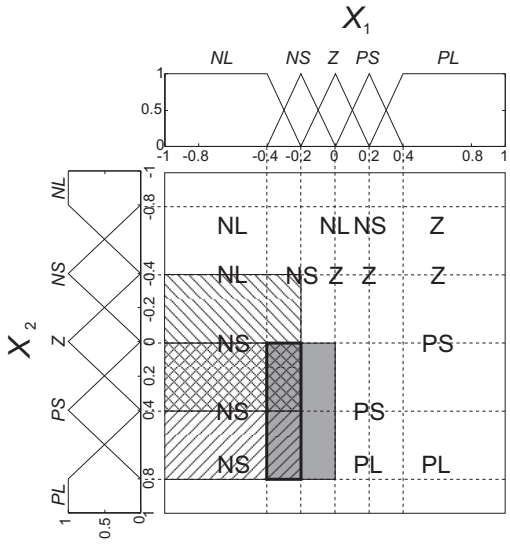


Figura 1: Rejilla difusa del espacio de entrada de un modelo inicial con entradas X_1 y X_2

4 Integración de la precisión en el algoritmo de hormigas

Integrar la precisión en el algoritmo de hormigas requiere evaluar el error introducido por la generalización de las reglas iniciales. En una primera aproximación, se empleó una muestra de puntos del espacio de entrada correspondiente a las zonas de solapamiento entre las reglas del modelo inicial y las añadidas por el modelo reformulado, y se calculó la diferencia entre las salidas generadas por sendos modelos. Sin embargo, el coste computacional de esta integración, que requiere evaluar el error introducido por cada solución construida por cada hormiga, con la consiguiente necesidad de realizar un número elevado de inferencias, hacía esta técnica inviable para dimensionalidades altas.

Se presenta aquí como alternativa una medida de error heurística que permite asociar a priori un error a cada regla susceptible de ser añadida a una región de entrada en función de la disposición de las reglas iniciales, lo que disminuye drásticamente el coste computacional. En concreto, a continuación se describe la medida heurística empleada para evaluar el error introducido cuando se añade una regla simple en una región de entrada que, en el modelo inicial, carece de regla. El lugar y modo de integración de esta medida en el algoritmo de hormigas se describe en la Sección 4.2.

4.1 Medida heurística del error

Para ilustrar la idea que subyace a la medida heurística que se presenta, obsérvese la Figura 1, que representa un modelo difuso inicial con 2 variables de entrada y 5 etiquetas asociadas a cada dominio difuso $\{NL, NS, Z, PS, PL\}$. Ciertas regiones de entrada se encuentran ocupadas por reglas iniciales (por ejemplo, la región $[NL, PS]$ está ocupada por la regla $R^{1,4}_{NS} : NL, PS \rightarrow NS$) y otras no (por ejemplo, la región $[NS, PS]$). Supóngase que durante la construcción de la solución por una hormiga, una regla ocupa la región de entrada $[NS, PS]$. La heurística aplicada para evaluar el error introducido por esta regla sobre el modelo difuso inicial se basa en que la inferencia de cada regla inicial que solapa con $[NS, PS]$ se verá afectada en mayor o menor grado en función, por un lado, de la diferencia entre los consecuentes de la nueva regla y de la regla inicial, y por otro, de la extensión de la zona de solapamiento. Así, por un lado, una regla en $[NS, PS]$ con consecuente PL introduciría más error sobre el modelo inicial que otra con consecuente Z, pues PL se encuentra a más distancia que Z del consecuente NS, el único que, según las reglas iniciales adyacentes, afecta inicialmente a la salida en $[NS, PS]$. Igualmente, una regla con consecuente NS no introduciría ningún error. Por otro lado, cabe esperar que el error que una nueva regla en $[NS, PS]$ introduzca sobre la regla $R^{1,4}_{NS} : NL, PS \rightarrow NS$ sea mayor que el que introduzca sobre la regla $R^{1,3}_{NS} : NL, Z \rightarrow NS$, pues la zona de solapamiento en el primer caso ($[-0.4, -0.2] \times [0, 0.8]$) es mayor que en el segundo ($[-0.4, -0.2] \times [0, 0.4]$).

Con este objeto, en primer lugar se define el grado de solapamiento entre dos etiquetas difusas como

$$O_L(LX_{i,j}, LX_{i,k}) = max(min(r_j, r_k) - max(l_j, l_k), 0), \quad (3)$$

donde $[r_j, l_j]$ y $[r_k, l_k]$ son los soportes de $LX_{i,j}$ y $LX_{i,k}$, respectivamente, y se define el grado de solapamiento entre dos regiones de entrada como

$$O_R(R^{i_1 \cdots i_n}, R^{j_1 \cdots j_n}) = \prod_{k=1}^{n} O_L(LX_{k,i_k}, LX_{k,j_k}). \quad (4)$$

A continuación, se define el error heurístico de una regla simple con respecto a la base de reglas inicial como

$$EH(R^{i_1 \cdots i_n}_{LY^i}) = \frac{\sum_{j_1,\ldots,j_n} O_R(R^{i_1 \cdots i_n}, R^{j_1 \cdots j_n}) \times \Delta(LY^i, LY^j)}{\sum_{j_1,\ldots,j_n} O_R(R^{i_1 \cdots i_n}, R^{j_1 \cdots j_n})} \quad (5)$$

siendo $R^{j_1 \cdots j_n}_{LY^j}$ las reglas iniciales que solapan con $R^{i_1 \cdots i_n}_{LY^i}$ y $\Delta(LY^i, LY^j)$ la distancia entre sus consecuentes. En concreto, como medida de distancia entre conjuntos difusos se empleó el índice de desajuste [10].

Obsérvese que, dado que el error heurístico en (5) solo depende de las reglas iniciales y del consecuente de la regla simple evaluada, puede calcularse a priori el error asociado a cada regla susceptible de ser añadida al modelo inicial, sin necesidad de que esta evaluación deba realizarse durante la ejecución del algoritmo de hormigas, con el consiguiente ahorro computacional. Además, en este cálculo previo solo

se obtendrá el error de las reglas situadas en regiones *relevantes*, es decir, en espacios de entrada no cubiertos por el modelo inicial que solapen con reglas iniciales.

4.2 Integración de la medida heurística del error en el algoritmo de hormigas

En un algoritmo de hormigas, la función heurística mide el nivel de ideoneidad local de cada paso válido en un estado de la hormiga. En el algoritmo original, esta función heurística consideraba la amplitud de la zona de amplificación asociada al paso (es decir, el nuevo espacio de entrada cubierto por el paso) para establecer dicha idoneidad, basándose en criterios guiados exclusivamente por la interpretabilidad de la regla amplificada. Una forma de integrar la precisión consiste en evaluar el grado de error introducido por dicha zona de amplificación en base a la definición de error heurístico de la Sección 4.1, según la ecuación

$$EH(ZA) = \frac{\sum_{(i_1,\ldots,i_n)\in ZA} EH(R_{LY^i}^{i_1\ldots i_n})}{Reg_{max}}, \qquad (6)$$

siendo $(i_1,\ldots,i_n)$ las regiones relevantes dentro de la zona de amplificación ZA, LY^i el consecuente de la regla amplificada y Reg_{max} el número máximo de regiones de entrada que pueden cubrirse por una amplificación de la variable considerada en el paso.

Ahora bien, sería deseable integrar algún mecanismo que permitiera regular el balance interpretabilidad/precisión considerado por el algoritmo, pudiendo variar desde esquemas que solo contemplaran criterios de interpretabilidad hasta aquéllos donde la mejora de la interpretabilidad quedara supeditada al mantenimiento riguroso de la precisión del modelo inicial. Por ello, se definió una medida de error heurístico acotado

$$\widetilde{EH}(ZA) = \begin{cases} 1 & \text{si } EH(ZA) > \tau_{EH} \\ EH(ZA) & \text{en otro caso} \end{cases} \qquad (7)$$

donde $\tau_{EH} \in [0,1]$ representa el umbral de máximo error admitido en una zona de amplificación, y se redefinió la función heurística del algoritmo de hormigas como

$$\eta = \frac{Reg_{cov}}{Reg_{max}} \cdot (1 - \widetilde{EH}(ZA))^{\sigma_1}, \qquad (8)$$

añadiendo, por lo tanto, un segundo factor que mediante el parámetro $\sigma_1 \in \{0,1\}$ establece la consideración o no de la precisión en el valor heurístico de cada paso.

Obsérvese que un umbral $\tau_{EH} = 1$ equivale a no condicionar la validez de un paso al grado de error que introduzca en el modelo (si bien un grado de error alto disminuirá su valor heurístico), mientras que un umbral $\tau_{EH} = 0$ equivale a considerar como pasos válidos solo a aquellos que no introducen error alguno en el modelo, asegurando por tanto el mantenimiento riguroso de la precisión del modelo inicial.

Como se ha mencionado anteriormente, la función heurística original se basaba en el grado de amplitud de la zona de amplificación del paso. Dado que un paso de amplificación consiste en la adición de una etiqueta a una premisa, la zona de amplificación asociada al paso dependerá de qué otras amplificaciones se hayan realizado antes sobre esa regla, es decir, de la posición que dicho paso ocupe en el recorrido de la hormiga. Esta observación puede llevar a pensar en la necesidad de evaluar la función heurística para todos los pasos válidos en cada estado de la hormiga, aunque su valor ya hubiera sido calculado para algunos de esos pasos en estados anteriores. Sin embargo, un análisis más profundo permitió concluir en [3] que el recálculo sólo era necesario para los pasos de la regla amplificada en el estado inmediatamente anterior que no afectaran a la premisa modificada en dicho paso, pues el resto de pasos mantendrán intacta su zona de amplificación asociada. Es decir, si el paso inmediatamente anterior de la hormiga fue $< i, j, \cdot >$ (que amplificó SX_j en la regla R^i), para la selección del paso actual solo será necesario recalcular el valor heurístico de los pasos $< i, k, \cdot >$, para $k \neq j$. Esto redujo drásticamente el número de evaluaciones de la función heurística necesarias para seleccionar cada paso, con el consiguiente ahorro computacional.

Con la modificación que aquí se propone, es necesario reconsiderar si sigue siendo válido este ahorro computacional, pues ahora el valor heurístico no solo depende de la amplitud de la zona de amplificación asociada al paso, sino también de su error. Ahora bien, una determinada zona de amplificación de una regla compuesta siempre tendrá asociado el mismo error, ya que ese error depende exclusivamente del consecuente que se sitúe en las regiones sin reglas iniciales (que siempre será el mismo para esa regla compuesta) y de las reglas del CRI con las que solape (que siempre serán las mismas). Por lo tanto, el error sólo variará cuando varíe la zona de amplificación asociada al paso, es decir, en los mismos casos contemplados por el ahorro computacional original. En consecuencia, dicho ahorro computacional puede mantenerse.

Cabe también observar que en el algoritmo de hormigas propuesto no hay pasos atrás o, dicho de otro modo, el mismo paso de amplificación realizado en momentos posteriores a uno dado solo podrá aumentar su zona de amplificación asociada. Dado que el error de cada región simple es fijo desde el principio para un mismo consecuente y, según la Ec. (6), $EH(ZA') \geq EH(ZA)$ para un ZA' más amplio que ZA, esto supone que un mismo paso de amplificación dado en momentos posteriores nunca podrá reducir su error asociado. Por este motivo, cada paso que proporcione un $EH(ZA) > \tau_{EH}$, que resultará en $\widetilde{EH}(ZA) = 1$ y un $\eta = 0$ (o lo que es lo mismo, en un paso no válido) se incluirá en la lista tabú y será así descartado en transiciones posteriores de la hormiga.

5 Resultados experimentales

Se realizaron experimentos para evaluar la propuesta presentada en la Sección 4. Para ello, se compararon los resultados del algoritmo original que sólo considera criterios de interpretabilidad ($\sigma_1 = 0$) con los resultados obtenidos al integrar la medida de error en la función heurística ($\sigma_1 = 1$). Además, se analizó la eficacia del umbral τ_{EH} para controlar el balance interpretabilidad/precisión utilizando para dicho umbral los valores {0, 0.25, 0.5, 1} (valores superiores a 0.5 mostraban resultados equivalentes al umbral máximo). Cada configuración del algoritmo se ejecutó 5 veces y se obtuvieron los valores medios.

Para la implementación del algoritmo de hormigas, se utilizó la versión ACS propuesta en [7] con los parámetros $\alpha = \rho = 0.1$ y $q_0 = 0.7$. Además, se emplearon 5 hormigas, el número máximo de ciclos se fijó en 40 y se consideró como condición de convergencia para finalizar la ejecución del algoritmo que todas las hormigas construyeran el mismo CRC durante 5 ciclos consecutivos.

La Tabla 2 resume los resultados obtenidos tras la experimentación, en la que se aplicó el algoritmo de hormigas a modelos difusos con distinta dimensionalidad n y número de etiquetas por variable p_i, empleando para cada configuración $[n, p_i]$ tres bases de reglas generadas automáticamente (BR^1, BR^2 y BR^3). Para ello, se implementó un generador automático de bases de reglas compuestas donde tanto el conjunto de etiquetas asociado a cada premisa SX_j^i como el consecuente LY^i se generaban aleatoriamente con la restricción de que las bases de reglas resultantes no presentaran inconsistencias. Así, el algoritmo de hormigas recibiría como entrada un conjunto de reglas simples para las que existe una descripción compacta mediante reglas compuestas y puede analizarse si el algoritmo de hormigas consigue, al menos, encontrar dicha descripción.

La tabla muestra, para cada configuración $[n, p_i]$, en la primera fila, los datos relativos a la base de reglas inicial sobre la que se aplica el algoritmo y, en el resto de filas, los resultados promediados relativos a las distintas configuraciones de los parámetros σ_1 y τ_{EH} en dicho algoritmo.

En concreto, para cada modelo difuso al que se aplica el algoritmo, se muestra la siguiente información: complejidad acumulada del modelo difuso descrito mediante reglas compuestas ($\sum_i C(R^i)$ en (2)); entre llaves, número de reglas iniciales del CRI proporcionado como entrada del algoritmo de hormigas; entre corchetes, tiempo medio en segundos empleado en cada ejecución del algoritmo; entre paréntesis angulares, número de reglas compuestas que describen al modelo difuso; entre paréntesis, error medio del modelo reformulado con respecto al modelo inicial.

Para el cálculo del error, se consideró:

$$E = \frac{\sum_{i=1}^{t} |y_N(\mathbf{x}^i) - \hat{y}_N(\mathbf{x}^i)|}{t} \qquad (9)$$

donde el conjunto de test utilizado $\{\mathbf{x}^1, \cdots, \mathbf{x}^t\}$ consistió en puntos aleatoriamente distribuidos en el espacio de entrada cubierto por el CRI y donde $y_N(\mathbf{x}^i)$ e $\hat{y}_N(\mathbf{x}^i)$ representan las salidas normalizadas generadas por el modelo inicial y el reformulado, respectivamente.

En base a los resultados mostrados, puede concluirse que la integración de la precisión consigue los objetivos buscados:

- Por un lado, la consideración de la precisión en (8) con $\sigma_1 = 1$ permite establecer un compromiso entre mejora de la interpretabilidad y mantenimiento de la precisión del modelo inicial en función del umbral τ_{EH} utilizado. Así, con $\tau_{EH} = 0$ se consigue mantener estrictamente la precisión del modelo inicial, si bien la mejora de la intepretabilidad no es tan alta como en el resto de casos[1], y a medida que se relaja el umbral τ_{EH}, la interpretabilidad conseguida aumenta junto con la diferencia entre la precisión del modelo inicial y del reformulado, y en general, consiguiendo una reducción en el error asociado a $\sigma_1 = 0$.
- Por otro lado, si bien en general el coste computacional se incrementa ligeramente cuando se considera la precisión, este incremento no es significativo ni se encuentra afectado por la dimensionalidad del modelo inicial, e incluso se reduce el tiempo necesario para encontrar una solución cuando $\tau_{EH} = 0$, al disminuir considerablemente el número de pasos válidos contemplados en la construcción de las soluciones.

6 Conclusiones y Trabajos Futuros

En este trabajo se propone la consideración de la precisión como criterio adicional sobre el algoritmo de hormigas presentado por los autores en [3, 4] para una reformulación más interpretable de modelos difusos ya existentes. El objetivo de esta propuesta es permitir el establecimiento de un balance entre mejora de la interpretabilidad y mantenimiento de la precisión del modelo inicial.

Con este propósito, se definió una medida heurística del error entre el modelo difuso reformulado y el inicial que permitiera integrar la precisión de forma eficiente en la función heurística que mide localmente el grado de adecuación de cada paso dado por una hormiga. Además, dicha integración contempló el uso de un umbral de error máximo admitido, pudiendo de esta forma condicionar la mejora de la interpretabilidad al grado de mantenimiento de la precisión deseado (incluso, al mantenimiento estricto de la misma).

En base a la experimentación realizada, puede concluirse que la propuesta cumple sus objetivos: permite establecer un balance entre mejora de intepretabilidad y modificación

[1] No obstante, debe advertirse que el algoritmo no parte de las reglas compuestas proporcionadas por el generador de bases de reglas, sino de las reglas simples que las integran.

Tabla 2: Resultados obtenidos con las distintas versiones del método propuesto: complejidad acumulada, {número de reglas simples}, ⟨numero de reglas compuestas⟩, [tiempo en segundos], (error medio)

n/p_i	σ_1/τ_{EH}	RB_a^1	RB_a^2	RB_a^3
2/9	BR inicial	3.64 {49} ⟨5.0⟩	2.71 {47} ⟨4.0⟩	3.00 {40} ⟨4.0⟩
	0.00/1.00	2.57 [94s] ⟨4.0⟩ (0.0743)	1.93 [101s] ⟨4.0⟩ (0.0359)	2.43 [94s] ⟨4.0⟩ (0.0267)
	1.00/0.00	3.64 [92s] ⟨5.0⟩ (0.0000)	2.57 [88s] ⟨4.0⟩ (0.0000)	3.00 [68s] ⟨4.0⟩ (0.0000)
	1.00/0.25	3.36 [106s] ⟨5.0⟩ (0.0280)	1.96 [100s] ⟨4.0⟩ (0.0223)	2.79 [88s] ⟨4.0⟩ (0.0041)
	1.00/0.50	2.71 [106s] ⟨4.0⟩ (0.0622)	1.93 [106s] ⟨4.0⟩ (0.0357)	2.71 [92s] ⟨4.0⟩ (0.0242)
	1.00/1.00	2.71 [97s] ⟨4.0⟩ (0.0620)	1.93 [105s] ⟨4.0⟩ (0.0359)	2.43 [97s] ⟨4.0⟩ (0.0265)
3/7	BR inicial	3.56 {88} ⟨5.0⟩	5.72 {145} ⟨8.0⟩	4.17 {98} ⟨6.0⟩
	0.00/1.00	2.00 [73s] ⟨5.0⟩ (0.0572)	3.70 [86s] ⟨7.2⟩ (0.0614)	1.96 [69s] ⟨5.0⟩ (0.0523)
	1.00/0.00	3.50 [51s] ⟨5.0⟩ (0.0000)	5.72 [68s] ⟨8.0⟩ (0.0000)	4.11 [62s] ⟨6.0⟩ (0.0000)
	1.00/0.25	2.52 [83s] ⟨5.0⟩ (0.0370)	4.36 [93s] ⟨7.0⟩ (0.0178)	2.63 [92s] ⟨5.2⟩ (0.0262)
	1.00/0.50	2.06 [91s] ⟨5.0⟩ (0.0462)	4.01 [110s] ⟨7.0⟩ (0.0300)	2.17 [86s] ⟨5.0⟩ (0.0334)
	1.00/1.00	2.02 [94s] ⟨5.0⟩ (0.0527)	3.64 [112s] ⟨7.0⟩ (0.0631)	2.03 [83s] ⟨5.0⟩ (0.0460)
4/5	BR inicial	14.55 {250} ⟨20.0⟩	7.70 {177} ⟨10.0⟩	3.35 {112} ⟨5.0⟩
	0.00/1.00	10.52 [201s] ⟨18.8⟩ (0.0677)	3.40 [127s] ⟨9.0⟩ (0.0982)	1.50 [70s] ⟨5.0⟩ (0.0852)
	1.00/0.00	14.55 [135s] ⟨20.2⟩ (0.0000)	7.70 [48s] ⟨10.0⟩ (0.0000)	3.10 [60s] ⟨5.0⟩ (0.0000)
	1.00/0.25	12.65 [212s] ⟨19.2⟩ (0.0191)	6.38 [147s] ⟨10.0⟩ (0.0169)	2.45 [88s] ⟨5.0⟩ (0.0084)
	1.00/0.50	11.11 [263s] ⟨19.0⟩ (0.0417)	3.62 [174s] ⟨9.0⟩ (0.0980)	1.66 [100s] ⟨5.0⟩ (0.0710)
	1.00/1.00	10.68 [256s] ⟨19.0⟩ (0.0614)	3.54 [166s] ⟨9.0⟩ (0.0942)	1.53 [95s] ⟨5.0⟩ (0.0826)

de la precisión inicial y lo hace utilizando una medida de error eficiente que posibilita su uso en el entorno de los algoritmos de optimización considerados.

Como trabajo futuro, se pretende implementar esta mejora en otro modelo propuesto por los autores [5] que emplea excepciones en las reglas difusas. Su implementación previa en el modelo sin excepciones utilizado aquí se debe a que su mayor simplicidad lo hace más apropiado para analizar los resultados en un primer acercamiento al problema.

Agradecimientos

Este trabajo ha sido cofinanciado por la Comunidad Autónoma de Extremadura y el FEDER.

Referencias

[1] J.M. Alonso, S. Guillaume y L. Magdalena. A hierarchical fuzzy system for assessing interpretability of linguistic knowledge bases in classification problems. In *Proc. IPMU'06*, 348–355, Paris, Francia, 2006.

[2] J.M. Alonso, L. Magdalena y G.G. Rodríguez. Looking for a good fuzzy system interpretability index: an experimental approach. *Int. J. Approx. Reasoning*, 51:115–134, 2009.

[3] P. Carmona y J.L. Castro. Using ant colony optimization for learning maximal structure fuzzy rules. En *Proc. FUZZ-IEEE'05*, 702–707, Reno, Nevada, 2005.

[4] P. Carmona y J.L. Castro. Mecanismos de mejora para la identificación de modelos difusos interpretables mediante algoritmos de hormigas. En *Actas del II Simposio sobre Lógica Fuzzy y Soft Computing*, 30–44, Zaragoza, España, 2007.

[5] P. Carmona y J.L. Castro. Algoritmo de hormigas mejorado para la descripción de modelos difusos interpretables mediante reglas con excepciones. En *Actas ESTYLF'08*, 185–190, Langreo-Mieres, España, 2008.

[6] J.L. Castro, J.J. Castro-Schez y J.M. Zurita. Learning maximal structure rules in fuzzy logic for knowledge acquisition in expert systems. *Fuzzy Sets Syst.*, 101:331–342, 1999.

[7] M. Dorigo y L.M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evol. Comput.*, 1(1):53–66, 1997.

[8] M. Dorigo y T. Stützle. *Ant Colony Optimization*. MIT Press, 2004.

[9] M.J. Gacto, R. Alcalá y F. Herrera. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Inf. Sci.*, 181(20):4340–4360, 2011.

[10] A. Kaufmann y M.M. Gupta. *Introduction to Fuzzy Arithmetic, Theory and Applications*. Van Nostrand Reinhold, New York, NY, USA, 1985.

[11] C. Mencar y A. Fanelli. Interpretability constraints for fuzzy information granulation. *Inf. Sci.*, 178(24):4585–4618, 2008.

SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS DIFUSAS Y SISTEMAS NÍTIDOS ROBUSTOS ENTRENADOS EN PRESENCIA DE RUIDO DE CLASE: UN CASO DE ESTUDIO

José A. Sáez[1], **Julián Luengo**[2], **Francisco Herrera**[1]

[1] *Dpto. de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, CITIC-UGR, 18071, Granada, España,* {smja, herrera}@decsai.ugr.es

[2] *Dpto. de Ingeniería Civil, Universidad de Burgos, 09006, Burgos, España,* jluengo@ubu.es

Resumen

La presencia de ruido es común en cualquier conjunto de datos y puede afectar a la precisión, tiempo de construcción y complejidad de los clasificadores. Algunos algoritmos, llamados robustos, incorporan mecanismos para reducir estos efectos negativos del ruido. Por otra parte, a los Sistemas de Clasificación Basados en Reglas Difusas se les ha atribuido tradicionalmente la capacidad de tratar con datos imperfectos.

Esta contribución analiza el comportamiento de los Sistemas de Clasificación Basados en Reglas Difusas frente a los sistemas nítidos robustos cuando el ruido afecta a la clase de los ejemplos de entrenamiento. Para ello, comparamos el algoritmo difuso *Fuzzy Unordered Rule Induction Algorithm* respecto al algoritmo nítido robusto C4.5. A partir de los resultados obtenidos, se observa que los Sistemas de Clasificación Basados en Reglas Difusas tienen una mayor precisión y robustez que los sistemas nítidos al ruido de clase.

Palabras Clave: Datos con Ruido; Ruido de Clase; Sistemas basados en Reglas Difusas; Algoritmos Robustos; Clasificación.

1 Introducción

Los Sistemas de Clasificación Basados en Reglas Difusas (SCBRDs) [10] se utilizan por su capacidad de construir modelos lingüísticos interpretables por los usuarios, con la posibilidad de mezclar información de diferentes fuentes (conocimiento de expertos e información de modelos matemáticos). Entre las aplicaciones de los SCBRDs encontramos propuestas en campos como la clasificación estándar [7] o aplicaciones médicas [12].

El objetivo de los algoritmos de clasificación es crear una generalización a partir de un conjunto de ejemplos etiquetados de forma que la precisión al clasificar otros ejemplos no observados previamente sea maximizada. La precisión del clasificador construido estará por tanto determinada por la calidad de los datos de entrenamiento. La calidad de los datos depende, a su vez, de varios componentes [14], entre los que están la fuente y la introducción de los datos, ambas sujetas a error. Por tanto, los conjuntos de datos reales suelen presentar estos tipos de corrupciones que afectan a las interpretaciones, las decisiones tomadas y los modelos creados a partir de los datos.

De esta forma, la precisión de clasificación máxima depende de la calidad de los datos de entrenamiento, pero también de la correcta elección del algoritmo de clasificación para esos datos. Saber qué tipo de algoritmos de clasificación son más adecuados cuando se trabaja con datos con ruido es una cuestión relevante.

En este trabajo se va a estudiar la idoneidad de los SCBRDs frente a los sistemas nítidos robustos cuando entrenan con datos con ruido en las etiquetas de clase de los ejemplos. Nos centraremos en el algoritmo difuso *Fuzzy Unordered Rule Induction Algorithm* (FURIA) [7], que es actualmente uno de los algoritmos con mejor comportamiento para sistemas difusos en clasificación, y que será comparado con C4.5 [11], el cual se considera robusto al ruido y se puede convertir a un conjunto de reglas. La capacidad de un algoritmo de clasificación para evitar el sobreajuste a los datos con ruido es una cuestión clave en este escenario [13]. Debido a las características propias de las reglas difusas y los procesos de inferencia de los SCBRDs que difieren de los que tienen los sistemas nítidos, se espera que los clasificadores construidos por los SCBRDs generalicen mejor los ejemplos del problema, evitando el sobreajuste a los datos con ruido y, por tanto,

obtengan clasificadores más precisos y robustos.

Para comparar FURIA con C4.5, consideraremos 19 conjuntos de datos del repositorio KEEL-dataset [1]. En ellos introduciremos 4 niveles de ruido en los datos de entrenamiento (5 %, 10 %, 15 % y 20 %) mediante la consideración de 2 tipos diferentes de ruido de clase. Por tanto, crearemos 171 conjuntos de datos con ruido. Después, obtendremos la precisión en test de los clasificadores y utilizaremos el test estadístico de Wilcoxon [5] para determinar la significación de las diferencias encontradas. También propondremos dos medidas para estimar la robustez de los algoritmos, basándonos en el deterioro de la precisión y el aumento del número de reglas de los clasificadores entrenados con ruido respecto a aquellos construidos sobre datos sin ruido.

El resto de esta contribución se organiza de la siguiente manera. La Sección 2 presenta una introducción a la clasificación con datos con ruido. A continuación, la Sección 3 describe el método FURIA. La Sección 4 muestra los detalles del marco experimental. La Sección 5 incluye el análisis y los resultados experimentales obtenidos por FURIA frente a C4.5. A continuación, en la Sección 6 se analizan las causas del buen comportamiento de los SCBRDs cuando se trata con ruido de clases. Finalmente, en la Sección 7 presentamos nuestras conclusiones.

2 Clasificación con Datos con Ruido

Los conjuntos de datos reales nunca son perfectos y suelen sufrir de corrupciones que afecten a la interpretación de los datos, los modelos creados y las decisiones tomadas. En clasificación, el ruido puede afectar negativamente al rendimiento del sistema en términos de la precisión de clasificación, el tiempo de construcción y el tamaño e interpretabilidad de los clasificadores [16].

Las etiquetas de clase y los atributos son dos fuentes de información que pueden influir en la calidad de un conjunto de datos. La calidad de las etiquetas de clase representa si la clase de cada ejemplo está asignada correctamente, y la calidad de los atributos indica en que grado los atributos caracterizan correctamente a los ejemplos para realizar la clasificación.

Basándonos en estas dos fuentes de información que definen la calidad de un conjunto de datos de clasificación, podemos distinguir dos tipos de ruido [15]:

1. *Ruido de clase.* Se produce cuando un ejemplo pertenece a una clase incorrecta. El ruido de clase puede atribuirse a varias causas, entre las que están la subjetividad en el proceso de etiquetado, los errores de entrada de datos o no disponer de información suficiente para etiquetar cada ejemplo. Hay dos posibles tipos de ruido de clase:
   - *Ejemplos contradictorios*: los mismos ejemplos aparecen más de una vez y se etiquetan con diferentes clases [6].
   - *Ejemplos mal etiquetados*: los ejemplos están etiquetados con la clase incorrecta [17].
2. *Ruido de atributos.* Hace referencia a corrupciones en los valores de uno o más atributos de ejemplos en el conjunto de datos. El ruido de atributos puede referirse tanto a valores erróneos, como perdidos o "sin importancia".

En la literatura existen principalmente dos formas para tratar con el ruido:

- *Algoritmos robustos.* Se caracterizan porque el ruido les afecta en menor grado. Un ejemplo es C4.5 [11], que utiliza estrategias de poda para reducir las posibilidades de que los árboles que construye se vean sobreajustados a los datos con ruido. Sin embargo, cuando el nivel de ruido es relativamente alto, incluso un algoritmo robusto puede obtener una baja precisión de clasificación.
- *Técnicas de pre-procesamiento de ruido* [3]. Tratan de eliminar los efectos negativos del ruido en los conjuntos de datos antes de crear un clasificador sobre los datos originales.

En esta contribución estudiaremos el ruido de clase producido por ejemplos mal etiquetados, ya que éstos son muy comunes en conjuntos de datos reales [16, 17]. Este tipo de errores se pueden producir en situaciones en las que diferentes clases tienen síntomas similares, como ocurre en los límites de las clases. Compararemos el comportamiento del algoritmo difuso FURIA frente al algoritmo robusto C4.5, con el fin de comprobar el efecto del ruido de clase en la precisión y tamaño de los clasificadores construidos por ambos algoritmos.

3 Sistemas de Clasificación Basados en Reglas Difusas

En esta sección, primero introducimos la notación básica que usaremos después para describir FURIA en la Sub-sección 3.1.

Cualquier problema de clasificación consiste en w ejemplos de entrenamiento $x_p = (x_{p1}, \ldots, x_{pn})$, $p = 1, 2, \ldots, w$, etiquetados con una de M posibles clases $\mathbb{L} = \{\lambda_1, \ldots, \lambda_M\}$, donde x_{pi} es el i-ésimo valor de atributo ($i = 1, 2, \ldots, n$) del p-ésimo ejemplo de entrenamiento. En esta contribución usamos reglas difusas

con una sola clase y un peso asociado a la misma en el consecuente [9]:

$$\begin{array}{c}\text{Regla } R_j : \texttt{IF } x_1 \texttt{ is } A_j^1 \texttt{ AND } \ldots \texttt{ AND } x_n \texttt{ is } A_j^n \\ \texttt{THEN CLASS} = C_j \texttt{ WITH } RW_j\end{array} \quad (1)$$

donde R_j es la etiqueta de la j-ésima regla, $x = (x_1, \ldots, x_n)$ es un vector n-dimensional, A_j^i es un conjunto difuso, C_j es una etiqueta de clase y RW_j es el peso de la regla [8].

3.1 Fuzzy Unordered Rule Induction Algorithm

FURIA [7] se construye sobre el algoritmo RIPPER [4]. El clasificador construido por FURIA usa reglas difusas de la forma presentada en la Ecuación (1), donde A_j^k es un conjunto difuso $I^F = (\phi^{s,L}, \phi^{c,L}, \phi^{c,U}, \phi^{s,U})$ con una función trapezoidal de pertenencia

$$I^F(v) = \begin{cases} 1, & \text{if} \quad \phi^{c,L} \leq v \leq \phi^{c,U} \\ \frac{v-\phi^{s,L}}{\phi^{c,L}-\phi^{s,L}}, & \text{if} \quad \phi^{s,L} \leq v \leq \phi^{c,L} \\ \frac{\phi^{s,U}-v}{\phi^{s,U}-\phi^{c,U}}, & \text{if} \quad \phi^{c,U} \leq v \leq \phi^{s,U} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

y $C_j \in \mathbb{L} = \{\lambda_1, \ldots, \lambda_M\}$ es una etiqueta de clase. El peso RW_j de la regla R_j se calcula como

$$RW_j = \frac{2\frac{|D_T^{(c)}|}{|D_T|} + \sum_{x \in D_T^{(c)}} \mu_{R_j}(x)}{2 + \sum_{x \in D_T} \mu_{R_j}(x)} \quad (3)$$

donde D_T denota el conjunto de entrenamiento, $D_T^{(c)}$ el subconjunto de ejemplos de entrenamiento con etiqueta λ_c y $\mu_{R_j}(x) = \prod_{i=1\ldots n} I_i^F(x_i)$.

Para asignar la clase a un nuevo ejemplo, supongamos que las reglas difusas $R_1, \ldots, R_k$ se han aprendido para la clase λ_c. Para un nuevo ejemplo de test x, el soporte de esta clase está definido por

$$s_c(x) = \sum_{j=1}^{k} \mu_{R_j}(x) RW_j \quad (4)$$

La clase predicha por FURIA es la que tiene el máximo soporte. En caso de un empate entre clases, se elige la que tiene mayor frecuencia. Cuando el ejemplo no es cubierto por ninguna regla, se usa un método que modifica localmente las reglas para que cubran el nuevo ejemplo. Para ello, se comprueba el orden en que los antecedentes aparecen en la regla y se eliminan todos a partir del primero que no empareja con el ejemplo.

FURIA construye la base de reglas difusas a través de dos pasos fundamentales:

1. Aprender un conjunto de reglas para cada clase λ_c del problema, utilizando una estrategia de descomposición uno-contra-todos. Para hacer esto, el algoritmo utiliza RIPPER, que consta de dos pasos fundamentales descritos en [7]: la fase de construcción y la de optimización.

2. Obtener las reglas difusas a través de *fuzzificar* las reglas finales del paso anterior. Cada regla es *fuzzificada* conservando la misma estructura que la original y reemplazando los intervalos originales en el antecedente con intervalos difusos. Para *fuzzificar* un intervalo, se necesitan calcular los cuatro parámetros necesarios para el conjunto difuso trapezoidal a partir del intervalo original (el procedimiento completo se describe en [7]).

4 Marco Experimental

En esta sección, primero se describe en la Sub-sección 4.1 los conjuntos de datos originales. En la Sub-sección 4.2, se describe el proceso de introducción de ruido. Finalmente, en la Sub-sección 4.3 se establece la metodología de comparación entre FURIA y C4.5.

4.1 Conjuntos de Datos Originales

La experimentación se basa en 19 conjuntos de datos del repositorio KEEL-dataset [1] [1]. La Tabla 1 resume sus propiedades. Para cada uno, se muestran el número de ejemplos (#Ej), el número de atributos numéricos (#At), junto con el número de atributos reales y enteros (R/E) y el número de clases (#Cl).

Tabla 1: Conjuntos de datos originales.

Conjunto	#Ex	#At (R/E)	#Cl	Conjunto	#Ex	#At (R/E)	#Cl
contraceptive	1473	9 (0/9)	3	satimage	6435	36 (0/36)	7
ecoli	336	7 (7/0)	8	segment	2310	19 (19/0)	7
glass	214	9 (9/0)	7	sonar	208	60 (60/0)	2
heart	270	13 (1/12)	2	spambase	4597	57 (57/0)	2
ionosphere	351	33 (32/1)	2	thyroid	7200	21 (6/15)	3
iris	150	4 (4/0)	3	twonorm	7400	20 (20/0)	2
page-blocks	5472	10 (4/6)	5	wdbc	569	30 (30/0)	2
penbased	10992	16 (0/16)	10	wine	178	13 (13/0)	3
pima	768	8 (8/0)	2	yeast	1484	8 (8/0)	10
ring	7400	20 (20/0)	2				

La precisión de cada clasificador se estima mediante 5 ejecuciones de una validación cruzada estratificada de 5 particiones. El conjunto de datos se divide en 5 partes con el mismo número de ejemplos y manteniendo la proporción entre clases. Cada partición se usa para evaluar el clasificador construido sobre las cuatro particiones restantes. Este procedimiento se repite cinco veces. Usamos 5 particiones ya que, si cada partición tiene un mayor número de ejemplos, los efectos del ruido son más notables, lo que facilita su análisis.

[1] `http://www.keel.es/datasets.php`

Los algoritmos se han ejecutado con KEEL [2] [2] con los parámetros recomendados por los autores.

4.2 Introducción de Ruido

Debido a que la cantidad de ruido inicialmente presente en los conjuntos de datos no se conoce, se va a utilizar un mecanismo para introducir ruido en los datos de forma controlada. A partir de los 19 conjuntos de datos originales se han creado otros nuevos con ruido de clase en los datos de entrenamiento. Se han tenido en cuenta 4 niveles de ruido: $x = 5\,\%$, $x = 10\,\%$, $x = 15\,\%$ y $x = 20\,\%$.

Para introducir un nivel de ruido de clase del $x\,\%$, se utilizan dos esquemas distintos:

- *Esquema de ruido de clase por pares.* El ruido se introduce siguiendo el esquema de pares utilizado en [17]: dado un par de clases (X, Y), con X la clase mayoritaria e Y la segunda clase mayoritaria, y un nivel de ruido $x\,\%$, un ejemplo con la etiqueta X tiene una probabilidad de $x/100$ de ser incorrectamente etiquetado como Y.
- *Esquema de ruido de clase aleatorio.* En este esquema, un nivel de ruido del $x\,\%$ supone que el $x\,\%$ de los ejemplos se corrompen. Las etiquetas de clase de estos ejemplos se cambian aleatoriamente por otras del dominio de la clase.

Para crear un conjunto de datos con ruido a partir del original, se siguen los siguientes pasos:

1. Un nivel de ruido del $x\,\%$ de un tipo de ruido de clase se introduce en una copia del conjunto de datos original completo.
2. Ambos conjuntos de datos, el original y la copia con ruido, se dividen en cinco partes equivalentes. Es decir, los ejemplos en cada parte de la copia con ruido son los mismos que en la parte correspondiente del conjunto de datos original.
3. Utilizamos una validación cruzada estratificada de 5 particiones para los nuevos conjuntos de datos con ruido. Los conjuntos de entrenamiento se crean con la copia con ruido y los conjuntos de test con la copia original.

4.3 Metodología de Comparación

Para conocer qué tipo de algoritmos, nítidos o difusos, son más precisos y robustos cuando se trata con ruido de clase, comparamos FURIA con C4.5. Esta comparación se realiza usando tres métodos distintos:

[2] `www.keel.es`

1. La precisión proporcionada por los algoritmos de clasificación en los conjuntos de test para cada nivel de ruido, definida como la precisión promedia sobre todos los problemas de clasificación. Para comprobar la significación de las diferencias de precisión en test encontradas, usamos el test estadístico de Wilcoxon [5]. Para cada nivel de ruido, se compara FURIA con C4.5 y se obtienen los p-valores asociados con estas comparaciones.
2. La *pérdida relativa de precisión* (RLA) (Ecuación 5). Se usa para observar la forma en que se ve afectada la precisión del clasificador al aumentar el nivel de ruido con respecto al caso sin ruido:

$$RLA_{x\,\%} = \frac{Acc_{0\,\%} - Acc_{x\,\%}}{Acc_{0\,\%}} \tag{5}$$

   donde $RLA_{x\,\%}$ es la pérdida relativa de precisión en un nivel de ruido del $x\,\%$ y $Acc_{y\,\%}$ es la precisión media en test con un nivel de ruido del $y\,\%$.
3. El *incremento relativo del número de reglas* (RIR) (Ecuación 6). Otro de los aspectos que pueden ser afectados por el ruido es el tamaño del modelo [16, 3] y por lo tanto, el número de reglas puede estar relacionado con la robustez del clasificador:

$$RIR_{x\,\%} = \frac{Rules_{x\,\%} - Rules_{0\,\%}}{Rules_{0\,\%}} \tag{6}$$

   donde $RIR_{x\,\%}$ es el incremento relativo en el número de reglas a un nivel de ruido $x\,\%$ y $Rules_{y\,\%}$ es el número medio de reglas del clasificador aprendido del conjunto de entrenamiento con un nivel de ruido $y\,\%$.

5 Efecto del Ruido de Clase en la Precisión y la Robustez

En esta sección nos centramos en el análisis de la precisión y robustez de FURIA frente a C4.5 cuando entrenan con datos con ruido de clase.

La Tabla 2 muestra los resultados en test de ambos algoritmos en cada nivel de ruido junto con los p-valores asociados. La precisión en test de FURIA es mayor que la de C4.5 en cada nivel y esquema de ruido. Los bajos p-valores obtenidos muestran que esas diferencias son significativas.

Además, cabe destacar el mejor comportamiento de FURIA especialmente con el esquema de ruido de clase aleatorio. Las precisiones en test de C4.5 se ven bastante más afectadas en este caso que las de FURIA. Los p-valores también reflejan este hecho, ya que los p-valores más bajos se obtienen generalmente en este esquema de ruido de clase.

Tabla 2: Precisión en test y p-valores.

		Precisión en test					p-valores				
	Ruido %	**0 %**	**5 %**	**10 %**	**15 %**	**20 %**	**0 %**	**5 %**	**10 %**	**15 %**	**20 %**
Pares	FURIA	85.81	85.37	84.74	84.23	83.10	1.1444E-5	1.9074E-5	9.652E-4	1.6404E-4	2.022E-3
	C4.5	83.93	83.66	82.81	82.25	81.41					
Aleat.	FURIA	85.81	85.17	84.54	84.06	83.66	1.1444E-5	7.63E-6	3.356E-4	3.814E-5	1.1444E-5
	C4.5	83.93	82.97	82.38	81.69	80.28					

Tabla 3: Resultados de RLA y RIR.

		Resultados RLA				Resultados RIR			
	Ruido %	**5 %**	**10 %**	**15 %**	**20 %**	**5 %**	**10 %**	**15 %**	**20 %**
Pares	FURIA	0.0051	0.0127	0.0191	0.0331	-0.03	-0.07	-0.08	-0.10
	C4.5	0.0035	0.0141	0.0212	0.0323	0.05	0.11	0.16	0.18
Aleat.	FURIA	0.0073	0.0152	0.0207	0.0253	-0.01	-0.04	-0.07	-0.08
	C4.5	0.0124	0.0204	0.0285	0.0466	0.10	0.14	0.27	0.30

Los resultados de precisión en entrenamiento, sin embargo, muestran que las diferencias entre FURIA y C4.5 son menores, incluso siendo C4.5 superior en algunos casos. Esto se debe a que C4.5 produce un mayor sobreajuste que FURIA a los datos con ruido.

La Tabla 3 muestra los resultados de RLA y RIR de cada algoritmo en cada nivel de ruido. Como se muestra en esta tabla, los valores de RLA son menores para FURIA que para C4.5 en todos los niveles de ruido considerados, para los esquemas de ruido. Sin embargo, con 5 % y 20 % del esquema de ruido de clase por pares esto no ocurre, aunque estos valores son muy cercanos. Esto demuestra una vez más la mayor robustez de FURIA cuando trata con datos mal etiquetados.

FURIA especialmente destaca en sus resultados de RIR, incluso reducidos cuando se considera ruido respecto al caso sin ruido. El número de reglas de FURIA es en promedio menor que el de C4.5. El mayor incremento del número de reglas de C4.5 que de FURIA se debe a que las reglas de la clase mayoritaria se ven fraccionadas en mayor medida en C4.5 por la aparición de nuevas reglas del resto de clases cuando se introduce ruido, con lo cual sus reglas son más específicas. El método de modificación local de las reglas en FURIA puede que influya también en el menor número de reglas de este algoritmo, pero aún así, se puede concluir que FURIA tiene una mayor robustez que C4.5 respecto al ruido de clase.

6 Análisis de la Mayor Robustez de FURIA con Ruido de Clase

Aunque la robustez de FURIA y C4.5 puede deberse a la estrategia seguida por cada uno para construir la base de reglas, existen otras diferencias que son relevantes para reducir la sensibilidad al ruido. Así, la mayor robustez de FURIA se puede deber a que éste tiene una serie de propiedades que lo hacen diferente del sistema nítido C4.5 cuando trata con problemas con ruido. Algunas de éstas son:

1. *El uso de los conjuntos difusos en los antecedentes de las reglas, en lugar de intervalos nítidos.* Esto permite, por ejemplo, controlar la importancia de la clase de un ejemplo predicha por una regla, en función de que éste caiga en una u otra zona de la función de pertenencia de los antecedentes. Los ejemplos con ruido pueden caer en las zonas con un menor valor de la función de pertenencia de los antecedentes de la regla, perteneciendo a una clase diferente de la predicha por la regla. Por lo tanto, estos ejemplos con ruido se verán influenciados en menor grado por la predicción de esta regla.

2. *La asignación de un peso a cada regla.* Esto, junto con los conjuntos difusos en los antecedentes, permite el solapamiento entre reglas difusas. Si varias reglas cubren un ejemplo, el peso de la regla (y la función de pertenencia de los conjuntos difusos) permitirá determinar la regla más adecuada que cubre este ejemplo. La posible superposición entre las reglas es un hecho muy importante cuando se trata con datos con ruido, ya que supone que las reglas se vean menos afectadas por el ruido que corrompe otras reglas.

3. *La agregación de las predicciones de las reglas difusas con el fin de predecir la clase final de un ejemplo.* Esta es una forma natural y robusta para hacer frente al ruido porque la predicción de la clase de un ejemplo no está determinada sólo por la acción de una sola regla, sino por la intervención de todas o parte de las reglas.

Estas propiedades ayudan a que FURIA sea menos afectado por el ruido de clase que C4.5. Así, este sistema logra un menor sobreajuste al ruido, aumentando la precisión al clasificar ejemplos de test no observados.

7 Conclusiones

En esta contribución se han analizado las ventajas de los SCBRDs, concretamente de FURIA, respecto al sistema nítido C4.5 cuando se entrena con datos con ruido. Para ello, se han considerado dos tipos diferentes de ruido de clase: el esquema de pares y el esquema aleatorio. Hemos creado conjuntos de datos con ruido basándonos en ellos con diferentes niveles de ruido.

Los resultados obtenidos indican que FURIA tiene una mayor precisión y robustez, en términos de la degradación de la precisión y el número de reglas del clasificador, que el sistema nítido robusto C4.5 cuando entrenan con datos con ruido de clase. Estos resultados son particularmente notables con el esquema de ruido aleatorio, el más disruptivo de los estudiados.

Agradecimientos

Apoyado por el Ministerio de Ciencia y Tecnología bajo el proyecto TIN2011-28488 y también por el Proyecto Regional P10-TIC-6858. J. A. Sáez disfruta de una beca FPU del Ministerio de Educación y Ciencia.

Referencias

[1] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, "KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework," *Journal of Multiple-Valued Logic and Soft Computing, in press*, 2010.

[2] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, and F. Herrera, "KEEL: a software tool to assess evolutionary algorithms for data mining problems," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 13, pp. 307–318, 2009.

[3] C. E. Brodley and M. A. Friedl, "Identifying Mislabeled Training Data," *Journal of Artificial Intelligence Research*, vol. 11, pp. 131–167, 1999.

[4] W. W. Cohen, "Fast effective rule induction," in *Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann Publishers, 1995, pp. 115–123.

[5] J. Demšar, "Statistical Comparisons of Classifiers over Multiple Data Sets," *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.

[6] M. A. Hernández and S. J. Stolfo, "Real-world Data is Dirty: Data Cleansing and The Merge/Purge Problem," *Data Mining and Knowledge Discovery*, vol. 2, pp. 9–37, 1998.

[7] J. C. Hühn and E. Hüllermeier, "FURIA: An Algorithm for Unordered Fuzzy Rule Induction," *Data Mining and Knowledge Discovery*, vol. 19, no. 3, pp. 293–319, 2009.

[8] H. Ishibuchi and T. Nakashima, "Effect of rule weights in fuzzy rule-based classification systems," *IEEE Transactions on Fuzzy Systems*, vol. 9, pp. 506–515, 2001.

[9] H. Ishibuchi, K. Nozaki, and H. Tanaka, "Distributed representation of fuzzy rules and its application to pattern classification," *Fuzzy Sets and Systems*, vol. 52, pp. 21–32, 1992.

[10] L. Kuncheva, *Fuzzy classifier design*. Springer—Verlag, 2000.

[11] J. R. Quinlan, *C4.5: programs for machine learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 1993.

[12] G. Schaefer, M. Závišek, and T. Nakashima, "Thermography based breast cancer analysis using statistical features and fuzzy classification," *Pattern Recognition*, vol. 42, no. 6, pp. 1133–1137, 2009.

[13] C.-M. Teng, "Correcting Noisy Data," in *Proceedings of the Sixteenth International Conference on Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 1999, pp. 239–248.

[14] R. Y. Wang, V. C. Storey, and C. P. Firth, "A Framework for Analysis of Data Quality Research," *IEEE Transactions on Knowledge and Data Engineering*, vol. 7, no. 4, pp. 623–640, 1995.

[15] X. Wu, *Knowledge acquisition from databases*. Norwood, NJ, USA: Ablex Publishing Corp., 1996.

[16] X. Zhu and X. Wu, "Class Noise vs. Attribute Noise: A Quantitative Study," *Artificial Intelligence Review*, vol. 22, pp. 177–210, 2004.

[17] X. Zhu, X. Wu, and Q. Chen, "Eliminating class noise in large datasets," in *Proceeding of the Twentieth International Conference on Machine Learning*, 2003, pp. 920–927.

QUALITY ASSESSMENT OF LINGUISTIC DESCRIPTION OF DATA

M. Pereira-Fariña[1], **Luka Eciolaza**[2], **Gracian Trivino**[2]

[1]*Centro de Investigación en Tecnoloxías da Información (CITIUS), University of Santiago de Compostela*, martin.pereira@usc.es
[2]*European Centre for Soft Computing (ECSC), Mieres, Asturias, Spain*, {luka.eciolaza,gracian.trivino}@softcomputing.es

Abstract

One important challenge in the field of automatic description of data consists of assessing the quality of the obtained results. This paper explores a possible approach to solve this problem based on analysis performed by readers that are familiar with the specific context. We show an example with a practical application.

Keywords: Linguistic description of data, quality assessment.

1 INTRODUCTION

Automatic linguistic description of data is a topic that has recently attracted a great attention. It represents an attempt to describe patterns emerging in applications with a strong human-machine interaction by means of linguistic expressions in natural language (NL). The available data about complex phenomena must be interpreted to make imprecise but practical representations that are easier to manage. The main goal of this technique is to make the data more understandable for human beings.

In previous works, we have developed several simple examples of linguistic descriptions of data, e.g., in [15], we describe the behavior of traffic in a roundabout, in [8], we provide linguistic descriptions about the different styles of driving. In [5], linguistic summarizations of weather data are proposed by other authors.

In this field, the evaluation of quality of automatically generated linguistic descriptions is still an open question [1, 6]. Once the automatic description of data is obtained, how can we evaluate its usability from the point of view of human consumption? What kind of indicators could be used to assess the quality of these linguistic descriptions?

Currently, the usual approach to this problem [1, 4, 5] is based on several *ad-hoc* indicators linked with particular applications. For instance, *good summaries* are associated to statements with a high degree of fulfilment. While other characteristics, as the length of the summary, its relevancy or its ambiguity are not considered although they are essential from the point of view of human consumers. In the field of Natural Language Generation, it is proposed a method for evaluating the quality of a generated text in the framework of cryptic crossword clues [12]. The procedure consists of a Turing-style test: the generated clue is the best when the human reviewers cannot distinguish the automatic generated text form the human text. The authors manage a concept of quality subjective, non-compositional and epiphenomenal. There are also some automatic evaluation metrics, but their behavior is not good enough [3].

In this paper, we propose an alternative interpretation about the quality of a text. Our concept of quality is intersubjective and compositional, considering different dimensions of the text that can be evaluated separately. According with the classification provided by authors in [12], our test is of the type *domain expert review*; that is, the reviewers of the text are familiar with the context of the text. The main point is to consider the obtained text as a communicative act and, by means of a set of basic questions, analyze the quality in its multiple dimensions. From each question, we obtain a partial rating that can be integrated with the others to build a global rating about the quality of the generated text.

Our fields of inspiration are Pragmatics [13], the discipline of linguistics that deals with the effects of *context* in any communicative act, and Systemic Functional Linguistics [10] (SFL), that refers to the view of language as a network of systems, or interrelated sets of options for making meaning. We take from them the different dimensions that are relevant in any communicative act to build our quality measure.

Our proposal is developed under the framework of the HITO project (Human Interface by Technology Observation). The main aim of this project is to develop an effective methodology and framework to evaluate on board devices

used in road transport using automatic linguistic descriptions (reports) in order to optimize reliability levels and to guarantee an ergonomic and safe workplace design.

This paper is organized as follows: in section 2 we describe the HITO project as a framework where we develop our first version of a form to measure the quality of a report; in section 3 we propose a hierarchical model of the quality of a linguistic report; in section 4 we explain some experimentation in the context of HITO project and, in section 5, the paper ends with some conclusions.

2 HITO PROJECT

Our participation in the HITO project (Human Interface by Technology Observation) deals with the development of a computational system capable of automatically generating linguistic descriptions of driving activity from vehicle simulator data. Our solution is built over the concept of Granular Linguistic Model of a Phenomenon (GLMP) [7, 8], that seeks to broaden the understanding of the Computational Theory of Perceptions (CTP) [16], representing our approach to it.

This project takes into account different on board devices (mobile phone, GPS, on board pc, tachometer, etc) used in road transport in order to optimize reliability levels and to guarantee an ergonomic and safe workplace design. Different representative parameters of vehicle, environment and driver behavior are collected and evaluated in order to know, depending on the nature and criticality of the information, WHAT, HOW, WHEN and WHERE the on board systems' information should be placed in the driving desk within the vehicles.

The GLMP consists of a network of Perception Mappings (PMs). Each PM receives a set of input Computational Perceptions (CPs) and transmits upwards another CP. We say that each output CP is explained by the PM using a set of input CPs. In the network, each CP covers specific aspects of the phenomenon with certain degree of granularity. Using different aggregation functions and different linguistic expressions, the paradigm GLMP allows the designer to model computationally his/her perceptions about the monitored phenomena.

Fig. 1 shows a GLMP created by the designer to model the perception of *driving quality* as required in the HITO application. This GLMP contains thirteen 1-PMs $p^1_1, p^1_2, \ldots, p^1_{13}$ that provide the corresponding 1-CPs. Examples of these 1CPs are, namely, the angle of the steering wheel, the percentage of applied brake, the angle with the road direction, etc. These 1-CPs are used as inputs of three 2-PMs p^2_1, p^2_2, p^2_3 that produce corresponding 2-CPs, namely, $P^2_1 =$ *steering wheel control* (*loss steering wheel control* is defined as an abrupt manoeuvre), $P^2_2 =$ *vehicle linearity* (refers to the uniformity of the vehicle trajectory) and $P^2_3 =$ *security distance* (it refers to the distance with respect to the vehicle in front). These 2-CPs will explain a top-order CP that answers a general question about the driving quality with three possible associated sentences given by the template "The driving quality is low | medium | high".

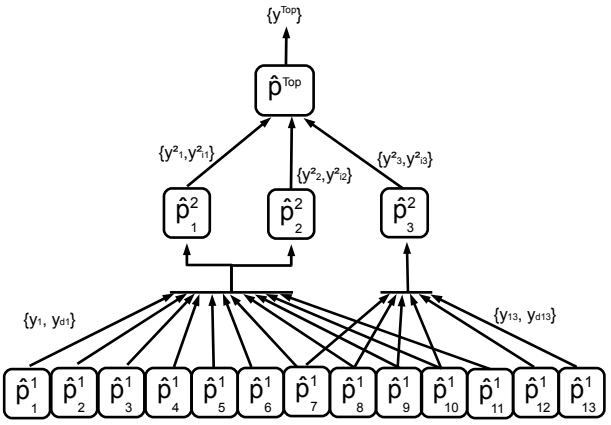


Figure 1: GLMP for representing different details of the perception of driving quality.

The generation of the summary report is performed following a report template customized for the analyzed application. A standard simulation exercise analyzed within the project contains data of 10-15 minutes. Figure 2 is a partial copy of one of this reports (see [7, 8] for more detail).

3 A HIERARCHICAL MODEL OF THE QUALITY OF A LINGUISTIC REPORT

From our point of view, the quality of a linguistic report is a complex concept that can be subdivided into different components until a basic level. We distinguish four layers: the most basic is *layer 1*, that consists of a set of simple questions that a human expert can answer easily, and the most abstract is *layer 4*, which provides the assessment of the quality of the report as a whole. *Layer 2* and *layer 3* are intermediate layers where the rating of the basic questions are aggregated (see Fig. 3).

These four layers are defined as follows:

3.1 LAYER 4

This layer provides a general assessment of the quality degree of the report. Our concept of quality is supported by the consideration of the automatic linguistic description as a communicative act: it refers to the specific scenario of a simulation driving report that must be useful for a human expert in the field.

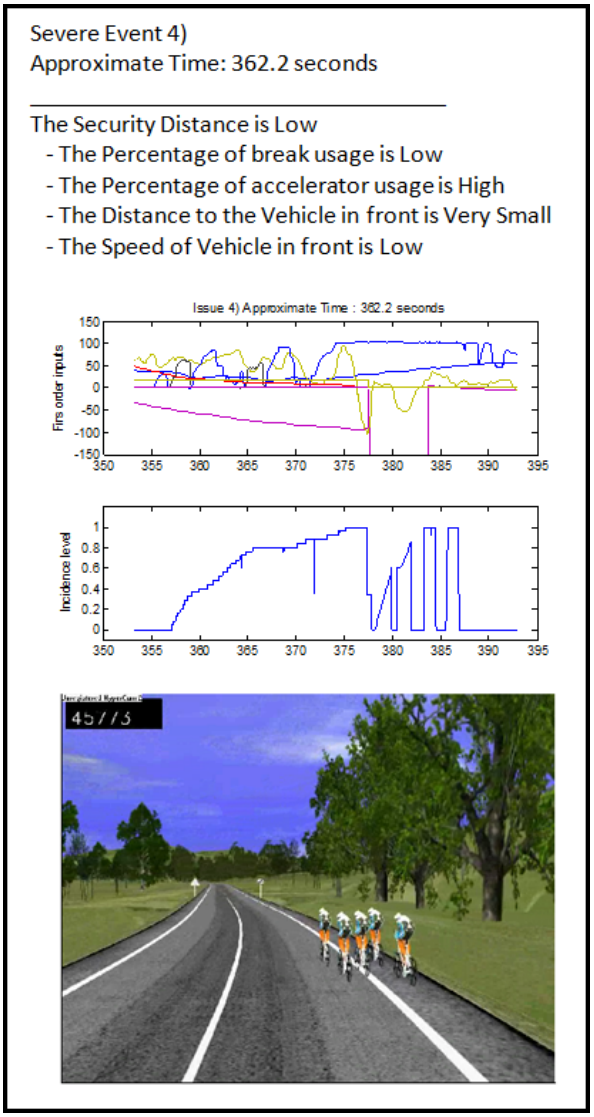


Figure 2: Fragment of a report generated in the HITO Project.

3.2 LAYER 3

The distinction between "what the text implicates" and "what the text says" is inspired in Austin's [2] ideas about NL: in any kind of utterance it is possible to distinguish between the output of the realm semantics ("what the text says") and what is conveyed or accomplished in particular linguistics and social context in or by saying something, including the communicative intentions of the speaker ("what the text implicates"). The borderline between both aspects is not a sharp line and there is a mutual influence between them.

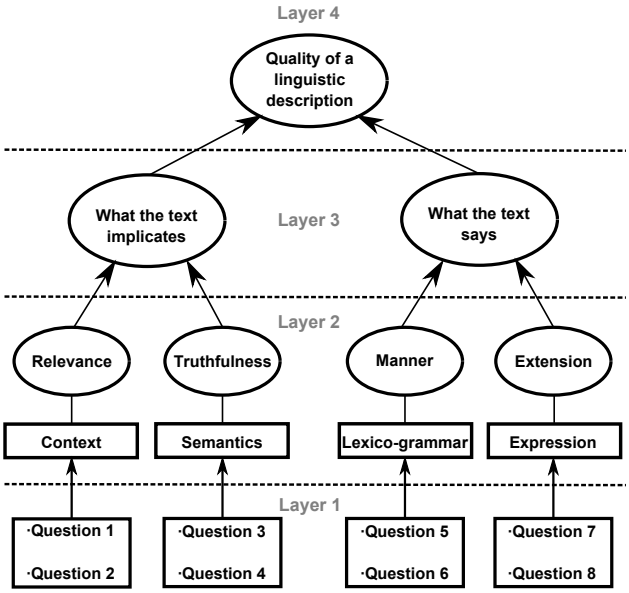


Figure 3: Hierarchical structure of our concept of the quality of a report.

3.3 LAYER 2

Here "what the text implicates" is subdivided into the concepts of *relevance* and *truthfulness* and "what the text says" is subdivided into the concepts of *manner* and *extension*.

The origin of these concepts is the Cooperative Principle [9, p. 24] proposed by H. P. Grice: "Make your conversation contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged". He postulates the existence of four rational principles that govern conversations:

1. Relation: Be relevant.
2. Quality: Try to make your contribution one that is true; do not say what you believe to be false; do not say that for which you lack adequate evidence.
3. Manner: Be perspicuous; avoid obscurity of expression; avoid ambiguity; be brief; be orderly; frame whatever you say in the form most suitable for any reply that would be regarded as appropriate; or, facilitate in your form of expression the appropriate reply.
4. Quantity: Make your contribution as informative as is required; do not make your contribution more informative than is required.

These four concepts belong respectively to the four basic strata of natural language according to SFL [11]. These four strata, that must be analyzed to explore the dimensions of the NL meaning, are briefly described as follows:

Context. The meaning of NL depends on the *domain of experience* (application domain) where is used. The *Context* stratum defines the limits of the specific *domain of experience* under consideration. The basic unit of context is called a *situation type*, and context is globally described by a network of *situation types*. For example, a *situation type* "cooking" can be considered close to a *situation type* "shopping food". However, a *situation type* "working on mathematics" is definitively far from "cooking". For the aim of this paper we can say that, the better defined the context the more relevant the report and the better the quality.

Semantic. The *Semantic* stratum describes the meaning resources stored in the NL-system that are acquired through experiences in terms of meaning features. Human experience are chunked into *figures*, i.e., a *figure* is interpreted as a unit of experience. A *figure* is classified into "doing", "sensing", "saying", and "being". For example in a domain of cooking, *figures* of "doing" are, namely, making soup, adding spices; *figures* of "saying" are, namely, giving an instruction of chef, a response of assistant and so on. For the aim of this paper, the *figures* in the expert's experience correspond with valid sentences in the report. Hence, the higher the degree of matching with the expert's experience the higher the quality.

Lexico-grammar. It contains the grammatical resources in the NL-system to realize the meanings at the *Semantic* stratum as a text. For example, *figures* are realized by a lexico-grammatical units called *clauses* which are, in short, simple sentences. Corresponding to *figures* "doing", "sensing", "saying" and "being", there are *clauses* "material", "mental", "verbal" and "relational". The better use of the resources provided by lexico-grammar the better the report design and the higher the quality.

Expression. It contains the resources to realize the final expression of texts, i.e., phonology or graphology. E.g., in computational applications, the so called TTS (text to speech generator) is a resource in this stratum that could include possibilities to express meaning using voice intonation and volume. For the aim of this paper, these resources are applied to generate the physical format given to the report. The better the use of these possibilities the more effective the report and the better the quality.

Note that, text generation goes from *Context* to *Expression* and text understanding follows the reverse path.

3.4 LAYER 1

In this first approach into the framework of HITO project, *layer 1* consists of a limited number of questions trying to avoid spending the time of the experts and focusing their attention to the more important features. For each one of the mentioned concepts, we propose two questions, which rating goes from 1 to 5.

3.4.1 Relevance

The concept of *relevance* is supported by the Maxim of Relation, by the Sperber and Wilson's Theory of Relevance [14] and linked with the stratum of context, fundamentally related to Field.

The relevance of a text is strictly linked with the speaker's communicative intention [14] into a Field: the hearer uses all kinds of information available to get at what the speaker intended to convey. So, assuming a representational theory of mind (the speaker/hearer can postulate the inferences that the other can do), two principles are established:

1. First (cognitive) principle of relevance: Human cognition is geared towards the maximization of relevance (that is, to the achievement of as many contextual (cognitive) effects as possible for as little processing effort as possible).

2. Second (communicative) principle of relevance: Every act of ostensible communication (e.g. an utterance) communicates a presumption of its own optimal relevance.

From a psychological point of view, the first principle evaluates the relevance of a text in terms of the steps that the hearer needs to discover the communicative intention of the speaker: the most likely meaning of an expression is that one which is achieved with a smaller number of intermediate steps. The second one indicates that the speaker also should make his intentions clear to the speaker for achieving a successful communication. Therefore, a text is relevant enough if it can be worth processing and if it is compatible with the communicator's abilities and preferences.

This definition of relevance is very subjective, since it depends of the background of each one of the speakers. However, all the speakers involved in a communicative situation can share the context in a very high degree; therefore, we can talk about an *intersubjective* concept of relevance, being the text relevant to all the involved people. For instance, to analyze the behavior of the engine in a truck, any expert considers that the fuel consumption is a critical value while other aspects, like its noise, are more debatable and depends more on the subjective considerations of the expert.

Table 1 contains our proposal of questions to evaluate the relevance: question 1 considers subjective relevance meanwhile question 2 considers the intersubjective relevance.

Table 1: Questions about Relevance.

1	Indicate in which degree the content of this report belongs to the application domain of the HITO project.
2	Indicate in which degree you identify the type of results expressed as the type of results expressed by yourself.

3.4.2 Truthfulness

The concept of *truthfulness* is inspired on the Maxim of Quality and on the stratum of Semantics in its component of *Ideational Semantics*. Now, the report is considered from the viewpoint of its truth conditions, that is, the *possible world* described in the report that allows to the reader understand if the expressed ideas can be true or not. For instance, in a report about driving, it is necessary to know the kind of road (motorway or highway) to determine if the driver drives rightly or not. Our proposal of questions to evaluate the truthfulness of a report are shown in Table 2.

Table 2: Questions about Truthfulness.

3	After observing the behavior of the driver in the simulator, do you agree with the assessed global quality?
4	Do you agree with the provided explanations?

With the ratings of the relevance and the truthfulness, we obtain the partial rating of the report respect to "what the text implicates" (layer 3).

3.4.3 Manner

The concept of *manner* is inspired on the Maxim of Manner and analyze the text in its lexico-grammar stratum. In this case, the report is evaluated in its utterance dimension, considering if it uses the adequate vocabulary, if the order of the ideas is the most appropriate or if the used expressions are the right ones. Table 3 contains our proposal of questions.

Table 3: Questions about Manner.

5	Indicate in which degree the vocabulary is correctly used.
6	Indicate in which degree the ideas are correctly ordered to facilitate the comprehension of the report.

3.4.4 Extension

It is inspired in the Maxim of Quantity and analyze the phonology-graphology stratum of the report. It tries to evaluate if the quantity of statements used in the report are the right for expressing the relevant ideas or if there is redundancy in the text. Bad *quantity* reports contains too much or insufficient statements respect to the fact that try to be described; good *quantity* reports contains the right statements to understand the fact. To evaluate this aspect, we propose two questions that are shown in Table 4.

Table 4: Questions about Quantity.

7	Indicate in which degree the format of the report, including the use of figures and punctuation, is the most adequate.
8	Indicate in which degree you consider that the extension of the report is the right respect to its content.

4 EXPERIMENTAL LAYOUT

In a typical experimental layout, we selected a group of five people familiar with the HITO project. We gave them three documents, namely, a description of the context of the simulation, the report obtained automatically (see Figure 2) and the form to evaluate it.

After 15 minutes, they had to fulfil the form with the different questions shown in Tables 1, 2, 3 and 4. The results of the different reviewers are shown in Table 5, where *R1*, *R2*, *R3*, *R4* and *R5* denotes each reviewer that participated in the experiment.

Table 5: Results of the evaluation.

	R1	R2	R3	R4	R5
Question 1	5	5	5	5	5
Question 2	4	3	3	5	2
Question 3	4	4	4	4	3
Question 4	5	4	3	5	3
Question 5	3	3	4	4	4
Question 6	4	4	2	5	4
Question 7	4	3	2	4	3
Question 8	4	4	4	5	5
Average	4.1	3.8	3.4	4.6	3.7

The data in Table 5 are an example of the type of practical results obtained during the development of the project.

The average denotes the global rating for the report. Two of the five ratings are higher than 4 and the minimum rating is 3.7. The average is calculated as the arithmetic mean of all questions; therefore, all the questions have the same importance for the global rating. It is relevant to note that all the reviewers agree with the maximum rating in question 1, the question that tries to determine if the content of the report belongs to the application domain. Question 7 has the worst ratings, following by question 2 and question 5.

After analyzing these results the report designers must interpret them in order to improve the HITO report generator. For instance, the rating of question 7 indicates that the format of the report and the different techniques for expressing the information must be improved, because the reviewers consider that it is not good enough. Question 2, that analyzes the way as the results are expressed, also has a low global rating. Therefore, this is other aspect that the designers should improve to make better reports.

5 CONCLUSIONS

This paper is a partial result of a long term research project aimed to generate linguistic descriptions of complex phenomena.

This is only a first approach to the problem the assessment quality of a generated linguistic report. In future works, we will explore into SFL looking for building a more detailed set of questions.

Here, we have described our solution to the problem of report validation in the framework of the HITO project. This solution is being used successfully to improve the quality of the generated reports.

Acknowledgements

This work was supported in part by the Spanish Ministry of Science and Innovation (grants TIN2008-00040, PSS-370100-2007-12 and TIN2008-06890-C02-01) and the Spanish Ministry for Education (FPU Fellowship Program).

References

[1] A. Alvarez-Alvarez, D. Sanchez-Valdes, and G. Trivino. Automatic linguistic description about relevant features of the mars' surface. In *Proc. 11th International Conference on Intelligent Systems Design and Applications (ISDA)*, Córdoba, Spain, 2011.

[2] J. L. Austin. *How to Do Things with Words*. Oxford University Press, Oxford, 1962.

[3] A. Belz and E. Reiter. Comparing automatic and human evaluation of nlg system. In *Proc. 1st International Natural Language Generation Conference (INLG-2000)*, pages 313–320, Trento, Italy, 2006.

[4] R. Castillo-Ortega, N. Marín, and D. Sánchez. A hierarchical fuzzy partition of time dimension for linguistic summarization of differential time series. In *Actas del XV Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF 2010)*, pages 267–272, 2010.

[5] F. Díaz-Hermida and A. Bugarín. Linguistic summarization of data with probabilistic fuzzy quantifiers. In *Actas del XV Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF 2010)*, pages 635–642, 2010.

[6] F. Díaz-Hermida, A. Ramos-Soto, and A. Bugarín. On the role of fuzzy quantified statements in linguistic summarization. In *Proc. 11th International Conference on Intelligent Systems Design and Applications (ISDA)*, Córdoba, Spain, 2011.

[7] L. Eciolaza and G. Trivino. Linguistic reporting of driver behavior: Summary and event description. In *Proc. 11th International Conference on Intelligent Systems Design and Applications (ISDA)*, Córdoba, Spain, 2011.

[8] L. Eciolaza, G. Trivino, B. Delgado, J. Rojas, and M. Sevillano. Fuzzy linguistic reporting in driving simulators. In *Proc. 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems*, pages 953–960, Paris, France, 2011.

[9] H. P. Grice. *The Logic of Grammar*, chapter Logic and conversation, pages 64–75. Dickenson, Encino, 1967.

[10] M. A. K. Halliday and M. I. Matthiessen. *Construing Experience through Meaning: A Language-based Approach to Cognition*. Continuum, 1999.

[11] M.A.K. Halliday. *Introduction to Functional Grammar*. Edward Arnold, London, 1994.

[12] D. Hardcastle and D. Scott. Can we evaluate the quality of generated text? In *Proc. Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may 2008. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/.

[13] K. Korta and J. Perry. Pragmatics, Summer 2011. Edward N. Zalta (ed.).

[14] D. Sperber and D. Wilson. *Relevance: Communication and Cognition*. Blackwell, Oxford, 1986.

[15] G. Trivino, A. Sanchez, A. S. Montemayor, J. J. Pantrigo, R. Cabido, and E. G. Pardo. Linguistic description of traffic in a roundabout. In *Proc. 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 2158–2165, Barcelona, Spain, July 18-23 2010.

[16] L. A. Zadeh. From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. *IEEE Transactions on Circuits and Systems*, 45(1):105–119, 1999.

LINGUISTIC DESCRIPTION OF THE TRAFFIC EVOLUTION IN ROADS

Daniel Sanchez-Valdes, **Alberto Alvarez-Alvarez**, **Gracian Trivino**

European Centre for Soft Computing (ECSC), Mieres (Asturias)
{daniel.sanchezv, alberto.alvarez, gracian.trivino}@softcomputing.es

Abstract

We present some results of a long term research project in the field of automatic linguistic description of phenomena. In this paper, we address the problem of generating linguistic descriptions of how a phenomenon changes between two states. We use a Fuzzy Finite State Machine to model phenomena evolving in time. Then we aggregate the information about different states to generate valid and relevant descriptions of changes. We demonstrate our approach using a practical example.

Keywords: Traffic Analysis, Linguistic Summarization.

1 INTRODUCTION

New technologies allow us to acquire and store data about complex phenomena. However, in order to convert data into relevant linguistic descriptions, it is needed to interpret and represent the data in an understandable way, according to each type of situation, and specially, making it useful to achieve specific goals.

For instance, in traffic management systems, one important problem consists of dynamically generating the most adequate natural language messages to communicate with drivers using electronic panels installed in the roads. Currently, these messages are produced by human experts but this task can be tough and tedious. Moreover, the relation between the amount of data to analyze and the number of experts available is growing dramatically. As SCADA systems were created to control and monitor graphically industrial or infrastructure processes, our goal is create computational systems that can interpret and describe linguistically the large amount of available information.

During the last years, we are working in a long term research project in the field of linguistic descriptions of data (see, e.g., [2, 5, 6, 7]) including a first work where we generated linguistic descriptions about the traffic behavior in a roundabout [10].

In this paper, we focused on the perception of change. We explored possibilities to perform linguistic descriptions of how the traffic evolves in time. We investigated how to model the meaning of sentences such as *"the phenomenon is changing from state A to state B"*. This idea has also been explored by other researchers with the aim of summarizing network flow statistics [8].

In order to model the evolution of phenomena in time, we have used our previous works on Fuzzy Finite State Machines (FFSM). In that context, the main contribution of this work consists of extending the use of the FFSM's output function to be used with this aim.

The remainder of this paper is organized as follows. Section 2 presents the main concepts related to fuzzy finite state machines. Section 3 describes how to use FFSMs for modeling the level of service in roads. Section 4 explains how we generate linguistic descriptions about the traffic evolution. A practical example with simulated data is developed in Section 5. Finally, Section 6 draws some conclusions and introduces some future research works.

2 FUZZY FINITE STATE MACHINES

In this section, we describe the main concepts and elements of a FFSM as a paradigm for phenomena modeling allowing experts to build comprehensible fuzzy linguistic models in an easier way. In our framework, a FFSM is a tuple $\{Q, U, f, Y, g\}$, where:

- Q is the state of the monitored phenomenon.
- U is the input vector of the FFSM.
- f is the transition function which calculates the state of the phenomenon.

- Y is the output vector of the FFSM.
- g is the output function which calculates the output vector.

Each of these components are described in the following subsections. The interested reader can refer to our previous works [1, 11] for a more detailed description.

2.1 FUZZY STATES (Q)

The state of a monitored phenomenon (Q) is defined as a linguistic variable [14] that takes its values in the set of linguistic labels $\{q_1, q_2, \ldots, q_n\}$, with n being the number of fuzzy states. Every fuzzy state represents the pattern of a repetitive situation and it is represented numerically by a state activation vector:

$S[t] = (s_1[t], s_2[t], \ldots, s_n[t])$, where $s_i[t] \in [0,1]$ and $\sum_{i=1}^{n} s_i[t] = 1$. S_0 is defined as the initial value of the state activation vector, i.e., $S_0 = S[t=0]$.

2.2 INPUT VECTOR (U)

U is the input vector $(u_1, u_2, ..., u_{n_u})$, with n_u being the number of input variables. U is a set of linguistic variables obtained after fuzzification of numerical data. Typically, u_i can be directly obtained from sensor data or by applying some calculations to raw measures, e.g., the derivative or integral of a signal, or the combination of several signals. The domain of values that u_i can take is represented by a set of linguistic labels, $A_{u_i} = \{A^1_{u_i}, A^2_{u_i}, \ldots, A^{n_i}_{u_i}\}$, with n_i being the number of linguistic labels of the linguistic variable u_i.

2.3 TRANSITION FUNCTION (f)

The transition function (f) calculates, at each time instant, the next value of the state activation vector: $S[t+1] = f(U[t], S[t])$. It is implemented by means of a set of fuzzy rules.

Once the expert has identified the relevant states in the model, she/he must define the allowed transitions among states. There are rules R_{ii} to remain in a state q_i, and rules R_{ij} to change from state q_i to state q_j. If a transition is forbidden in the FFSM, it will have no fuzzy rules associated.

A generic expression of a rule is of the form:

R_{ij}: **IF** ($S[t]$ is q_i) **AND** C_{ij} **THEN** $S[t+1]$ is q_j

where:

- The first term in the antecedent ($S[t]$ is q_i) computes the degree of activation of the state q_i in the time instant t, i.e., $s_i[t]$. With this mechanism, we only allow the FFSM to change from the state q_i to the state q_j (or to remain in state q_i, when $i = j$).
- The second term in the antecedent C_{ij} describes the constraints imposed on the input variables. For example: C_{ij} = ($u_1[t]$ is $A^3_{u_1}$) **AND** ($u_2[t]$ is $A^4_{u_2}$ **OR** $u_3[t]$ is $A^1_{u_3}$).
- Finally, the consequent of the rule defines the next value of the state activation vector $S[t+1]$. The activation degree of each state q_j (s_j) will be the sum of the firing degrees of the rules predicting that state (ω_{ij}) as shown in Equation 1. The firing degrees ω_{ij} are calculated using the minimum for the AND operator and the maximum for the OR operator.

$$s_j[t+1] = \begin{cases} \frac{\sum_{i=1}^{n} \omega_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} \omega_{ij}} & \text{if} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_{ij} \neq 0 \\ s_j[t] & \text{if} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_{ij} = 0 \end{cases} \tag{1}$$

2.4 OUTPUT VECTOR (Y)

Y is the output vector: $(y_1, y_2, ..., y_{n_y})$, with n_y being the number of output variables. Y is a summary of the characteristics of the system evolution that are relevant for the application.

2.5 OUTPUT FUNCTION (g)

The output function (g) calculates, at each time instant, the next value of the output vector: $Y[t] = f(U[t], S[t])$. For example, the most simple implementation of g is $Y[t] = S[t]$.

3 LEVEL OF SERVICE MODELING USING A FUZZY FINITE STATE MACHINE

This section presents the design of the main elements needed to build a FFSM for modeling the level of service in roads (*LOS*). The *LOS* is a measure used by traffic engineers to determine the effectiveness of elements of transportation infrastructure. *LOS* is most commonly used to analyze highways by categorizing traffic flow with corresponding safe driving conditions. The concept has also been applied to intersections and water supply.

3.1 FUZZY STATES

The Highway Capacity Manual [3] lists the following levels of service:

- $A \rightarrow$ Free-flow operation.

- $B \rightarrow$ Reasonably free flow, with ability to maneuver is only slightly restricted and the effects of minor incidents still are easily absorbed.
- $C \rightarrow$ Stable flow, speeds at or near free-flow and queues may form.
- $D \rightarrow$ Approaching unstable flow, speeds decline slightly with increasing flows while density increases more quickly.
- $E \rightarrow$ Unstable flow, with operation near or at capacity and no usable gaps in the traffic stream.
- $F \rightarrow$ Forced or breakdown flow, queues form behind breakdown points and demand is greater than capacity.

Therefore, in this application, we have defined six different fuzzy states which directly describe the *LOS*. Therefore, the set of fuzzy states will be: $Q = LOS = \{A,B,C,D,E,F\}$.

3.2 INPUT VECTOR

In our experiments, we have used two input variables, the traffic density (*density*), and its trend (*density trend*). The set of linguistic labels for each linguistic variable are strong fuzzy partitions (SFPs) [9], uniformly distributed over the universe of discourse of each input variable:

- $density \rightarrow \{EL,VL,L,H,VH,EH\}$
- $density\ trend \rightarrow \{N,Z,P\}$

where the terms EL,VL,L,H,VH,EH denote extremely low, very low, low, medium, high, very high, and extremely high respectively; and the terms N,Z,P denote a negative, zero (stable), and positive trend.

3.3 TRANSITION FUNCTION

The definition of which transitions are allowed and which are not can be easily represented by means of the state diagram, which captures the expert knowledge about the application. Figure 1 shows how we use the FFSM to define constraints on the possibilities to change of state. From this state diagram, it can be recognized that there are 16 fuzzy rules overall in the system: 6 rules to remain in each state and other 10 to change between states.

This rule base is defined using expert knowledge based on the descriptions of the Highway Capacity Manual [3], which links terms related to traffic density and its evolution along time. In Figure 2, can be seen the set of rules. It is clear how the system evolution is given by the traffic density, which has a different associated linguistic term for each *LOS*; and its trend, which governs the change to a better or worse state if the density trend is negative or positive respectively.

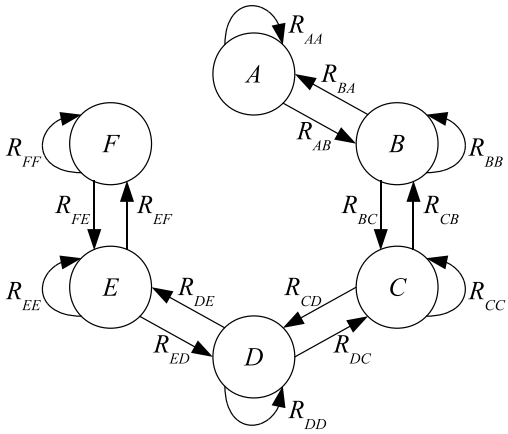


Figure 1: State diagram of the FFSM for *LOS* modeling.

3.4 OUTPUT VECTOR AND OUTPUT FUNCTION

In previous works [2], we have defined the concept of Computational Perception (CP). A CP is the computational model of a unit of information acquired by the designer of a computational application about the phenomenon to be modeled. A CP is a couple $(P,W) = \{(p_1,w_1),(p_2,w_2),\ldots,(p_n,w_n)\}$ where:

P is a set of natural language (NL) propositions that represent the meaning of the CP. These sentences can be either simple, e.g., $p_i =$*"The traffic density is low"* or more complex, e.g., $p_i =$*"The situation could get worse during the next hours"*.

$W \in [0,1]$ is the vector of validity degrees assigned to each p_i in the specific context. The concept of validity depends on the application, e.g., it is a function of the truthfulness and relevancy of the sentence in its context of use.

Here, the output vector Y is a CP, i.e., a set of possible sentences with associated degrees of validity. As it was said before, the main contribution of our work consists in, not only inform about the *LOS* in each time instant, but inform if the *LOS* is changing from one state to another. In the same way, we inform periodically that the *LOS* is keeping constant if it is not changing in time. Every sentence in the report is the one with the highest validity degree at each time instant. The set of sentences is comprised by four types of NL propositions for each level L at each time instant t:

1. $p_1^L[t]$: If the *LOS* is keeping constant: *"The level of service is keeping constant in level L"*.

R_{AA}: **IF** ($LOS[t]$ is A) **AND** ($density[t]$ is EL) **THEN** $LOS[t+1]$ is A
R_{BB}: **IF** ($LOS[t]$ is B) **AND** ($density[t]$ is VL) **THEN** $LOS[t+1]$ is B
R_{CC}: **IF** ($LOS[t]$ is C) **AND** ($density[t]$ is L) **THEN** $LOS[t+1]$ is C
R_{DD}: **IF** ($LOS[t]$ is D) **AND** ($density[t]$ is H) **THEN** $LOS[t+1]$ is D
R_{EE}: **IF** ($LOS[t]$ is E) **AND** ($density[t]$ is VH) **THEN** $LOS[t+1]$ is E
R_{FF}: **IF** ($LOS[t]$ is F) **AND** ($density[t]$ is EH) **THEN** $LOS[t+1]$ is F
R_{AB}: **IF** ($LOS[t]$ is A) **AND** ($density[t]$ is VL) **AND** ($density\ trend[t]$ is P) **THEN** $LOS[t+1]$ is B
R_{BC}: **IF** ($LOS[t]$ is B) **AND** ($density[t]$ is L) **AND** ($density\ trend[t]$ is P) **THEN** $LOS[t+1]$ is C
R_{CD}: **IF** ($LOS[t]$ is C) **AND** ($density[t]$ is H) **AND** ($density\ trend[t]$ is P) **THEN** $LOS[t+1]$ is D
R_{DE}: **IF** ($LOS[t]$ is D) **AND** ($density[t]$ is VH) **AND** ($density\ trend[t]$ is P) **THEN** $LOS[t+1]$ is E
R_{EF}: **IF** ($LOS[t]$ is E) **AND** ($density[t]$ is EH) **AND** ($density\ trend[t]$ is P) **THEN** $LOS[t+1]$ is F
R_{BA}: **IF** ($LOS[t]$ is B) **AND** ($density[t]$ is EL) **AND** ($density\ trend[t]$ is N) **THEN** $LOS[t+1]$ is A
R_{CB}: **IF** ($LOS[t]$ is C) **AND** ($density[t]$ is VL) **AND** ($density\ trend[t]$ is N) **THEN** $LOS[t+1]$ is B
R_{DC}: **IF** ($LOS[t]$ is D) **AND** ($density[t]$ is L) **AND** ($density\ trend[t]$ is N) **THEN** $LOS[t+1]$ is C
R_{ED}: **IF** ($LOS[t]$ is E) **AND** ($density[t]$ is H) **AND** ($density\ trend[t]$ is N) **THEN** $LOS[t+1]$ is D
R_{FE}: **IF** ($LOS[t]$ is F) **AND** ($density[t]$ is VH) **AND** ($density\ trend[t]$ is N) **THEN** $LOS[t+1]$ is E

Figure 2: Rule base.

2. $p_2^L[t]$: If the *LOS* is changing: *"The level of service is changing from level L to level L'"*.
3. $p_3^{L'}[t]$: If the *LOS* has recently changed: *"The level of service of the road has changed to level L'"*.
4. $p_4^L[t]$: If the change of *LOS* has not been completed: *"The level of service of the road has returned to level L"*.

Therefore, we have to define an output function g that calculates the validity degrees W of this set of NL propositions P. This output function analyzes the trend of each level of service, in such a way that at each time instant we can determine if a certain level L is decreasing (D_L), keeping constant (KC_L), or increasing (I_L) by fuzzifying its derivative. Moreover, we also know that there has been a change in the level of service (denoted by $C_{LL'}$), which occurs when a certain level L, which was higher than other level L', becomes smaller than L', this is a binary indicator which takes value 0 when there is not a change and 1 when a change is produced. After a carefully experimentation, we have defined the validity degrees of each type of sentence at each time instant t as follows:

1. $w_1^L[t] = \min(KC_L, 1 - C_{LL'})$. A *LOS* is keeping constant when the previous level was the same ($C_{LL'} = 0$) and it is keeping constant (KC_L).
2. $w_2^L[t] = \min(D_L, I_{L'}, 1 - C_{LL'})$. A level L is changing to a level L' when L is decreasing (D_L), L' is increasing ($I_{L'}$), and the change between levels has not been completed yet ($C_{LL'} = 0$).
3. $w_3^{L'}[t] = C_{LL'}$. A *LOS* has recently changed to a level L' when the *LOS* of the previous period was different to the actual one ($C_{LL'} = 1$).
4. $w_4^L[t] = \min(I_L, D_{L'}, w_2^L[t-1])$. In some many cases a *LOS* could have been changing but the change was not completed. For example, the *LOS* could be changing from level E to level F but suddenly the density decreases and the change is stopped, keeping at level E. This case is recognized when the system was changing from level L to level L' ($w_2^L[t-1]$) but the level L starts to increase (I_L) and the expected level L' starts to decrease ($D_{L'}$).

4 LINGUISTIC DESCRIPTION

Apart from the goal of obtaining suitable texts to be showed to drivers using the road screens, the linguistic reports can be studied by traffic experts with the aim of understanding changes in traffic behavior and anticipating possible effects that would occur, thanks to the study of past events and the accumulated knowledge. Due to the fact that our system provides us with information related to the evolution of the *LOS*, the developed application provides two types of linguistic description report.

4.1 PUNCTUAL TRAFFIC REPORT

A punctual or eventual report about the traffic behavior is given. The periodicity of these reports depends on the precision that the designer wants (one minute, five minutes, ten minutes, etc.). Each report informs about the *LOS* trend (changing, recently changed or keeping constant), the traffic density in the last period of time and the traffic density trend.

The information about traffic density and density trend in the last period of time is directly extracted from the input vector of the FFSM. Therefore, the type of sentences are of the form: *"In the last period the traffic density has been*

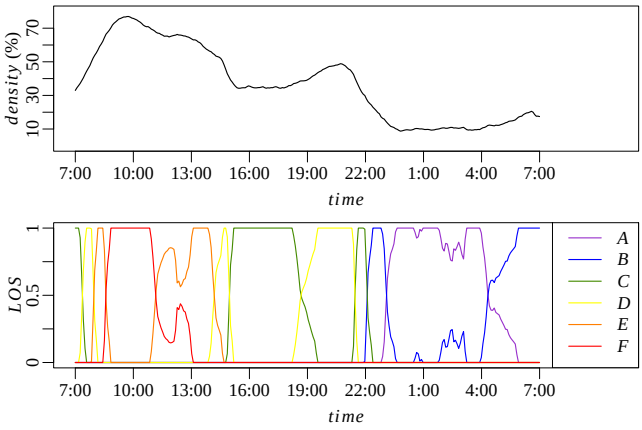


Figure 3: Graphical representation of the traffic density and the activation degrees of the different levels of service.

{extremely low | very low | low | high | very high | extremely high} and the traffic density is {decreasing | keeping constant | increasing}". Similarly to the *LOS* trend, the sentence with the highest validity degree is chosen at each time instant. This type of information shows us that it is possible that these variables are changing while the *LOS* is keeping constant, e.g., traffic density can be increasing inside the level *C* but this does not mean that *LOS* is changing from level *C* to level *D*. This is the difference between changes inside a state and changes between states.

4.2 TRAFFIC SUMMARY REPORT

Our application generates a traffic summary report that summarizes traffic behavior throughout a period of time, for example, a day. A full day is divided into a set of periods that can give relevant information about the traffic progress.

Traffic experts decide how many periods they want to analyze separately, in order to verify the differences existing among them. This type of information allows them to extract conclusions and to implement appropriate measures to improve the quality of traffic (improving infrastructure, notice drivers and so on). For example, one type of differentiation could be analyzing separately sunrise, morning, midday, afternoon, evening and night. In the same way, the designer could decide to segment the day into smaller periods and to extract information about periods of different sizes. The traffic summary report also gives information relative to the average traffic density and the average level of service in each period of time. The following sentences show an example of this type of report: *"Between 08:00 and 10:00, the average traffic density has been {extremely low | very low | low | high | very high | extremely high}. The average level of service was {A|B|C|D|E|F}"*.

Moreover, to complete the traffic summary, the report may include a global linguistic fuzzy summary based in the summarizers proposed by Yager [12]:

"Q the LOS has been R"

where

- *Q* is a fuzzy quantifier [13] applied on the cardinality of the perception *''the LOS has been R"*. The set of linguistic labels for each quantifier are strong fuzzy partitions denoted by the expressions *never*, *few*, *sometimes* and *many times*.
- *R* is the summarizer, in this case the set of possible levels of service.

Last information is really important to summarize traffic behavior once the day has finished; this fact allows to compare the state and trend of traffic in a specific road which study is interesting to obtain conclusions, check the need to redirect the traffic, etc.

5 PRACTICAL EXAMPLE

To test the performance of our application, we have used a simulator of traffic data that simulates real traffic behaviors. The performance of the simulator is based on the study and analysis of databases of expert traffic systems from traffic control centers of important cities, such as Madrid, Valencia, or Sevilla [4].

One of the main differences between our application and the current solutions is that we provide users with linguistic descriptions that complement numerical and graphical data. Based on the knowledge of traffic experts, we have used the simulator to generate random data that recreates the traffic behavior each five minutes. Fig. 3 shows the evolution of the traffic density along a typical working day and the activation degrees of the different levels of service.

Every report, daily or periodic, is accompanied by its reference of time, either the period of the day or the specific time of measurement. Results are consistent and show accurately the simulated situations. Some examples of sentences in the report are as follows:

- **Traffic summary report:**
  - Between 08:00 and 10:00, the average traffic density has been extremely high. The average level of service was *F*.
  - Between 16:00 and 18:00, the average traffic density has been high. The average level of service was *C*.
  - During the day, the level of service has been many times *A*, has been some times *B* and *C*, and has been few times *D*, *E* and *F*.

- **Punctual report:**
  - At 10:55, the level of service is changing from F to E. The traffic density is extremely high and it is decreasing.
  - At 11:10, the level of service of the road has changed to E. The traffic density is very high and it is decreasing.
  - At 11:20, the level of service is keeping constant in level E. The traffic density is very high and it is keeping constant.
  - At 12:10, the level of service is changing from E to F. The traffic density is very high and it is increasing.
  - At 12:20, the level of service of the road has returned to level E. The traffic density is very high and it is decreasing.
  - At 12:25, the level of service is keeping constant in level E. The traffic density is very high and it is keeping constant.

6 CONCLUDING REMARKS

The most of phenomena in our environment are dynamic rather than static. This paper deals with the challenge of generating linguistic descriptions of phenomena evolving in time.

This is a provisional result of a long term project devoted to linguistic descriptions of data, nevertheless, the presented practical application shows how this research line is already delivering useful practical results.

Acknowledgements

This work has been funded by the Spanish Government (MICINN) under project TIN2008-06890-C02-01.

References

[1] Alberto Alvarez-Alvarez, José M. Alonso, Gracian Trivino, Noelia Hernández, Fernando Herranz, Ángel Llamazares, and Manuel Ocaña. Human activity recognition applying computational intelligence techniques for fusing information related to WiFi positioning and body posture. In *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pages 1881–1885, July 2010.

[2] Alberto Alvarez-Alvarez, Daniel Sanchez-Valdes, and Gracian Trivino. Automatic linguistic description about relevant features of the Mars' surface. *Accepted for presentation in the 11th International Conference on Intelligent Systems Design and Applications (ISDA)*. 2011.

[3] National Research Council (U.S.). Transportation Research Board. *Highway capacity manual*. Special report. Transportation Research Board, National Research Council, 1985.

[4] Portal del tráfico urbano de la ciudad de Sevilla. http://www.trajano.com.

[5] L. Eciolaza, G. Trivino, B. Delgado, J. Rojas, and M. Sevillano. Fuzzy linguistic reporting in driving simulators. In *Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Paris, France*, pages 30–37, 2011.

[6] Luka Eciolaza and Gracian Trivino. Linguistic reporting of driver behavior: Summary and event description. *Accepted for presentation in the 11th International Conference on Intelligent Systems Design and Applications (ISDA)*.

[7] S. Mendez-Nunez and G. Trivino. Combining semantic web technologies and computational theory of perceptions for text generation in financial analysis. In *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pages 953–960, July 2010.

[8] F. M. Pouzols, A. Barriga, D. R. Lopez, and Sanchez S. Solano. In *Proceedings of the 2008 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Hong Kong, China*.

[9] Enrique H. Ruspini. A new approach to clustering. *Information and Control*, 15:22–32, 1969.

[10] G. Trivino, A. Sanchez, A. S. Montemayor, J. J. Pantrigo, R. Cabido, and E. G. Pardo. Linguistic description of traffic in a roundabout. In *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Barcelona, Spain*, pages 2158–2165, July 2010.

[11] Gracian Trivino, Alberto Alvarez-Alvarez, and Gonzalo Bailador. Application of the computational theory of perceptions to human gait pattern recognition. *Pattern Recognition*, 43(7):2572–2581, 2010.

[12] Ronald R. Yager. Fuzzy summaries in database mining. In *In Proceedings 11th Conference on Artificial Intelligence for Applications*, 1995.

[13] L. A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computing and Mathematics with Applications*, 9:149–184, 1983.

[14] Lotfi A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Parts I, II, and III. Information Sciences*, 8, 8, 9:199–249, 301–357, 43–80, 1975.

BUILDING LINGUISTIC SUMMARIES WITH F-CUBE FACTORY

Rita Castillo-Ortega[1], **Nicolás Marín**[1], **Carlos Molina**[2], **Daniel Sánchez**[1,3]

[1]*Department of Computer Science and A.I., University of Granada, Spain*, {rita,nicm}@decsai.ugr.es
[2]*Department of Computer Science, University of Jaén, Spain*, carlosmo@ujaen.es
[3]*European Centre for Soft Computing, Mieres, Asturias, Spain*, daniel.sanchezf@softcomputing.es

Abstract

In this paper, a new OLAP operation to obtain fuzzy linguistic summaries coming from time series data stored in data cubes is presented, implemented and integrated in a user-friendly tool within the F-Cube Factory platform. Giving a linguistic result for an OLAP operation has several advantages. First, natural language is one of the best ways of presenting results to human users as it is their inherent way of communication. Second, sentences in natural language are well suited to summarize a huge amount of data in an understandable manner. We show an illustrative example of the new operation using a beta implementation in F-Cube Factory.

Keywords: Linguistic summarization, Time series, Multidimensional data model, OLAP, Business Intelligence, Fuzzy Logic.

1 Introduction

Companies and organizations produce and consume great amounts of data during their daily operation. There is an increasing need for being able to handle these data that has motivated the development of many different data management tools. In this context, Business Intelligence solutions help managers to obtain a better understanding of their commercial procedures and operations in order to support better business decision-making, by means of tasks such as decision analysis, prediction, or forecasting [2]. The use of the multidimensional data model and OLAP (OnLine Analytical Processing) operations to query them is of crucial importance in Business Intelligence [12].

The multidimensional data model is based on the use of data cubes. Data cubes are sets of data related to a given fact, whose context is described by several dimensions that may be hierarchically defined. This way, each cell of the cube contains data concerning a determined context defined by given members of each of the dimensions. Multidimensional systems can also benefit on the use of fuzzy sets theory in order to improve their management of imperfect data and to develop a more friendly communication with users. In this sense, the F-Cube Factory system proposed in [5] is able to manage data cubes in which facts and/or dimensions are fuzzy, allowing the user to employ imprecise concepts in the representation and operation with data, including data mining tasks like the discovery of fuzzy association rules within data cubes.

In this paper we go a step forward in the way of getting a communication with the user through the use of natural language. We describe how we have incorporated further capabilities to F-Cube Factory related to querying data cubes with time dimension, providing linguistic results to the user. The temporal dimension is among the most popular dimensions within a data cube structure, due to the importance of time in all the activities carried out by humans. This fact has encouraged a big number of researches to develop different approaches to *time series data mining* [1] and *time series summarization* using natural language [9, 4, 1, 8], though most of them out of the area of multidimensional databases. The concept of fuzzy summaries obtained from multi-dimensional databases has been investigated for instance in [10]. Also related to linguistic summarization are the proposals in [13, 15]. Ontologies of terms have been used for a multi-criterion approach to data summarization [14]. However, these approaches to summarization within multidimensional databases have not focused in the particularities of the time dimension.

The paper is organized as follows: in Section 2 we present the fuzzy multidimensional data model and the F-Cube Factory system (2.1), and the problem of linguistic summarization of time series (2.2). In Section 3 the formulation of new fuzzy OLAP operations on data cubes with time dimension, and its integration into F-Cube Factory, are presented. We show an illustrative example in Section 4. Finally, some conclusions are given in Section 5.

2 Preliminary work

This section is devoted to present some concepts and tools developed in previous works, in which we can find the basis of the present work. We will start with some ideas about the multidimensional data model and its application in F-Cube Factory. In the second part of the section we will review a specific method with two different strategies to carry out linguistic summaries of time series data.

2.1 Multidimensional data model and F-Cube Factory

The fuzzy multidimensional model proposed in [5] can be briefly described as follows:

Definition 1 *A **dimension** is a tuple $d = (l, \leq_d, l_\perp, l_\top)$ where $l = \{l_1, \ldots, l_m\}$ so that each l_i is a set of values $l_i = \{c_{i1}, ..., c_{im_i}\}$ and $l_i \cap l_j = \emptyset$ if $i \neq j$, and $\leq_d$ is a partial order relation between the elements of l so that $l_i \leq_d l_k$ if $\forall c_{ij} \in l_i \Rightarrow \exists c_{kp} \in l_k / c_{ij} \subseteq c_{kp}$. $l_\perp$ and $l_\top$ are two elements of l so that $\forall l_i \in l \; l_\perp \leq_d l_i \leq_d l_\top$.*

We denote level to each element l_i. To identify the level l of the dimension d we will use $d.l$. The two special levels $l_\perp$ and $l_\top$ will be called *base level* and *top level* respectively. The partial order relation in a dimension is what gives the hierarchical relation between levels.

Definition 2 *For each pair of levels l_i and l_j such that $l_j \in H_i$, we have the relation $\mu_{ij} : l_i \times l_j \rightarrow [0,1]$ and we call this the* **kinship relation**. *H_i is the set of* children of l_i, *$H_i = \{l_j | l_j \neq l_i \wedge l_j \leq_d l_i \wedge \neg \exists l_k, l_j \leq_d l_k \leq_d l_i\}$*

If we use only the values 0 and 1 for kinship and we only allow an element to be included with degree 1 by an unique element of its parent levels, this relation represents a crisp hierarchy. If we relax these conditions and we allow to use values in the interval [0,1] without any other limitation, we have a fuzzy hierarchical relation.

Definition 3 *We say that any pair (h, α) is a* **fact** *when h is an m-tuple on the attributes domain we want to analyze, and $\alpha \in [0,1]$.*

The value α controls the influence of the fact in the analysis. The imprecision of the data is managed by assigning an α value representing this imprecision. Now we can define the structure of a fuzzy DataCube.

Definition 4 *A **DataCube** is a tuple $C = (D, l_b, F, A, H)$ such that $D = (d_1, ..., d_n)$ is a set of dimensions, $l_b = (l_{1b}, ..., l_{nb})$ is a set of levels such that l_{ib} belongs to d_i, $F = R \cup \emptyset$ where R is the set of facts and $\emptyset$ is a special symbol, A is an application defined as $A : l_{1b} \times ... \times l_{nb} \rightarrow F$, giving the relation between the dimensions and the facts defined, and H is an object of type history with information regarding the obtention of the current data cube.*

In order to provide support for the definition and management of data within this model, the system F-Cube Factory has been developed [5]. The system is built using client/server architecture. The server implements the main functionality over the data cubes (definition, management, queries, aggregation operators, user views operators, API for data cube access, etc.). The client is web based and is thought to be light enough to be used in a personal computer and to give an intuitive access to the functionality of the server (hiding the complexity of using a DML or DDL to the user). For a more detailed explication of the structure and the operations, see [11, 6].

2.2 Linguistic summarization of time series data

In our work we consider a linguistic summary of data as a set of type II quantified sentences of the form "Q of D are A", where Q is a linguistic quantifier, and A, D are fuzzy subsets of a finite crisp set X. In our case, D is related to the time dimension and A is related to the variable domain. An example of a produced sentence could be "Most of (Q) the days in the early morning (D) the wind speed is low or very low (A)".

In paper [3], two strategies to obtain linguistic summaries of historical information stored in data warehouses are presented. These approaches work with time series data obtained by means of OLAP operations on data cubes with time dimension.

The aforementioned strategies take advantage of a hierarchical fuzzy partition on dimensions inherited from the data cube structure. Thanks to the hierarchical organization of dimensions, granularity in the data cube can vary according to the selected level chosen in each dimension. The set of OLAP operations allows the user to take advantage of this fact and query the data cube in order to get the desired data in each moment.

As time is crucial for human activities a big part of the existing data cubes on data warehouses includes the so called temporal dimension. The application of OLAP operations to query data cubes with temporal dimension give us time series data as result. Suppose that the time dimension is described in its finest grained level of granularity by members $T = \{t_1, ..., t_m\}$. Then, the above mentioned query result will have the following form: $TS = \{< t_1, v_1 >, ..., < t_m, v_m >\}$, where every v_i is a value of the basic domain of the variable V.

In order to linguistically describe the information of this result, we have considered to use fuzziness as follows:

- The basic domain of variable V under study is partitioned by a set of linguistic labels E=$\{E_1, ..., E_s\}$.

- The time dimension is hierarchically organized in n levels, namely, L=$\{L_1, ..., L_n\}$. Each level L_i has associated a partition $\{D_{i,1}, ..., D_{i,p_i}\}$ of the basic time domain.

There is no restriction concerning the form of the membership function of a label apart from that it must be normalized. In our approach, we will use trapezoidal functions. When necessary, labels $D_{i,j}$ in the time dimension can be the union of a set of trapezoidal functions. Some constraints have been added regarding the way partitions are done and also regarding the structure of the hierarchy.

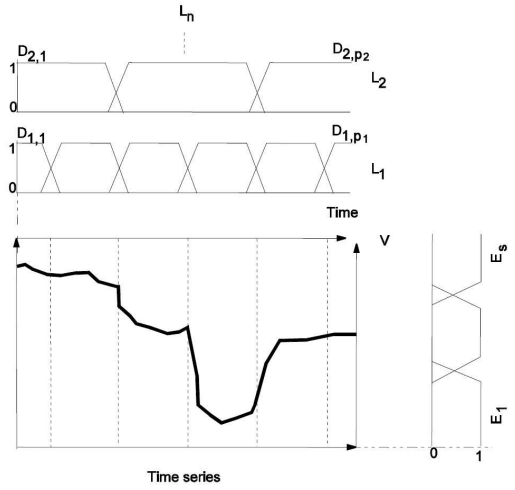


Figure 1: General context for the summarization of a time series.

Figure 1 illustrates the above described context for the time series. The way F-Cube factory defines its hierarchical dimensions fulfills the above mentioned restrictions.

In the context of our research some requirements are imposed to summaries in order to consider them as good quality summaries.

- The accomplishment degree of every sentence must be greater than or equal to a certain threshold τ (accuracy).
- The set of quantified sentences must be as small as possible (brevity).
- The union of the supports of all the time periods $D_{i,j}$ in the sentences of the summary must be T (coverage).

Each of the proposed methods implements an alternative *greedy* strategy in order to accomplish these requirements. The complete *coverage* of the time dimension is reached by means of several restrictions imposed to the time dimension and also by the exploration strategy of the algorithm. In order to asses the *accuracy* of the sentences we use the GD method (introduced in [7]) to calculate the degree of accomplishment, which has to be greater than a user-provided threshold τ. To ensure *brevity* of the summary the algorithm will prefer summaries of the more abstract levels and it will try to avoid the exploration of less abstract levels. Choosing one or another of the strategies is a question of preference, none of the algorithms is better than the other. Moreover, in some occasions, both greedy strategies to explore the solution space could lead to the same optimal result (or very similar ones) depending on the input data series, the linguistic framework, and the selected parameters.

Roughly speaking, the algorithms work as follows. Taking the before mentioned protoform "Q of D are A", the algorithms try to find the combination of Q and A that gives the better accomplishment degree for a certain D. The process is repeated for a certain set of periods $D_{i,j}$. In each step, the time period is obtained exploring the time hierarchy from more abstract levels to more precise levels. For each D combinations of Q and A are proved and the one with greatest GD value is chosen. If this value is greater than or equal to τ the resulting sentence will take part of the final summary, if not, the algorithms repeat the procedure for the labels $D_{i-1,j}$ in the next level (a more precise one) with no empty intersection with the studied $D_{i,j}$.

Summing-up, the user has to provide the following information:

- The strategy he/she desires to use.
- A time series TS.
- A hierarchical fuzzy partition of the time dimension D.
- A linguistic variable E.
- A totally ordered subset $\{Q_1..Q_{qmax}\}$ of a coherent family of quantifiers Q.
- A minimum accomplishment degree threshold τ for quantified sentences.
- For each level i, the maximum number of: a) quantifiers to use ($Qbound_i$), and b) labels to group together ($Gbound_i$).

Many of these are suggested and/or given default values in order to facilitate the communication with the user. The result calculated by the algorithm and given to the user will be a summary of TS comprised of a set of quantified sentences. Further information is provided in [3].

3 New querying capabilities on F-Cube Factory

F-Cube Factory implements the usual OLAP operations *slice and dice*, related to the selection of data in a cube, *roll-up and drill-down*, devoted to change the granularity of the data cube going up and down through the different abstraction levels in the dimensions' hierarchies, and *pivot* (also known as *rotate*), for providing an alternative presentation of the data. For numerical data, usual aggregation functions include the maximum/minimum, average, addition, etc. All the above-mentioned operations are extended to fuzzy facts and dimensions. F-Cube Factory incorporates data mining abilities (fuzzy association rules can be obtained from data cubes) as well.

We are interested here in particular case of aggregation function that can be used, for example, in particular cases of the *roll up* operation. That is, in a data cube with a time dimension, we can allow the user to obtain a new data cube without this dimension, where cells values are appropriate linguistic summaries of the time series that has to be aggregated in the roll up.

Our proposal in this paper is to provide a linguistic summary as the result of this operation, by using the time series summarization techniques we have previously proposed. We have implemented this operation within F-Cube Factory, defining the time hierarchy as a fuzzy dimension using the capabilities of the system. We have also incorporated the operation in a beta extension of the interface of F-Cube Factory, allowing the user to see a graphical depiction of the time series and the linguistic description. We provide more information about this interface by means of an illustrative example in the next section.

4 Experimentation

In this section we will see how to use the proposed aggregation function with F-Cube Factory using a simple data set. The medical centre data set is a toy set representing the patient inflow along the time to a certain medical centres (365 measures per year, centre, and gender). Figure 2 represents the structure of the data cube of the medical centre data set, and the structure of the cube obtained after operations performed with F-Cube Factory, including a linguistic summary.

Figures 3 and 4 will assist us to clarify the linguistic background that will help the user to obtain a more understandable result. Figure 3 represents the time dimension that has a hierarchical structure with 3 levels. The least abstract level is composed by 12 labels representing "fuzzy months", the following ones are based on meteorological concepts describing somehow the seasonal behavior of weather.

Figure 4 depicts the domain of the variable under study. The domain is partitioned using five fuzzy sets. The sets are obtained utilizing a relative partition taking into account the minimum and maximum value of the observed measures. There are various ways of building the partition but the study of them is above our scope in this moment.

The partition of the variable and the hierarchy in the time dimension are created and stored in the database using the capabilities of F-Cube Factory. Figure 5 shows the interface for defining the different levels of the time hierarchy. A similar interface is available for the dimensions related to the value of the variable.

The following parameters have been employed for the linguistic summarization: the family of quantifiers is Most of = (0, 0.8, 0.9, 1), At least 80% = (0, 0.7, 0.8,

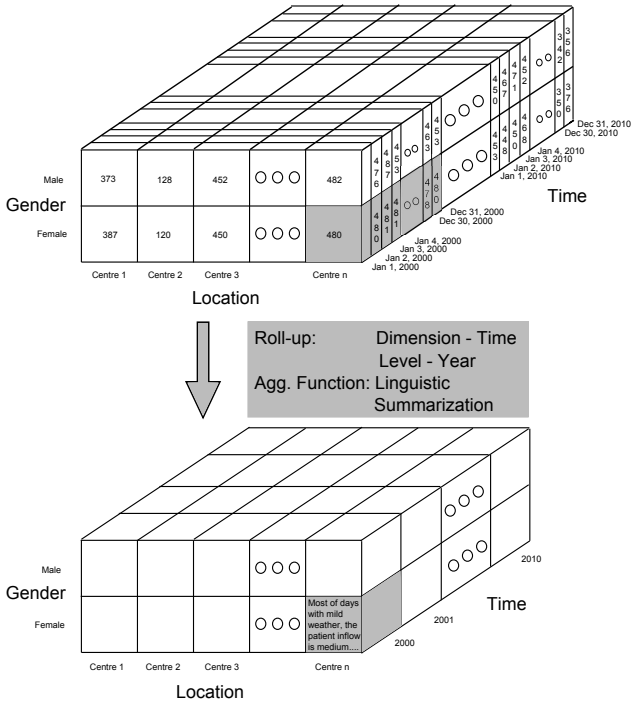


Figure 2: Medical centre data cube structure and cube resulting from operations involving a linguistic summary.

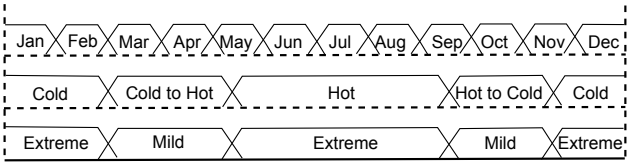


Figure 3: Fuzzy hierarchy on time dimension.

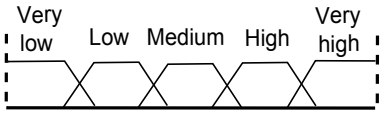


Figure 4: Fuzzy partition on patient inflow.

1), At least 70% = (0, 0.6, 0.7, 1), At least 60% = (0, 0.5, 0.6, 1)}, the accuracy threshold has been set to $\tau = 0.8$, and the boundaries are $Qbound_i = 3$ and $Gbound_i = 2$ in all the levels of the time dimension. These parameters are offered by default by F-Cube Factory, that allows the user to specify others if desired by using the graphic interface shown in Figure 6.

Once the new operation is applied over the medical centre data cube a new data cube is created. This cube is composed by a linguistic summary in each cell instead of sets of numerical measures. The next step is to select the specific result that the user need to analyze. This is, select the linguistic summary describing a specific time period, gender and center.

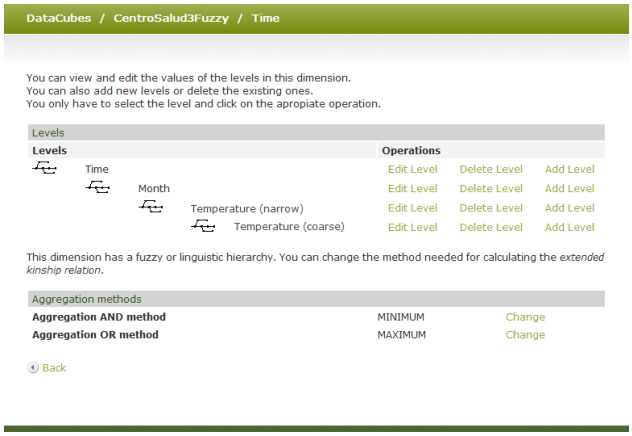


Figure 5: F-Cube Factory interface for managing dimensions, showing levels of time dimension in our example.

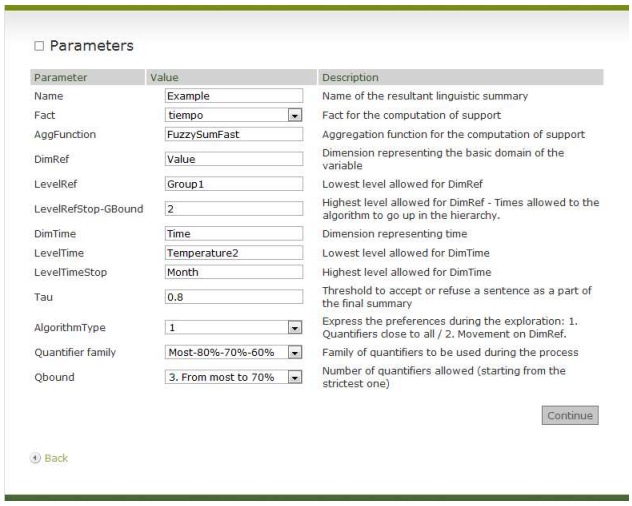


Figure 6: Patient inflow data series: Parameters.

Figure 7 depicts the obtained solution of the new operation in a specific data cube's cell. The figure has three well differentiated areas. The superior part represents graphically the patient inflow series. Below, the interface of F-Cube Factory shows the linguistic summary that describes the given data series. This final result is obtained after a postprocessing step which presents the list of sentences in a more concise way. However, the user can also see in a detailed way the list of quantified sentences as they are obtained by the algorithm (raw result). Both, the processed result and the raw result can be listened by the user. A multimedia reproductor is available for listening the processed result. Regarding the quantified sentences, each of them can be selected and heard as well. Apart from that utility, when the user marks a quantified sentence, the appropriate fuzzy area in the graph is highlighted (a crisp approximation). This way the user can analyze and study in more detail the final result.

5 Conclusions and future work

In this work we have worked on a new OLAP aggregation operation called *linguistic summary* that can be used together with OLAP operations that involve aggregation, such as roll up. With this new capability the system is able to offer the user a linguistic summary of a data cube containing at least two dimensions. The new operation has been integrated F-Cube Factory, a human-friendly tool for interaction with data based on a fuzzy multidimensional model. In the future we will work on integrating other fuzzy operations yielding linguistic summaries in F-Cube Factory.

Acknowledgements

Part of the research was supported by the Andalusian Government (Junta de Andalucía, Consejería de Innovación, Ciencia y Empresa) under project P07-TIC-03175 *Representación y Manipulación de Objetos Imperfectos en Problemas de Integración de Datos: Una Aplicación a los Almacenes de Objetos de Aprendizaje*. Part of the research was supported by the Spanish Government (Science and Innovation Department) under project TIN2009-08296.

References

[1] I. Z. Batyrshin and L. Sheremetov. Perception-based approach to time series data mining. *Appl. Soft Comput.*, 8(3):1211–1221, 2008.

[2] I. Z. Batyrshin and T. Sudkamp. Perception based data mining and decision support systems. *Int. J. Approx. Reasoning*, 48(1):1–3, 2008.

[3] R. Castillo-Ortega, N. Marín, and D. Sánchez. A fuzzy approach to the linguistic summarization of time series. *Journal of Multiple-Valued Logic and Soft Computing (JMVLSC)*, 17(2,3):157–182, 2011.

[4] D. Chiang, L. R. Chow, and Y. Wang. Mining time series data by a fuzzy linguistic summary system. *Fuzzy Sets Syst.*, 112:419–432, June 2000.

[5] M. Delgado, C. Molina, L. Rodríguez Ariza, D. Sánchez, and M. A. Vila Miranda. F-cube factory: a fuzzy olap system for supporting imprecision. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15:59–81, 2007.

[6] M. Delgado, C. Molina, D. Sánchez, M. A. Vila, and L. Rodriguez-Ariza. A linguistic hierarchy for datacube dimensions modelling. In *Current Issues in Data and Knowledge Engineering*, pages 167–176. EXIT, 2004.

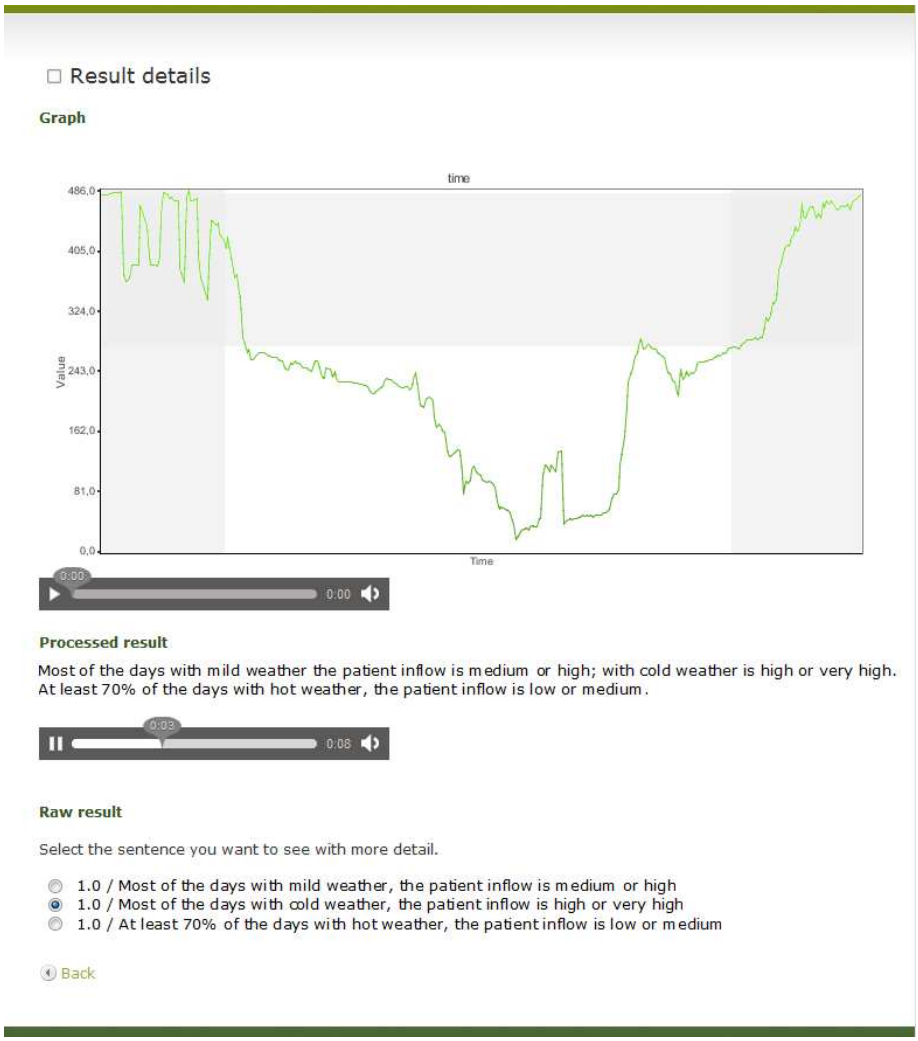


Figure 7: Patient inflow data series: Result of the aggregation of the time series in one cell of the data cube.

[7] M. Delgado, D. Sánchez, and M.A. Vila. Fuzzy cardinality based evaluation of quantified sentences. *International Journal of Approximate Reasoning*, 23:23–66, 2000.

[8] J. Kacprzyk, A. Wilbik, and S. Zadrozny. Linguistic summarization of time series using a fuzzy quantifier driven aggregation. *Fuzzy Sets and Systems*, 159(12):1485–1499, 2008.

[9] I. Kobayashi and N. Okumura. Verbalizing time-series data: With an example of stock price trends. In *IFSA/EUSFLAT Conf.*, pages 234–239, 2009.

[10] A. Laurent. A new approach for the generation of fuzzy summaries based on fuzzy multidimensional databases. *Intell. Data Anal.*, 7:155–177, April 2003.

[11] C. Molina, L. Rodríguez Ariza, D. Sánchez, and M. A. Vila Miranda. A new fuzzy multidimensional model. *IEEE T. Fuzzy Systems*, 14(6):897–912, 2006.

[12] J. A. O'Brien and G. M. Marakas. *Management information systems*. McGraw-Hill, 8 edition, 2008.

[13] D. Pilarski. Linguistic summarization of databases with quantirius: a reduction algorithm for generated summaries. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 18(3):305–331, 2010.

[14] R R. Yager and F. E. Petry. A multicriteria approach to data summarization using concept ontologies. *IEEE T. Fuzzy Systems*, 14(6):767–780, 2006.

[15] L. Zhang, Z. Pei, and H. Chen. Extracting fuzzy linguistic summaries based on including degree theory and fca. In *Proceedings of the 12th international Fuzzy Systems Association world congress on Foundations of Fuzzy Logic and Soft Computing*, IFSA '07, pages 273–283, Berlin, Heidelberg, 2007. Springer-Verlag.

SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS BORROSAS BIPOLARES

J. Tinguaro Rodríguez[1], Begoña Vitoriano[1], Javier Montero[1]

[1] *Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid*
{jtrodrig,bvitoriano,monty}@mat.ucm.es

Resumen

A menudo, el conjunto de clases de un problema de clasificación presenta una estructura en relación con las características propias de cada contexto de aplicación. Sin embargo, los modelos de clasificación habituales no suelen contemplar esa estructura en sus procesos de aprendizaje o razonamiento. Mediante el uso de pesos con carácter positivo y negativo (i.e. pesos bipolares), este trabajo propone una revisión de las nociones básicas de los SCBRB dirigida a extender su potencia y su adaptación a conjuntos de clases estructurados.

Palabras Clave: sistemas de clasificación basados en reglas borrosas, bipolaridad

1 INTRODUCCIÓN

Este trabajo propone la introducción de un enfoque bipolar (véase [5] para una introducción al concepto de *bipolaridad*) en el marco de los sistemas de clasificación basados en reglas borrosas (SCBRB, ver por ejemplo [4,6]). Esta propuesta esta enfocada, por tanto, a la replicación, en un contexto de clasificación supervisada borrosa, de la habilidad humana de aprender y razonar en base a información con carácter positivo y negativo.

Es preciso observar que este trabajo no propone una nueva técnica sofisticada de aprendizaje o razonamiento (aunque los clasificadores bipolares aquí propuestos admiten la aplicación posterior de procesos adaptativos, por ejemplo), sino que más bien plantea una revisión (o mejor dicho una extensión) de las bases conceptuales de los SCBRB con el objetivo de ampliar su capacidad de generalización y su adaptabilidad a los diferentes contextos de aplicación.

En particular, como se verá, a través del enfoque propuesto aquí y con base en la noción de disimilaridad propuesta en [9,10], es posible dotar de estructura al conjunto de clases e introducir esa estructura en los procesos de aprendizaje y razonamiento propios de los SCBRB.

2 FUNDAMENTOS

2.1. CLASIFICADORES BORROSOS

Un problema de clasificación consiste en asignar una *clase* C_j, de entre un conjunto de N_C clases $\zeta = \{C_1,...,C_{N_C}\}$, a un *objeto* (o patrón, instancia, ejemplo, etc.) x descrito por medio de un vector $x = (x_1,...,x_n) \in U_{X_1} \times ... \times U_{X_n} = U_X$ de *atributos* $X_1,...,X_n$, donde U_{X_i} es el rango del atributo X_i, $i = 1,...,n$. En un marco de *aprendizaje supervisado*, se provee un conjunto de ejemplos de entrenamiento $\{(x_1^p,...,x_n^p;C^p), p = 1,...,m\}, C^p \in \zeta \; \forall p$, y el objetivo es *aprender* una función $C: U_{X_1} \times ... \times U_{X_n} \to \zeta$ que generalice la muestra de entrenamiento de una manera óptima respecto a cierto criterio $\varepsilon(C)$ que mide la precisión del clasificador. Normalmente, el objetivo es obtener un clasificador que asigne clases a las instancias de una muestra de validación con el menor error posible. Diferentes metodologías pueden ser aplicadas en el desarrollo de un clasificador: reglas borrosas, redes neuronales, árboles de clasificación, maquinas de soporte vectorial, etc. Los clasificadores que utilizan reglas borrosas son conocidos como *sistemas de clasificación basados en reglas borrosas* (SCBRB) o *clasificadores borrosos*.

En general, una regla borrosa *R* es una expresión del tipo

$$R: \textbf{si } X_1 \text{ es } A_{1j_1} \text{ y} ... \text{ y } X_n \text{ es } A_{nj_n} \textbf{ entonces } Y \text{ es } (r_1,...,r_{N_C}), \quad (1)$$

que valora la relación existente entre unos valores o *etiquetas* lingüísticas $A_{1j_1},...,A_{nj_n}$ de los atributos $X_1,...,X_n$ (situados en la *premisa* de la regla, i.e., antes de la palabra entonces) y las clases $C_j \in \zeta$ que pueden ser asignadas a la variable de clasificación Y (en el *consecuente* de la regla, i.e. después de *entonces*). Los *grados de confianza* (o pesos) $r_j \in [0,1]$ en el consecuente cuantifican la fuerza del vínculo entre la premisa y cada clase C_j.

El significado o *semántica* de las etiquetas A_{ij_i} se define a través de *funciones de pertenencia* $\mu_{A_{ij_i}} : U_{X_i} \to [0,1]$ de modo que, para todo $x_i \in U_{X_i}$, el valor $\mu_{A_{ij_i}}(x_i)$ especifica el grado en que la afirmación "x_i is A_{ij_i}" es verdadera. Por supuesto, esto es lo mismo que decir que, para cada $i = 1,...,n$, X_i es una *variable lingüística* [12] que toma valores sobre los l_i *conjuntos borrosos* A_{ij_i}, $j_i = 1,...,l_i$, definidos sobre el *universo* U_{X_i}. La *compatibilidad* $\mu_A(x)$ de un vector $x = (x_1,...,x_n)$ con el conjunto borroso $A_{1j_1} \wedge ... \wedge A_{nj_n}$ (y por tanto con la premisa de *R*) es evaluada habitualmente mediante el producto

$$\mu_A(x) = \mu_{A_{1j_1}}(x_1) \cdot ... \cdot \mu_{A_{nj_n}}(x_n). \quad (2)$$

Para realzar la interpretabilidad del clasificador resultante, se suele asumir que los conjuntos borrosos A_{ij_i} son especificados por el usuario o decisor previamente a la etapa e aprendizaje. Esto lleva a considerar una partición sobre U_X, de modo que cada celda $A = A_{1j_1} \times ... \times A_{nj_n}$ es identificada con la premisa de una regla. Por tanto, la etapa de aprendizaje puede ser reducida a especificar, para cada premisa *A*, los pesos $r_j(A)$ que miden la validez de cada subregla $R_j : A \Rightarrow C_j$. En minería de datos y aprendizaje automático, la medida habitual de validez de una subregla $R_j : A \Rightarrow C_j$ viene dada, a partir de una muestra de entrenamiento $(x_1^p,...,x_n^p;C^p)$, $p = 1,...,m$, por su *confianza*

$$cf(A \Rightarrow C_j) = \frac{\sum_{p=1/C^p=C_j}^{m} \mu_A(x^p)}{\sum_{p=1}^{m} \mu_A(x^p)} \quad (3)$$

Nótese que la confianza $cf(A \Rightarrow C_j)$ es una aproximación de la probabilidad condicional $P(C_j \mid A)$, por lo que su uso como peso de reglas está justificado en el caso de los SCBRB (en consonancia con [3,7]). Una posible definición del peso $r_j(A)$ es entonces

$$r_j^I(A) = cf(A \Rightarrow C_j). \quad (4)$$

Otros autores (ver por ejemplo [7]) prefieren compensar de algún modo este valor mediante el grado de confianza obtenido por las otras clases $C_l, l \neq j$, y definen los pesos como

$$r_j^{II}(A) = cf(A \Rightarrow C_j) - \frac{1}{N_C - 1} \sum_{\substack{l=1 \\ C_j \neq C_l}}^{N_C} cf(A \Rightarrow C_l), \quad (5)$$

o también

$$r_j^{III}(A) = cf(A \Rightarrow C_j) - \sum_{\substack{l=1 \\ C_j \neq C_l}}^{N_C} cf(A \Rightarrow C_l). \quad (6)$$

2.2. EXCEPCIONES MENORES Y SIGNIFICATIVAS

Obsérvese que una subregla $R_j : A \Rightarrow C_j$ produce una 3-partición del universo $U = U_X \times \zeta$ dado por [5]:

- el conjunto de ejemplos de la regla: $A \wedge C_j$,
- su conjunto de contraejemplos: $A \wedge C_j^c$ y
- el conjunto de instancias irrelevantes: A^c.

Por tanto, una regla de este tipo puede ser evaluada por medio de las proporciones de ejemplos positivos $P(A \cap C_j)$ y negativos $P(A \cap C_j^c)$. Esta concepción de regla es equivalente a la noción habitual en minería de datos y aprendizaje automático, en la que la validez de una regla es evaluada en términos de dos cantidades, el *soporte* $P(A)$ y la *confianza* $P(C_j \mid A)$, puesto que

$$P(A) = P(A \cap C_j) + P(A \cap C_j^c), \quad (7)$$

$$P(C_j \mid A) = \frac{P(A \cap C_j)}{P(A \cap C_j) + P(A \cap C_j^c)}. \quad (8)$$

Por tanto, esta concepción de una regla $R_j : A \Rightarrow C_j$, implícitamente basada en la 3-partición anterior, conduce a un marco de evaluación respecto al espacio de datos *U* en el espíritu de la bipolaridad de tipo 2 (ver [5]), i.e., un marco en el que una regla es evaluada mediante dos valores independientes (aunque relacionados), uno exhibiendo un carácter positivo (o a favor de la regla) mientras que el segundo muestra un carácter negativo (o en contra de la regla). Además es posible la existencia de información neutral o irrelevante, como se requiere a menudo cuando reglas con diferentes premisas han de ser comparadas. Así, es posible medir qué regla tiene una mayor proporción de información localmente favorable (luego mayor confianza) así como mayor relevancia global (i.e., mayor soporte).

Sin embargo, nótese que cuando se comparan subreglas $R_j : A \Rightarrow C_j$, $j = 1,...,N_C$, i.e. reglas con la misma premisa pero distinto consecuente, (lo cual es bastante habitual en el contexto de la clasificación basada en reglas), esta concepción de regla degenera en un marco de bipolaridad tipo 1 (de nuevo [5]). Esto es así porque en este caso el universo de interés es solo la celda $A \times \zeta$ y los patrones contenidos en ella y, en esta situación, la información positiva $P(A \cap C_j)$ es el complementario con respecto a $P(A)$ de la información negativa $P(A \cap C_j^c)$, pues $P(A \cap C_j) = P(A) - P(A \cap C_j^c)$ o, equivalentemente,

$$P(C_j \mid A) = \frac{P(A \cap C_j)}{P(A)} = 1 - \frac{P(A \cap C_j^{\,c})}{P(A)} = 1 - P(C_j^{\,c} \mid A) \quad (9)$$

Así pues, en este marco las evaluaciones positiva y negativa no son independientes, y toda la información contenida en la celda $A \times \zeta$ es juzgada como positiva o negativa, sin dejar lugar para información neutral, i.e. todo lo que no es un ejemplo es considerado como un contraejemplo [8]. Por ello, se asume implícitamente que cada clase C_j es igualmente opuesta a las demás.

Sin embargo, esta asunción puede ser demasiado restrictiva en algunos contextos de aplicación, en los que la relación de oposición entre las clases podría ser gradual y/o variable: por ejemplo, al valorar las consecuencias potenciales de un desastre a partir de información histórica (ver por ejemplo [9,11]), la cantidad de víctimas mortales podría ser evaluada en términos de las clases o etiquetas {*sin víctimas*, *muy pocas*, *pocas*, *bastantes*, *muchas*}. Dada una cierta premisa A y una regla $R: A \Rightarrow$ *sin víctimas*, un ejemplo de la clase *muchas víctimas* verificando A constituye una excepción mayor y más significativa a R que otra instancia en A con *muy pocas víctimas*.

Otro ejemplo: en el contexto del análisis de cestas de la compra [1], si se descubre una asociación entre compra de *cereales* (premisa A) y compra de *películas para adultos* (consecuente C_j), podría ser interesante contar con un marco de representación del conocimiento capaz de advertir que entre los restantes compradores de cereales hay un grupo importante que adquiere *artículos religiosos* (excepción significativa desde un punto de vista moral), distinguiéndolos de aquellos que compran *fruta* en su lugar (excepción menor o no significativa).

Estos ejemplos sugieren que puede ser importante distinguir la situación en que los contraejemplos de una regla $R_j: A \Rightarrow C_j$ presentan una cantidad relevante de *excepciones significativas* de aquella en que solo ocurren *excepciones menores*. La confianza de la regla debería ser menor en la primera situación que en la segunda. Sin embargo, esto no es posible en el marco habitual basado en confianza y soporte, que como se ha visto solo es capaz de distinguir entre ejemplos y contraejemplos.

En este sentido, la distinción de estos últimos entre excepciones menores y significativas implica la adopción de un marco de evaluación bipolar de tipo 2 en el interior de cada celda A del espacio de datos, en el que la información positiva, como de costumbre, viene dada por los ejemplos, la información negativa es ahora asociada a las excepciones significativas y se permite la aparición de información de carácter neutral (las excepciones menores).

Por otro lado, en tanto que la información negativa deja así de ser el complementario de la positiva, es entonces necesario dar un procedimiento que permita obtener ambas de manera separada.

En este trabajo, la información negativa será definida de manera precisa y práctica a través del uso de matrices de disimilaridad $\Delta = (d_{ij})_{N_C \times N_C}$ (ver [10]), tales que el valor $d_{ij} \in [0,1]$ expresa el grado en la clase C_j es opuesta, antagónica o disimilar a la clase $C_i, i, j = 1,...,N_C$. Se asume en principio que $d_{ii} = 0$ para todo i, de manera que una clase no puede ser opuesta a sí misma. No obstante, se admite que Δ pueda ser no simétrica, de algún modo permitiendo que la noción de oposición o disimilaridad subyacente sea asimétrica.

Así, por ejemplo, la matriz $\Delta^I = 0$ representa una situación en que ninguna clase es opuesta a las demás. En el otro extremo, la matriz $\Delta^{III} = 1 - Id$ describe una situación en que cada clase es totalmente opuesta a las demás. Por tanto, entre esos extremos, existe un amplio rango de posibilidades para modelar condiciones de oposición o disimilaridad específicas. La elección de una matriz de disimilaridad concreta dependerá, por supuesto, de los requisitos y restricciones característicos de cada contexto de aplicación.

Antes de proceder a exponer el modelo de clasificación bipolar, se introducirá alguna notación útil para conectar las matrices de disimilaridad con la especificación de los pesos $r_j(A)$ de las reglas.

Sea $\Delta = (d_{ij})_{N_C \times N_C}$ una matriz de disimilaridad y $\zeta = \{C_1,...,C_{N_C}\}$ el conjunto de clases en consideración. Dada una clase $C_i \in \zeta$, se llamará función de disimilaridad de la clase C_i a la función $\mu_{\Delta C_i}: \zeta \to [0,1]$ dada por $\mu_{\Delta C_i}(C_j) = d_{ij}$. Nótese que $\mu_{\Delta C_i}$ puede ser interpretada como la función de pertenencia del predicado borroso $\Delta C_i =$ "ser opuesta a C_i" definido sobre el conjunto ζ.

3 APRENDIZAJE BIPOLAR

Sobre la base de las consideraciones anteriores acerca de especificación de pesos de reglas y dada una matriz de disimilaridad $\Delta = (d_{ij})_{N_C \times N_C}$ y las correspondientes funciones de disimilaridad $\mu_{\Delta C_i}, i = 1,...,N_C$, se propone evaluar grados de confianza (o pesos) positivos y negativos $r_j^+(A)$ y $r_j^-(A)$, $j = 1,...,N_C$, respectivamente a favor y en contra de una subregla $R_j: A \Rightarrow C_j$, por medio de las siguientes expresiones:

$$r_j^+(A) = \frac{\sum_{p=1/C^p=C_j}^{m} \mu_A(x^p)}{\sum_{p=1}^{m} \mu_A(x^p)} = cf(A \Rightarrow C_j) \qquad (10)$$

$$r_j^-(A) = \frac{\sum_{p=1}^{m} \mu_A(x^p) \cdot \mu_{\Delta C_j}(C^p)}{\sum_{p=1}^{m} \mu_A(x^p)} = cf(A \Rightarrow \Delta C_j) \qquad (11)$$

donde $(x^p; C^p) = (x_1^p, ..., x_n^p; C^p)$ es el *p*-ésimo patrón de aprendizaje. Nótese que $r^+, r^- \in [0,1]$ y que se cumple que $r^+ + r^- \leq 1$ (véase [3,5]).

Claramente, $r_j^+(A)$ tiene la misma semántica y recoge la misma información que el grado de confianza $r_j^I(A) = cf(A \Rightarrow C_j)$ usado habitualmente en la literatura. Lo mismo se aplica a la evaluación negativa $r_j^-(A) = cf(A \Rightarrow \Delta C_j)$ con respecto al predicado antagónico ΔC_j. De cualquier modo, el valor $r_j^-(A)$ será interpretado como la proporción de información negativa o conflictiva sobre la clase C_j que se acumula en la celda $A \times \zeta$. Estas consideraciones permiten considerar el par $(r_j^+(A), r_j^-(A))$ como un grado de confianza (o peso) *bipolar*.

Dada una regla $R_j : A \Rightarrow C_j$, el comportamiento de C^p en las instancias relevantes (i.e., aquellas compatibles con *A*) con respecto a C_j y ΔC_j produce diferentes escenarios significativos para la evaluación de la regla, que se traducen en diferentes valores del par $(r_j^+(A), r_j^-(A))$. En particular, es por ejemplo posible distinguir una situación de aceptación relativamente fuerte de la regla, dada por ejemplo por el par $(r_j^+(A), r_j^-(A)) = (\frac{1}{2}, 0)$, de otra de fuerte conflicto, como la representada por el par $(r_j^+(A), r_j^-(A)) = (\frac{1}{2}, \frac{1}{2})$. Nótese que estos casos son indistinguibles en el marco habitual de evaluación, en el que ambos producirían la evaluación $cf(A \Rightarrow C_j) = \frac{1}{2}$, a pesar de que podría ser relevante distinguir las situaciones en que la mitad restante de contraejemplos son excepciones menores de aquella en que se comportan como excepciones significativas, como se ejemplificó antes en los contextos de gestión de desastres y del análisis de asociación en consumo.

Por tanto, es claro que la distinción de los contraejemplos en excepciones menores y significativas, llevada a cabo mediante la consideración de una matriz de disimilaridad Δ, permite un marco de evaluación de reglas más expresivo, en el que es posible distinguir situaciones relevantes que de otro modo serían modelizadas como idénticas. En particular, el marco de evaluación bipolar propuesto constituye una generalización del marco estándar basado en soporte-confianza.

Por otro lado, el marco estándar de razonamiento borroso [4] considera los pesos como grados unidimensionales, por lo que la evaluación bipolar (r_j^+, r_j^-) no puede ser asignada como peso de una regla al ser bidimensional. Así pues, es entonces necesario desarrollar métodos de razonamiento capaces de manejar grados de confianza bidimensionales o, alternativamente, producir una agregación unidimensional de la evaluación bipolar (r_j^+, r_j^-) que pueda ser asignada como peso de una regla. Nótese que la primera opción, que será estudiada en la próxima sección, es más general que la segunda e implica que el proceso de aprendizaje finaliza en la obtención de los pares (r_j^+, r_j^-).

Sin embargo, es interesante comprobar cómo la segunda opción, que continua el proceso de aprendizaje para adaptarse a los métodos de razonamiento habituales, permite generalizar de manera sencilla y significativa buena parte de los pesos o grados de confianza habituales mediante la consideración del grado de *veracidad*

$$t_j(A) = \max\left\{r_j^+(A) - r_j^-(A), 0\right\} \qquad (12)$$

como peso (unidimensional) de la regla $R_j : A \Rightarrow C_j$. Nótese que $t_j(A)$ está definido mediante la t-norma de Lukasiewicz $T_W(x,y) = \max\{x + y - 1, 0\}$ de la evidencia positiva ($r_j^+(A)$) y la no-negativa $(1 - r_j^-(A))$, pues

$$t_j(A) = T_W(r_j^+(A), 1 - r_j^-(A)) \qquad (13)$$

Así pues, la veracidad $t_j(A)$ refleja la conjunción de la evidencia favorable y no-desfavorable, midiendo el grado en que la información positiva excede a la negativa o, en otras palabras, *cuánto mayor* es la proporción de ejemplos que la de excepciones significativas.

El grado de veracidad permite introducir, de una manera compensativa, la información negativa dada por la matriz Δ en los pesos de las reglas, pues se cumple que

$$t_j(A) = \max\left\{cf(A \Rightarrow C_j) - \sum_{l \neq j} d_{jl} cf(A \Rightarrow C_l), 0\right\}. \qquad (14)$$

Esto es, la confianza o evidencia positiva para la regla $R_j : A \Rightarrow C_j$ es reducida por la presencia de evidencia positiva relevante para aquellas clases C_i tal que $d_{ji} > 0$, i.e. para las clases disimilares a C_j. Si la evidencia contra una regla excede la favorable, entonces el peso de la regla es igual a cero y la regla es por tanto eliminada.

Atendiendo a (14) es ahora sencillo observar que los pesos definidos en (4) - (6) son casos particulares de este grado de veracidad bajo diferentes hipótesis de disimilaridad:

- Particularmente, tomando $\Delta = \Delta^I \equiv 0$, se sigue que $r_j^-(A) = 0$, por lo que entonces $t_j(A) = r_j^I(A)$.
- Si $\Delta = \Delta^{III} = 1 - Id$, entonces es claro que $r_j^-(A) = \sum_{l \neq j} cf(A \Rightarrow C_l)$, luego se tiene $t_j(A) = r_j^{III}(A)$.
- También, si $\Delta = \Delta^{II} = \frac{1}{N_C - 1}\Delta^{III}$, entonces se cumple que $r_j^-(A) = \frac{1}{N_C - 1}\sum_{l \neq j} cf(A \Rightarrow C_l)$, y por tanto $t_j(A) = r_j^{II}$.

Como los pesos $r_j^I, ..., r_j^{III}$ negativos se suelen truncar por 0, es claro que el grado de veracidad t_j los reproduce correctamente bajo las diferentes condiciones de oposición entre clases dadas por cada matriz Δ.

4 RAZONAMIENTO BIPOLAR

De manera más general, es posible extender el marco de razonamiento borroso propuesto en [4] para dar cabida a métodos de razonamiento basados en reglas evaluadas de manera bipolar, y, en particular, a los métodos que, como el expuesto en la sección anterior, se engloben dentro de tal marco estándar mediante el uso de una agregación unidimensional de la evaluación bipolar (r_j^+, r_j^-). Siguiendo [4], se propone la siguiente formulación del marco de razonamiento bipolar:

Formulación general de los métodos bipolares de razonamiento borroso para clasificación

1. **Grado de emparejamiento**: Se calcula, usando un operador de conjunción T_1, el grado de emparejamiento de la instancia *x* con el antecedente o premisa de todas las reglas R^q:
$$\mu_{A^q}(x) = T_1(\mu_{A_1^q}(x_1), ..., \mu_{A_n^1}(x_n)),\ q = 1, ..., N_R.$$
2. **Grado de asociación bipolar:** Se computan los grados de asociación positiva y negativa, b_j^{q+} y b_j^{q-}, del ejemplo *x* con cada una de las clases C_j según la información bipolar $(r_j^+(A), r_j^-(A))$ de cada regla R^q, para todo $q = 1, ..., N_R$,
$(b_j^{q+}, b_j^{q-}) = T_2(\mu_{A^q}(x), (r_j^+(A^q), r_j^-(A^q))),\ j = 1, ..., N_C$
con T_2 un operador de agregación de información escalar y bipolar.
3. **Ponderación:** Se ponderan los valores obtenidos, usando una función $g: [0,1]^2 \to [0,1]^2$ que, normalmente, incrementará los valores altos y penalizará los bajos, esto es, $(w_j^{q+}, w_j^{q-}) = g(b_j^{q+}, b_j^{q-})$.
4. **Grado de certeza bipolar de la clasificación del patrón en cada una de las clases:** se utiliza una función de agregación de información bipolar *f* que combine, para cada clase C_j, los grados de asociación ponderados, positivos y negativos, y produzca un grado de certeza bipolar en la clasificación de la instancia *x* en cada clase,
$(gc_j^+(x), gc_j^-(x)) = f((w_j^{q+}, w_j^{q-}), q = 1, ..., N_R),\ j = 1, ..., N_C$.
5. **Clasificación:** Se aplica un procedimiento de *defuzzification* que transforme los grados de certeza borrosos obtenidos en el paso anterior en un conjunto nítido formado por una o varias clases, que constituye la salida del clasificador. En otras palabras, este paso consiste en aplicar un operador de decisión $F: ([0,1]\times[0,1])^{N_C} \to \wp(C(Y))$ que asigne una o varias clases C_j a la instancia *x* en función de los grados de certeza bipolares $(gc_j^+(x), gc_j^-(x))$ de cada clase.

Para obtener clasificadores concretos, una opción sencilla consiste en usar $T_1 = T_2 = prod$ en los pasos 1 y 2, obviar el paso 3 y aplicar los métodos de adición normalizada (AN) y de la regla ganadora (RG) en el paso 4 para obtener diferentes grados de certeza, bien en función de la veracidad $t_j(A^q)$ de las reglas, esto es

$$t_j(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) t_j(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \quad \text{y} \quad t_j^{RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) \cdot t_j(A^q) \right\},$$

bien directamente a partir de la evaluación bipolar $(r_j^+(A^q), r_j^-(A^q))$ de cada regla R^q,

$$\pi_j^+(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) r_j^+(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \text{ y } \pi_j^-(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) r_j^-(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)},$$

$$\pi_j^{+RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) r_j^+(A^q) \right\} \text{ y } \pi_j^{-RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) r_j^-(A^q) \right\},$$

o bien a través de sus grados de *falsedad* $f_j(A^q)$,

$$f_j(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) f_j(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \text{ y } f_j^{RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) \cdot f_j(A^q) \right\},$$

donde $f_j(A^q) = \max\left\{ r_j^-(A^q) - r_j^+(A^q), 0 \right\}$. Estas diversas expresiones pueden combinarse como se muestra en la Tabla 1, dando lugar a clasificadores con diferentes grados de certeza bipolares $(gc_j^+(x), gc_j^-(x))$.

Si es preciso dar predicciones nítidas, a partir de estos grados de certeza bipolares es posible aplicar diversos procesos de *defuzzification* que valoren de manera diferente la presencia o no de evidencia negativa. Por ejemplo, es posible asignar el objeto *x* a la clase con mayor certeza positiva, o con menor certeza negativa, o

desarrollar estrategias compensativas similares a la del grado de veracidad, definiendo el grado de veracidad final

$$vf_j(x) = \max\left\{gc_j^+(x) - gc_j^-(x), 0\right\},\ j = 1,...,N_C$$

asignándose *x* a la clase C_h tal que $vf_h(x) = \max_{j=1,...,N_C} vf_j(x)$.

Otra posibilidad es usar la certeza negativa como un *veto*, de manera que *x* se asigne a la clase con mayor certeza positiva de entre aquellas con certeza negativa menor que un umbral dado.

Tabla 1: Ejemplos de clasificadores bipolares en función de sus grados de certeza positivos y negativos.

Denominación	$gc_j^+(x)$	$gc_j^-(x)$
Veracidad aditiva unidimensional (VA1)	$t_j(x)$	0
Máxima veracidad unidimensional (MV1)	$t_j^{RG}(x)$	0
Veracidad aditiva bidimensional (VA2)	$t_j(x)$	$f_j(x)$
Máxima veracidad bidimensional (MV2)	$t_j^{RG}(x)$	$f_j^{RG}(x)$
Posibilidad aditiva bidimensional (PA2)	$\pi_j^+(x)$	$\pi_j^-(x)$
Máxima posibilidad bidimensional (MP2)	$\pi_j^{+RG}(x)$	$\pi_j^{-RG}(x)$

5 CONCLUSIONES

La separación de los contraejemplos a una regla borrosa de clasificación en excepciones menores y significativas, llevada a cabo a través de la consideración de una estructura de disimilaridad sobre el conjunto de clases, permite la obtención de un marco bipolar para la evaluación de reglas que generaliza el marco habitual y aumenta el poder expresivo y de generalización de los SCBRB. En particular, la utilidad de la metodología y de algunos de los clasificadores bipolares propuestos aquí ha sido validada en [9] en el contexto de la gestión de desastres, mejorando los resultados obtenidos en [11] mediante un SCBRB no bipolar..

El trabajo futuro en relación a esta propuesta consistirá, entre otras cosas, en el desarrollo de modelos de aprendizaje adaptativo sobre la matriz de disimilaridad Δ, y en el testeo de esos modelos sobre datasets de referencia, a ser posible con conjuntos de clases estructurados. También se estudiará la aplicación de un enfoque similar en otras metodologías de aprendizaje automático, como el aprendizaje basado en datos o la minería de reglas de asociación.

Agradecimientos

Este trabajo ha sido realizado en el marco del proyecto TIN2009-07901.

Referencias

[1] Agrawal, R., Imielinski, T., Swami, A. (1993) Mining Association Rules Between Sets of Items in Large Databases, SIGMOD Conf. 207-216

[2] Atanassov, KT (1999) Intuitionistic fuzzy sets: theory and applications, Physica-Verlag, Heidelberg

[3] van den Berg, J., Kaymak, U. van den Bergh, W. -M. (2002) Fuzzy classification using probability based rule weighting, Proc. of 11th IEEE Int. Conf. on Fuzzy Systems, pp. 991-996, Honolulu, May 2002.

[4] Cordón, O., del Jesús, M. J., Herrera, F. (1999) A proposal on reasoning methods in fuzzy rule-based classification systems, International Journal of Approximate Reasoning, 20 (1) 21-45

[5] Dubois, D., Prade, H. (2008) An introduction to bipolar representations of information and preference, Int. J. of Intell. Syst., 23 (8) 866-877)

[6] Ishibuchi, H., Nakashima, T. (2001) Effect of rule weights in fuzzy rule-based classification systems, IEEE Trans. on Fuzzy Syst., 9 (4) 506-515

[7] Ishibuchi, H., Yamamoto, T. (2005) Rule weight specification in fuzzy rule-based classification systems, IEEE Trans. on Fuzzy Systems, 13 (4) 428-435

[8] Montero, J., Gomez, D., Bustince, H. (2007) On the relevance of some families of fuzzy Sets, Fuzzy Sets and Systems, 158 (22) 2429-2442

[9] Rodríguez, JT (2010) Clasificación de desastres y emergencias con representación bipolar del conocimiento, Tesis Doctoral, Universidad Complutense de Madrid

[10] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) An axiomatic approach to the notion of semantic antagonism, Procs IFSA-AFSS'11 FT104-1/6

[11] Rodríguez, J. T., Vitoriano, B., Montero, J. (2009) A general methodology for data-based rule building and its application to natural disaster management, Computers & Operations Research, 39 (4) 863-873

[12] Zadeh, L. A. (1975) Concept of a Linguistic Variable and Its Application to Approximate Reasoning 1, Information Sciences, 8 (3) 199-249

Una propuesta para la mejora de los resultados de predicción en modelos basados en Prototipos Deformables Borrosos

Ma. del Rosario Vázquez[1], Francisco P. Romero[2], José A. Olivas[2]

[1] *Universidad Autónoma del Carmen*, {mvazquez}@pampano.unacar.mx

[2] *Universidad de Castilla-La Mancha*, {franciscopascual.romero, joseangel.olivas}@uclm.es

Resumen

En el área de la Inteligencia Artificial, específicamente en aplicaciones de Sistemas Basados en Conocimiento existe el inconveniente de la categorización exclusiva de cada nueva situación al patrón más similar previamente definido. Sin embargo, en el mundo real una situación puede coincidir con las características de varios patrones y no solo con uno. Mediante la utilización del modelo Prototipos Deformables Borrosos como herramienta de predicción se interpreta la situación a evaluar deformando los prototipos para describirla. El objetivo de este trabajo es la definición de dos técnicas que mejoran los resultados de predicción obtenidos tras la utilización del modelo Prototipos Deformables Borrosos. Para probar la efectividad de las técnicas propuestas, se ha aplicado a un caso de estudio cuyos datos se han extraído del concurso KDD Cup 2010. Se obtuvieron mejores resultados con las técnicas propuestas que con la técnica del modelo de predicción.

Palabras Clave: Descubrimiento de Conocimiento Prototípico Borroso, Prototipos, Prototipos borrosos.

1 INTRODUCCIÓN

En la actualidad es habitual que a partir de grandes cantidades de datos se requiera extraer información útil. Un análisis manual de estas enormes colecciones de datos necesitará una gran cantidad de tiempo y esfuerzo; razón por la cual han surgido varias técnicas, métodos y metodologías que apoyan el proceso de extracción de conocimiento a partir de grandes volúmenes de datos. La metodología más ampliamente utilizada es el Descubrimiento de conocimiento en Bases de Datos (KDD) [6] que se define como el proceso no trivial de identificar patrones validos, nuevos, potencialmente útiles y comprensibles a partir de los datos. En este trabajo se considera un patrón como un subconjunto de datos, el cual tiene asociado un modelo. A lo largo del trabajo se referirá al patrón como prototipo.

Debido al inconveniente del KDD para manejar la imprecisión de los datos, en [3,4] propusieron una extensión del KDD llamada Descubrimiento de Conocimiento Prototípico Borroso (DCPB). El DCPB añade conocimiento en cada una de las fases del KDD, con el fin de obtener prototipos conceptuales de acuerdo a la teoría de prototipos de Zadeh [7].

Por otra parte, en las aplicaciones de sistemas basados en conocimiento simulan el conocimiento experto para interpretar situaciones, de tal manera que cuando una nueva situación se presenta, la ajustan al prototipo más similar. Sin embargo en el mundo real, una situación puede coincidir con las características de varios prototipos y no solo con uno.

Para dar solución al problema anterior, en [5] se propuso un modelo de predicción que hace uso del proceso DCPB y del modelo Prototipos Deformables Borrosos, en el cual se automatiza el conocimiento del experto con el fin de facilitar la interpretación de nuevas situaciones; esta interpretación se representa como un proceso de modificación (deformación) de prototipos a la situación real, como el experto lo hace. Dicha deformación servirá de base para el proceso de predicción.

En este trabajo se pretende ponderar los resultados de predicción que se obtienen al utilizar la técnica original del modelo propuesto en [5]. Con dicha técnica los prototipos con mayor grado de afinidad prevalecen sobre los de menor, afectando los resultados de predicción. Se proponen modificaciones a dicha técnica obteniéndose tras su aplicación mejores resultados de predicción. Para probar la efectividad de las modificaciones propuestas, se han aplicado a un caso de estudio cuya colección de datos pertenece al concurso KDD Cup 2010.

El resto del trabajo se encuentra organizado de la siguiente manera, en la sección 2 se describe el modelo de predicción en el cual se basa este trabajo. En la Sección 3 se describen las técnicas propuestas. En la sección 4 se presenta un caso de estudio. Finalmente en la sección 5 se presentan las conclusiones.

2 PRELIMINARES

En esta sección se describe de manera breve el proceso *Descubrimiento de Conocimiento Prototípico Borroso* y el modelo *Prototipos Deformables Borrosos*, los cuales son la base del modelo de predicción utilizado en este trabajo. Finalmente, se describe dicho modelo de predicción.

2.1. DESCUBRIMIENTO DE CONOCIMIENTO PROTOTIPICO BORROSO

El proceso de Descubrimiento de Conocimiento Prototípico Borroso (DCPB) es una extensión del KDD, en el cual se obtienen prototipos de datos de acuerdo a la teoría de prototipos de Zadeh [7]. Estos prototipos se representan a través de tablas (marcos), los cuales permitirán realizar predicciones.

Las etapas del DCPB son las mismas que las del KDD en las cuales interviene el conocimiento experto, se hace uso de éste para definir relaciones que existen entre ciertos atributos, dichas relaciones son definidas en la etapa de transformación y contribuyen a simplificar la etapa de minería de datos a través del uso de un solo algoritmo. En la Figura 1 se muestran las fases del proceso DCPB.

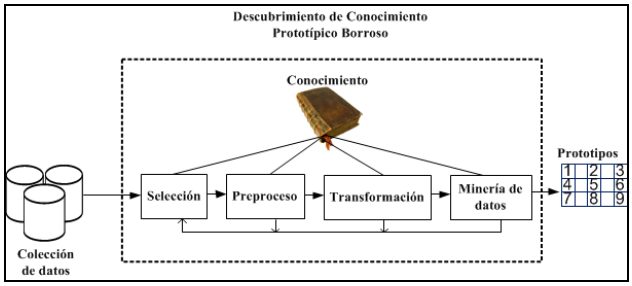


Figura 1: Fases del proceso DCPB.

2.2. PROTOTIPOS DEFORMABLES BORROSOS

En el modelo Prototipos Deformables Borrosos (PDB), el conocimiento experto en el dominio proporciona ejemplos relevantes en el dominio que se ajustan a un prototipo dado. Estos ejemplos están compuestos por una serie de características (atributos).

Los prototipos generados a partir de los ejemplos se representan formalmente mediante la utilización de números borrosos, los cuales permiten obtener el grado de afinidad con la nueva situación. Por simplicidad, comúnmente se utiliza la función triangular para representar de manera grafica a los prototipos, para ello se calculan los siguientes factores:

1. El centro de cada prototipo, el cual se obtiene aplicando alguna función al conjunto de ejemplos de cada prototipo.
2. El soporte de cada prototipo, en [5] definen varias medidas para calcularlo.

2.3. MODELO DE PREDICCIÓN BASADO EN DCPB Y PDB

El modelo de predicción en el cual se basa este trabajo está compuesto por los siguientes pasos:

1. Mediante el proceso DCPB se obtiene la definición paramétrica de cada uno de los prototipos. Esta definición está representada por una serie de marcos (de acuerdo a la complejidad del problema) con sus correspondientes características, que serán fundamentales para el modelo de predicción.
2. A partir de los ejemplos definidos en el modelo PDB por el conocimiento experto, se construyen prototipos. De esta manera cada prototipo vendrá representado formalmente mediante un número borroso y definido paramétricamente por un marco (se agregan las distintos marcos que pudiera tener el dominio real).

   Construido el modelo y ante una nueva situación, se establece un método para obtener la predicción de esa situación. Esto se realizará mediante los siguientes pasos.

1. Se calcula el valor que definirá a la nueva situación a través de la agregación de sus características.
2. Se calcula la afinidad prototípica entre el valor que define a la nueva situación con cada uno de los prototipos borrosos.
3. Se modifica la definición que se obtuvo mediante DCPB en base al grado de compatibilidad con los prototipos deformables borrosos. La técnica de deformación de los prototipos utilizada por el modelo se muestra en la ecuación 1, la cual consiste en una combinación lineal cuyos coeficientes son las afinidades de los prototipos con la nueva situación, los vectores son lo marcos que se obtuvieron en el proceso DCPB.

$$Caso\,\mathrm{Re}\,al_j(w_{1j},..w_{nj}) = \sum_{i=1}^{n} \mu\rho_{ij} v_{ij} \qquad (1)$$

Donde: w_{nj} son las características de la nueva situación j, μp_{ij} es la afinidad del prototipo i con la situación j; y v_{ij} es el marco i para la situación j.

3 PROPUESTA

En este trabajo se pretende ponderar los resultados de predicción que se obtienen al utilizar la técnica original del modelo definido en [5], por lo que se proponen 2 técnicas que añaden ciertas modificaciones a la técnica original de predicción.

La Figura 2 muestra el diagrama de bloques del modelo de predicción, el cual involucra los siguientes procesos: *definición paramétrica mediante DCPB, construcción de prototipos borrosos, evaluación de afinidad prototípica* y finalmente el proceso *modificación de afinidad paramétrica*, en el cual reside la propuesta de este trabajo.

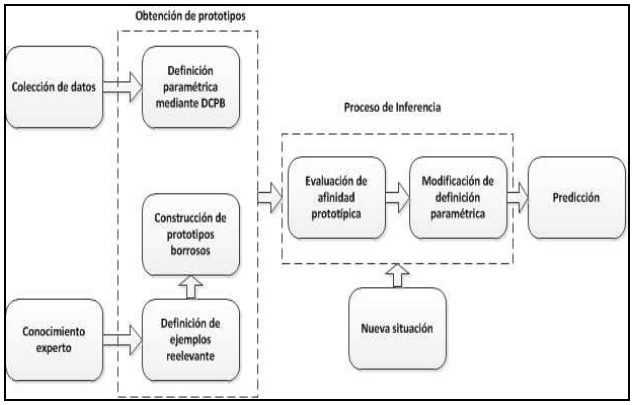


Figura 2: Diagrama de bloques del modelo de predicción

La contribución de este trabajo consiste en modificaciones realizadas a la técnica de predicción del modelo propuesto por Olivas en [5] (ver Ecuación 1). En la primera técnica propuesta definida por la Ecuación 2, se añadió un factor divisor que consiste en la agregación de las afinidades de la situación a evaluar con cada uno de los prototipos, con el fin de ponderar los resultados de la predicción.

$$Caso\,\mathrm{Re}\,al_j(w_{1j},..w_{nj}) = \frac{\left|\sum_{i=n}^{n} \mu\rho_{ij} v_{ij}\right|}{\sum_{i=1}^{n} \mu\rho_{ij}} \qquad (2)$$

Donde: w_{nj} son las características de la situación real *j*, μp_{ij} es el grado de pertenencia del prototipo *i* con la nueva situación *j*; y v_{ij} es el marco *i* de la situación *j*.

En la segunda técnica propuesta definida por la Ecuación 3, además de añadir el factor divisor de la ecuación 2, se añadió un peso a fin de resaltar el prototipo de mayor grado de afinidad con la situación a evaluar. En resumen, la segunda técnica propuesta consiste en una combinación lineal ponderada (ver Ecuación 3), en la que intervienen los grados de afinidad ($\mu\rho_{ij}$) de la nueva situación *j* con cada uno de los prototipos *i* definidos; el peso (w_{ij}) correspondiente a cada prototipo *i*, y el vector (v_{ij}) representa al marco *i* de la situación *j*.

$$Caso\,\mathrm{Re}\,al_j(w_{1j},..w_{nj}) = \frac{\left|\sum_{i=n}^{n} \mu\rho_{ij} w_{ij} v_{ij}\right|}{\sum_{i=1}^{n} \mu\rho_{ij} w_{ij}} \qquad (3)$$

Se utilizará un ejemplo con el fin de describir la forma de calcular el peso (w_{ij}) a partir de la Ecuación 4 para cada uno de los prototipos *i* afines a la nueva situación *j*. Supongamos que desde el conocimiento experto se generaron tres prototipos borrosos, los cuales se representan formalmente como números borrosos triangulares (ver Figura 3). La letra M_j representa el valor borroso que definirá a la nueva situación *j*, dicho valor se obtuvo a partir de la agregación de las características de la nueva situación, para este ejemplo el valor de M_j es de 0.6.

$$w_{ij} = f(M_j) = \begin{cases} 0 & \text{Si } M_j \le a_i \\ M_j - a_i & \text{Si } M_j \in (a_i, c_i] \\ b_i - M_j & \text{Si } M_j \in (c_i, b_i) \\ 0 & \text{Si } M_j \ge b_i \end{cases} \qquad (4)$$

Donde:

w_{ij} es el peso del prototipo *i* de la situación *j*, a_i es el límite inferior del soporte del prototipo *i*, b_i es el límite superior del soporte del prototipo *i*, c_i es el centro del prototipo *i* y M_j el valor que define a la situación *j*.

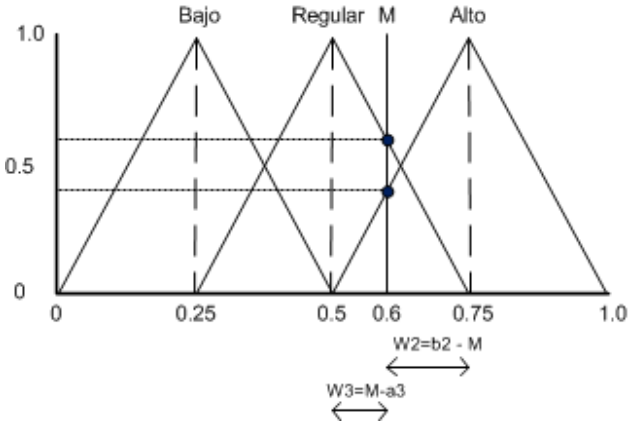


Figura 3: Prototipos borrosos.

En la Figura 3 se observa que existe afinidad entre la situación a evaluar (M_1) y los prototipos 2 (*Regular) y* 3 (*Alto)*. Para calcular los pesos correspondientes a dichos prototipos, se evalúan las condiciones definidas en la Ecuación 4. Para el peso del prototipo *regular* ($w_{2,1}$) se satisface la condición $M_j \in (c_i, b_i)$; por lo que el valor de $w_{2,1}$ se obtiene a través de la operación aritmética b_i-M_j, *e*n la cual el valor de M_1 se resta del límite superior del soporte del prototipo *regular* b_2 *(ver* Figura 3). Para el peso del prototipo *alto* ($w_{3,1}$) se satisface la condición $M_j \in (a_i, c_i]$, por lo que el valor de a_3 se resta del valor de M_1.

4 CASO DE ESTUDIO

Con el objetivo de probar la efectividad de las técnicas propuestas se ha utilizado como caso de estudio, los datos propuestos en el concurso KDD Cup 2010 [2]. Estos datos provienen de la bitácora de un sistema tutor en el aprendizaje de Álgebra llamado "The Carnegie Learning Algebra system". El conjunto de datos contiene 809694 registros que corresponden a 574 estudiantes que tomaron el curso de algebra durante el período 2005-2006.

A lo largo de la sección se describe el proceso completo del modelo de predicción propuesto en [5] aplicado a este caso de estudio, con el fin de detallar la contribución.

4.1. DEFINICIÓN PARAMÉTRICA MEDIANTE DCPB.

En este apartado se describe cada una de las fases del Descubrimiento de Conocimiento Prototípico Borroso (DCPB) aplicadas al caso de estudio.

4.1.1. Selección y preproceso

Los datos bajo estudio contienen una gran cantidad de atributos tales como: el identificador del estudiante (student_id), la jerarquía del problema (hierarchy_problem), la cual contiene la sección y la unidad a la que pertenece el problema; el número de pasos en los que se descompuso el problema (problema_steps), el número de veces que intentó resolver el problema (problema_view) entre otros. Debido a la multidimensionalidad de los datos, se hizo uso del conocimiento experto con el fin de obtener un subconjunto reducido de atributos, el cual contendrá los atributos que contribuirán de manera significativa al proceso de predicción del futuro desempeño del estudiante para resolver problemas de Álgebra. Los atributos del subconjunto fueron: *student_id*, este atributo fue considerado debido a que permitió identificar todos los registros que pertenecen a cada estudiante; *correct_first_attempt* (respuestas correctas), *time* (tiempo en resolver el problema) y steps (número de pasos utilizados).

Para subsanar el problema de valores desconocidos en los datos, se rellenaron éstos con la media de los valores de cada atributo.

Por otra parte, con el fin de obtener una base de datos reducida, se calculó el desempeño global por estudiante para resolver problemas algebraicos, para facilitar el cálculo fue de vital importancia mantener todos los registros de cada estudiante almacenados secuencialmente. Una vez que se obtuvo el desempeño global del estudiante, se desechó el atributo student_id, ya que no era relevante conocer a que estudiante pertenecía cada registro.

4.1.2. Transformación

Los datos se normalizaron en el rango de 0 a 1, para ello se utilizó la técnica definida en la Ecuación 5.

$$v' = \frac{v}{\max} \tag{5}$$

Donde: *v* es el valor del atributo y *max* es el valor máximo de ese atributo.

Posteriormente, con el fin de mantener distribuidos de manera uniforme los datos normalizados de cada atributo, se utilizó la función definida por la Ecuación 6.

$$f(x) = Ax + B \tag{6}$$

Donde:

A = (rango_máximo -rango_mínimo)/(valor_máximo – valor_mínimo)

B = rango_mínimo - A *(valor_mínimo)

rango_mínimo y *rango_máximo* se refieren al mínimo y máximo valor que tendrán los datos.

Valor_mínimo y *valor_máximo* se refieren al valor mínimo y máximo de cada atributo.

Por otra parte, para determinar los diferentes tipos de desempeño académico que pudieran tener los estudiantes, se hizo uso de un algoritmo de clustering jerárquico a través de la técnica *Repertory Grids* [1]. Se obtuvo una matriz de similitud de 540x540, cuyos datos por encima de la diagonal muestran el grado de similitud de los estudiantes para resolver problemas algebraicos con base a su nivel cognitivo, rapidez y el número de pasos requeridos. Cada dato en la matriz es un objeto compuesto que representa la distancia que existe entre los estudiantes con base a estos 3 atributos. En la Figura 4 se muestra el dendrograma generado con el algoritmo de clustering, en el cual se distinguen 3 tipos de desempeño (alto, regular y bajo) que pudieran tener los estudiantes al resolver problemas algebraicos. Con el fin de mostrar un dendrograma pequeño se agruparon los elementos hojas con mayor grado de similitud a través de letras tales como: X, M, G, etc.

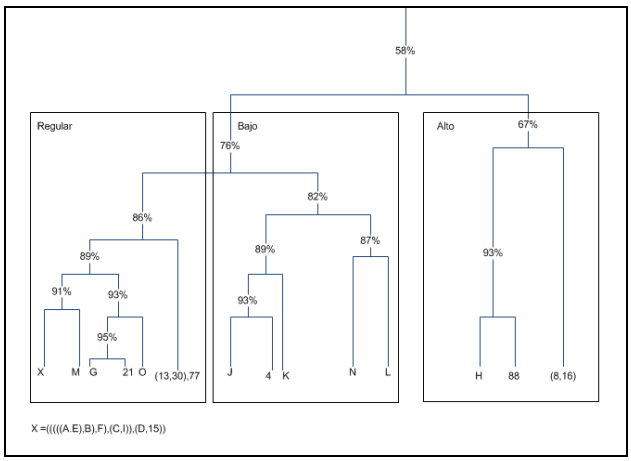


Figura 4: Dendrograma generado por el algoritmo de clustering jerárquico.

4.1.3. Minería de datos

En esta etapa se obtienen definiciones paramétricas de los prototipos a través de funciones de resumen. Se calcularon factores como la media, mínimo y máximo en cuanto a sus habilidades cognitivas (numero de respuestas correctas), rapidez (tiempo en resolver el problema) y numero de pasos utilizados (ver Tabla 1).

La Tabla 1 muestra las 3 definiciones paramétricas de los prototipos que se obtuvieron mediante el proceso DCPB. En la primera columna de la Tabla 1, la letra *C* representa las respuestas correctas, la *T* el tiempo empleado, y la *P* el número de pasos para resolver el problema. La palabra *Min* se refiere a la nota mínima obtenida, *Prom* a la nota promedio y *Max* a la nota máxima.

Como se observa en la Tabla 1, el prototipo *Alto* describe a estudiantes cuyo porcentaje de respuestas correctas al primer intento fueron del 75 a 100%; el tiempo utilizado para resolver un problema fue de 0 a 3 minutos; con un número de pasos requeridos de 1 a 17. En el prototipo *Regular*, se encuentran los estudiantes cuyas respuestas correctas fueron del 66 al 74%; el tiempo utilizado fue de 20 a 100 minutos; el numero de pasos utilizados fue de 18 a 26. Finalmente, en el prototipo *Bajo* se encuentran los estudiantes cuyas respuestas correctas fueron del 0 a 65%; en un tiempo de 4 a 19 minutos; con un numero de pasos de 29 a 100.

Tabla 1: Prototipos obtenidos a través del proceso DCPB.

	Prototipos del desempeño académico								
	Alto			**Regular**			**Bajo**		
	Min	**Prom**	**Max**	**Min**	**Prom**	**Max**	**Min**	**Prom**	**Max**
C	75	87.5	100	66	70	74	0	32.5	65
T	0	1.5	3	20	60	100	4	11.5	19
P	1	9	17	18	22	26	29	64.5	100

4.2. REPRESENTACIÓN DE LOS PROTOTIPOS

A través del conocimiento experto se generaron una serie de ejemplos, los cuales van a estar conformados por las siguientes características:

a) Complejidad del problema: vendrá dada por: {Alta/0.85, Media/0.45, Baja/0.15}.
b) Capacidad Cognitiva del estudiante: Se refiere a la capacidad cognitiva que posee el estudiante en el tema: {Alta/0.8, Media/0.5, Baja/0.2}.
c) Capacidad de comprensión del problema que se plantea: Se refiere a la capacidad que posee el estudiante de interpretar lo que se le pide y de estructurar el problema: {Alta/0.75, Media/0.45, Baja/0.15}.

Con base a la serie de ejemplos proporcionados por el conocimiento experto se crearon prototipos borrosos, los cuales se representaron formalmente a través de funciones de pertenencia triangulares y trapezoidales (ver Figura 5), para ello se ha utilizado el método explicado en la sección 2.2.

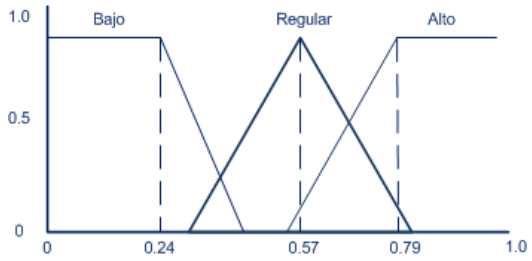


Figura 5: Representación formal de los prototipos.

4.3. OBTENCIÓN DE LOS RESULTADOS DE PREDICCIÓN

Una vez establecido el modelo, con el fin de obtener el valor que definirá a cada nuevo estudiante, se utilizó la media de los valores de las características (definidas en la sección 4.2) que identifican a ese estudiante.

La Tabla 2 muestra 10 situaciones, en la que cada una de éstas representa a un estudiante a evaluar (columna S), la columna M representa el valor que definirá al estudiante. Para el estudiante 1, el valor M es de 0.36; su grado de afinidad con el prototipo *Bajo* es de 0.5 y con el prototipo *Regular* es de 0.12.

Tabla 2: Nuevas situaciones.

S	**M**	**Grados de afinidad con los prototipos**		
		Bajo	**Regular**	**Alto**
1	0.36	0.5	0.12	0
2	0.4	0.33	0.29	0
3	0.46	0.08	0.54	0
4	0.5	0	0.7	0
5	0.6	0	0.87	0.2
6	0.67	0	0.58	0.5
7	0.7	0	0.45	0.62
8	0.73	0	0.33	0.75
9	0.76	0	0.2	0.87
10	0.9	0	0	1

Para predecir el futuro desempeño de un nuevo estudiante, se modifica la definición que se obtuvo mediante DCPB (ver Tabla1), lo cual se realiza en base al grado de compatibilidad con los prototipos deformables borrosos (ver Tabla 2).

En las Figuras 6 y 7 se observan las predicciones de los estudiantes de la Tabla 2, la cuales se obtuvieron en base a las respuestas correctas, para ello se hizo uso de las técnicas propuestas y la original del modelo.
De acuerdo con las Figuras 6 y 7, al utilizar la técnica original del modelo, con los estudiantes 1, 2, 3, 4 y 5 se obtuvieron resultados inferiores (pesimistas) a los que se obtuvieron con las técnicas propuestas. Por otra parte, con los estudiantes 6, 7, 8, 9 y 10 se obtuvieron resultados superiores (idealistas) a los de las técnicas propuestas. Los resultados que se obtuvieron con la técnica original del modelo se deben al hecho de que los prototipos con mayor grado de afinidad prevalecen sobre los de menor, afectando los resultados de predicción.

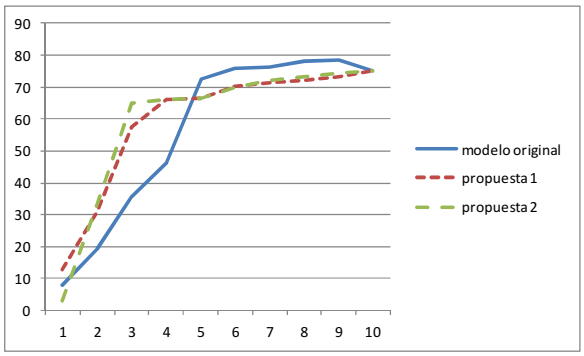


Figura 6: Resultados de predicción para la nota mínima que obtendrán los estudiantes

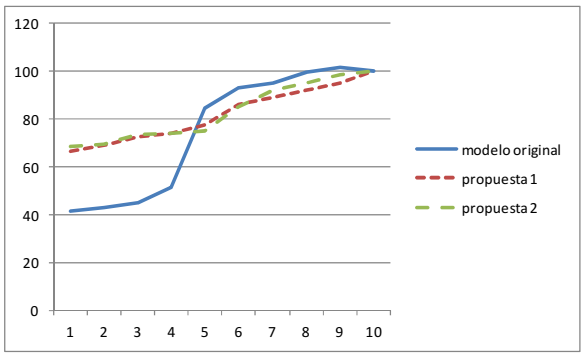


Figura 7: Resultados de predicción para la nota máxima que obtendrán los estudiantes

Por otro lado cabe señalar que cuando una nueva situación se evalúa y coincide con las características de un solo prototipo, las técnicas propuestas categorizan a ésta con el prototipo más similar sin deformar al prototipo. En esta situación, con la técnica original del modelo se obtienen mejores resultados.

Con base a los resultados de predicción para los estudiantes de la Tabla 2 (ver Figuras 6 y 7), el 70% de las predicciones obtenidas con las técnicas propuestas reflejan resultados que concuerdan con la afinidad de cada situación con los prototipos.

5 CONCLUSIONES

En este trabajo, se propusieron dos técnicas para mejorar los métodos de deformación en la inferencia sobre prototipos deformables borrosos. Las bondades de las técnicas propuestas se reflejan cuando la situación a evaluar es afín con más de un prototipo y no solo con uno. Para verificar la efectividad de las técnicas propuestas, se utilizó un caso de estudio cuya colección de datos se ha extraído del concurso KDD Cup 2010. Esta colección de datos contiene registros de estudiantes en sistemas tutores. A partir de esta colección de datos, se obtuvieron 3 definiciones paramétricas (prototipos) mediante DCPB: alto, regular y bajo; los cuales reflejan el desempeño de un estudiante al resolver problemas algebraicos. Posteriormente, se generaron 3 prototipos borrosos a partir de ejemplos proporcionados por el experto. A través de las técnicas propuestas y la original del modelo, se modificaron las definiciones que se obtuvieron mediante DCPB con base al grado de compatibilidad de la nueva situación con los prototipos borrosos. Se obtuvieron mejores resultados de predicción con las técnicas propuestas.

Como trabajo futuro se pretende utilizar diferentes colecciones de datos pertenecientes a distintas áreas tales como: medicina y fenómenos naturales, con el fin de probar el comportamiento de las técnicas propuestas.

Agradecimientos

Esta investigación está financiada por los proyectos FIDELIO (TIN2010-20395), MEC-FEDER, SCAIWEB2 (PEIC09-0196-3018) y PLINIO (POII10-0133-3516) de la Junta de Comunidades de Castilla-La Mancha y el Programa de Mejoramiento del Profesorado (PROMEP, México).

Referencias

[1] R. Bell: Analytic Issues in the Use of Repertory Grid Technique. Advances in Personal Construct Psychology 1, pp. 25-48, 1990.

[2] Educational Data Mining Challenge. 2010. https://pslcdatashop.web.cmu.edu/KDDCup/ [Consulta: 6 Septiembre 2010].

[3] J.A. Olivas, A. Sobrino: An application of Zadeh's Prototype Theory to the Prediction of Forest Fire in a Knowledge-based System. Proceedings of the 5th. International IPMU (Information Processing and Management of Uncertainty in Knowledge-based Systems), Paris, vol. II, pp. 747-752, 1994.

[4] J.A. Olivas, F.P. Romero: Fuzzy Prototypical Knowledge Discovery. Application to Forest Fire Prediction. Twelfth International Conference on Software Engineering and Knowledge Engineering (SEKE2000). Chicago, USA, 2000.

[5] J.A. Olivas: Contribución a un estudio experimental de predicción basada en Categorías Deformables Borrosas. Tesis de doctorado, Universidad de Castilla-La Mancha, España 2000.

[6] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth: The KDD Process for Extracting Useful Knowledge from Volumes of Data. pp. 27-34, 1996.

[7] L.A. Zadeh: A note on prototype set theory and fuzzy sets. Cognition 12, pp. 291-297, 1982.

Selección de Prototipos Basada en Conjuntos Rugosos Difusos

Nele Verbiest[1], **Chris Cornelis**[1,2], **Francisco Herrera**[2]

[1] *Universiteit Gent, Bélgica*, {Nele.Verbiest,Chris.Cornelis}@UGent.be
[2] *Universidad de Granada, España*, herrera@decsai.ugr.es

Abstract

En este trabajo abordamos uno de los principales problemas de k vecinos más cercanos (kNN): su sensibilidad al ruido. Llevamos a cabo Selección de Prototipos (SP), es decir, eliminamos instancias ruidosas para mejorar la calidad de la clasificación de k vecinos más cercanos. Concretamente, basándonos en un método existente de selección de instancias basada en conjuntos rugosos difusos, construimos un algoritmo de tipo envoltura que tiene en cuenta la granularidad óptima de la relación difusa de indiscernibilidad en cada conjunto de datos. Llamamos a este método Selección de Prototipos a base de Conjuntos Aproximados Difusos (SPCAD). La comparación del enfoque con el estado del arte en Selección de Prototipos confirma que nuestro método ofrece buenos resultados: supera a todos los métodos de selección de prototipos existentes con respeto a la precisión de la clasificación.

Palabras clave: conjuntos rugosos difusos, kNN, selección de prototipos

1. Introducción

El algoritmo kNN (k vecinos más cercanos ,[10]) es uno de los clasificadores no paramétricos más usados en el campo de minería de datos y de aprendizaje automático. En general, un clasificador dispone de datos de entrenamiento (un sistema de decisión X que consiste en instancias con sus valores de atributos y su clase) e intenta predecir la clase $d(t)$ de una instancia objetivo t. kNN resuelve este problema buscando las k instancias de X más similares a t y asignando t a la clase mayoritaria entre estas k instancias.

El clasificador kNN debe su popularidad al hecho de que es simple, fácil de implementar y presenta un sesgo reducido: para un problema de clasificación con dos clases, la tasa de error de 1NN nunca excederá el doble del error óptimo de Bayes.

Sin embargo, kNN tiene algunas desventajas. Uno de sus problemas es la alta dimensionalidad [4]. La selección de características hace frente a este problema seleccionando solo características relevantes y eliminando características redundantes (p.e. [18]).

En segundo lugar, kNN necesita mucho almacenamiento: el conjunto de entrenamiento completo tiene que ser almacenado para la clasificación de nuevas instancias. Además, kNN necesita mucho cómputo por que hace falta calcular las similitudes entre la instancia nueva y todas las instancias de entrenamiento. Finalmente, kNN es muy susceptible a los datos ruidosos, ya que considera todos los datos igual de relevantes, no teniendo en cuenta que el conjunto de entrenamiento pueda contener datos incorrectos.

La Selección de Prototipos (SP, [14]) hace frente a dichos problemas: reduce el conjunto de entrenamiento para seleccionar los vecinos más cercanos. Se ha realizado mucho trabajo en el campo de métodos de SP; en [15], se presenta una taxonomía extensa de métodos existentes. En primer lugar, podemos distinguir entre los métodos SP basándonos en el tipo de instancias que eliminan. Los métodos de edición quitan instancias ruidosas, mientras que los métodos de condensación principalmente eliminan instancias supérfluas, y los métodos híbridos intentan eliminar ambos tipos de instancias al mismo tiempo.

En segundo lugar, los métodos se agrupan en métodos de filtro y de envoltura. En el contexto de SP, los métodos de envoltura utilizan el conjunto entero de entrenamiento X para evaluar un subconjunto de instancias $S \subseteq X$. La calidad de S se mide clasificando X con una estrategia "leave-one-out": cada instancia x en X se clasifica buscando vecinos dentro de $S \setminus \{x\}$ si $x \in S$ y dentro de S en otro caso. Los métodos filtro solo usan datos parciales para decidir si una instancia dada debe ser retenida o eliminada. Por ejemplo, el método CNN selecciona una instancia si la regla 1NN ejecutada con el subconjunto actual de instancias la cla-

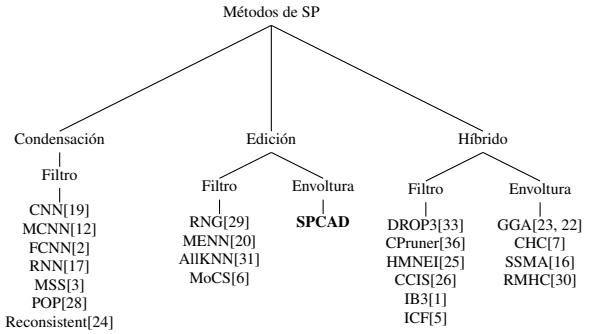


Figura 1: Taxonomía de Métodos de SP, incluido SPCAD

sifica mal. En general, los métodos de envoltura son más exactos pero también más lentos que los filtros. En la Figura 1, se representa la taxonomía de SP. Para cada grupo de métodos, se muestran los que tienen mejor desempeño.

El método que proponemos en este trabajo tiene como objetivo eliminar las instancias ruidosas y lo consigue evaluando la calidad de subconjuntos de instancias sobre el conjunto de entrenamiento completo. Por lo tanto, nuestra propuesta, Selección de Prototipos a base de Conjuntos Rugosos Difusos (SPCAD), es un método de envoltura y de edición. SPCAD utiliza los conjuntos rugosos difusos (CRD, [13]), un modelo híbrido que une la teoría de los conjuntos rugosos [27] que modelan el conocimiento imperfecto y la de los conjuntos difusos [35] que modelan datos imprecisos. La teoría CRD se ha explorado extensamente para la selección de características [9] pero su uso para SP todavía se encuentra en una etapa preliminar. En [21]) se presentó un primer método de selección de instancias basado en la teoría CRD. En este trabajo, extendemos este método y mejoramos su precisión usando un enfoque de envoltura.

En la siguiente sección, describimos SPCAD y en la Sección 3, mostramos su buen desempeño en problemas reales. Concluimos el trabajo en la sección 4.

2. Selección de Prototipos Basada en Conjuntos Rugosos Difusos

En lo sucesivo, se considera un sistema de decisión $(X, \mathscr{A} \cup \{d\})$, donde $X = \{x_1, \ldots, x_n\}$ es un conjunto de instancias, $\mathscr{A}$ es un conjunto de atributos $\{a_1, \ldots, a_m\}$, y d es el atributo de decisión (no incluido en $\mathscr{A}$). El valor de la instancia x para el atributo a se denota por $a(x)$ y puede ser continuo (real) o nominal. El atributo de decisión de una instancia x, $d(x)$, siempre es nominal. Sin perdida de generalidad, asumimos que todos los atributos continuos están normalizados: para cada $x \in X$, $a(x) \in [0,1]$.

El nucleo de la teoría CRD es el concepto de región positiva difusa. Para una instancia $x \in X$, el grado de pertenencia a la región positiva difusa está dado por:

$$POS^{\alpha}_{\mathscr{A}}(x) = \inf_{y \in X} \mathscr{I}(R^{\alpha}_{\mathscr{A}}(x,y), R_d(x,y)).$$

En esta fórmula, $\mathscr{I}$ es una implicación difusa y $R_d(x,y) = 1$ si las instancias x e y tienen la misma decisión y 0 en otro caso. $R^{\alpha}_{\mathscr{A}}(x,y)$ mide la indiscernibilidad de las instancias x e y con respeto a los atributos $\mathscr{A}$. Se define por:

$$R^{\alpha}_{\mathscr{A}}(x,y) = \underbrace{\mathscr{T}(\max(0, 1 - \alpha\delta_a(x,y)))}_{a \in \mathscr{A}},$$

donde $\mathscr{T}$ es una norma triangular (t-norma) y δ_a es una medida de distancia basada en el atributo a. Nótese que esta definición sigue siendo válida incluso si $\mathscr{A}$ contiene más de dos atributos, ya que una t-norma es asociativa y por lo tanto se puede extender a $[0,1]^m \to [0,1]$ inequívocamente. Usamos la siguiente medida de distancia en el caso de un atributo nominal a:

$$\delta_a(x,y) = \begin{cases} 0, & \text{si } a(x) = a(y) \\ 1, & \text{en otro caso} \end{cases}$$

Si a es continuo, usamos:

$$\delta_a(x,y) = (a(x) - a(y))^2.$$

Como asumimos que todos los atributos continuos están normalizados, $R^{\alpha}_{\mathscr{A}}$ producirá un valor entre 0 y 1.

El parámetro $\alpha \in [0, +\infty[$ se llama la *granularidad* y expresa lo grandes que las diferencias entre valores de atributos de instancias tienen que ser para poder distinguirse entre ellas. Cuando α es más pequeño, los valores de atributos de las instancias tendrán que diferenciarse más.

Para un α determinado, la región positiva evalua el grado en el que todas las instancias indiscernibles de una instancia x pertenecen a la misma clase. El algoritmo FRIS (Fuzzy Rough Instance Selection, [21]) emplea esta idea y elimina todas las instancias que no pertenecen completamente a la región positiva difusa ($POS_{\mathscr{A}}(x) < 1$).

FRIS usa un valor fijo para el parámetro de granularidad α. Sin embargo, los experimentos en [21] mostraban que el desempeño de FRIS depende altamente de la elección de α, y que el valor óptimo de α depende del conjunto de datos. Por tanto, SPCAD estima el valor óptimo de α para FRIS usando un enfoque de envoltura. A continuación, escribimos FRIS$^{\alpha}$ para denotar el algoritmo FRIS usando α como parámetro de granularidad.

Antes de introducir SPCAD, consideramos el siguiente teorema demostrable a partir de la teorïa CRD:

Teorema de Granularidad Mínima:
Sea $\mathscr{I}$ una implicación difusa tal que $\forall t \in [0,1]$, $\mathscr{I}(t,0) = 1 - t$ (se mantiene por ejemplo para las conocidas implicaciones de Łukasiewicz y Kleene-Dienes), y sea $x \in X$.

Entonces si $\mathscr{T}$ es el mínimo o el producto:

$$\sup\{\alpha \in [0,+\infty[| POS^{\alpha}_{\mathscr{A}}(x) < 1\} = \max_{y \notin [x]_d} \frac{1}{\max_{i=1}^{m} \delta_{a_i}(x,y)} \quad (1)$$

y si $\mathscr{T}$ es la t-norma de Łukasiewicz:

$$\sup\{\alpha \in [0,+\infty[| POS^{\alpha}_{\mathscr{A}}(x) < 1\} = \max_{y \notin [x]_d} \frac{1}{\sum_{i=1}^{m} \delta_{a_i}(x,y)}, \quad (2)$$

siempre que los denominadores sean distintos de 0.

Nótese que, en caso de que los denominadores son cero, significa que existen $x, y \in X$ tales que $\forall a \in \mathscr{A}, a(x) = a(y)$, pero $d(x) \neq d(y)$. Entonces, para cada $\alpha \in [0,+\infty[$, $POS^{\alpha}_{\mathscr{A}}(x) = POS^{\alpha}_{\mathscr{A}}(y) = 0$. En otras palabras, x e y son eliminadas por $FRIS^{\alpha}$ para cada valor de α.

Denotamos por $\alpha(x)$ el valor crítico de x y usamos la notación simbólica ∞ para instancias que siempre están eliminadas por $FRIS^{\alpha}$, independientemente del valor de α:

$$\forall x \in X : \alpha(x) = \begin{cases} \infty \text{ si } \exists y \notin [x]_d, \forall a \in \mathscr{A}\, a(x) = a(y) \\ \sup\{\alpha \in [0,+\infty[| POS^{\alpha}_{A}(x) < 1\} \text{ en otro caso} \end{cases}$$

Esto significa que si $FRIS^{\alpha}$ se ejecuta con un parámetro de granularidad α menor de $\alpha(x)$, elimina la instancia. Por otro lado, si α es mayor o igual a $\alpha(x)$, $FRIS^{\alpha}$ conserva la instancia. En el caso especial case donde $\alpha(x) = \infty$, $FRIS^{\alpha}$ siempre elimina x, independientemente de la elección de α.

Ahora podemos volver a formular el problema de encontrar un valor óptimo de α para FRIS. Asumimos que las instancias $x \in X$ están ordenadas según sus valores $\alpha(x)$:

$$\alpha(x_1) \leq \alpha(x_2) \ldots \leq \alpha(x_n),$$

donde asumimos que $\infty \leq \infty$ y $\forall r \in \mathbb{R}, r < \infty$.

El algoritmo $FRIS^{\alpha}$ elimina todas las instancias x para las que $\alpha < \alpha(x)$, es decir, existe un *índice de división* $i \in \{1,\ldots,n\}$ tal que las instancias $x_1,\ldots,x_i$ se conservan y que las instancias $x_{i+1},\ldots,x_n$ se eliminan. Esto se reduce a aplicar el algoritmo $FRIS^{\alpha(x_i)}$. Significa que encontrar un buen $\alpha \in [0,\infty[$ para $FRIS^{\alpha}$ es equivalente a encontrar un buen $\alpha \in \{\alpha(x_1),\ldots,\alpha(x_n)\}$ para $FRIS^{\alpha}$.
El algoritmo SPCAD explota esta idea. Evalua los subconjuntos $S = \{x_1,\ldots,x_i\}$ de instancias correspondientes a los algoritmos $FRIS^{\alpha(x_i)}$ $(i = 1,\ldots,n)$ calculando sus precisiones en entrenamiento usando leave-one-out. Selecciona los valores α que proporcionan un acierto de entrenamiento máximo. En caso de que distintos valores de α corresponden al mejor acierto en entrenamiento, tomamos su mediana.

En Algoritmo 1, se muestra el esquema final del algoritmo. En primer lugar, los valores $\alpha(x)$ se calculan para cada $x \in X$. Después, se eliminan los valores duplicados y se ordenan los restantes, teniendo en cuenta de nuevo que

Algorithm 1 SPCAD

1: **input**: Sistema de decisión $(X, \mathscr{A} \cup \{d\})$
2: Calcula $\alpha(x_1),\ldots,\alpha(x_n)$
3: Elimina duplicados y ordena los valores de α del paso 2: $\alpha_1 > \alpha_2 > \ldots > \alpha_p$
4: opt.alphas $\leftarrow \{\infty\}$
5: Calcula vecinos más cercanos de todas las instancias
6: acc.opt $\leftarrow$ accuracy$(X, \mathscr{A} \cup \{d\})$
7: acc.current $\leftarrow$ acc.opt
8: **for** $\alpha = \alpha_2,\ldots,\alpha_p$ **do**
9: Elimina instancias x para las que $\alpha(x) > \alpha$
10: **if** número de instancias restantes > 1 **then**
11: Recalcula vecinos más cercanos de instancias para las que el vecino más cercano actual es eliminado
12: Recalcula acc.current
13: **if** acc.current $>$ acc.opt **then**
14: opt.alphas $\leftarrow \{\alpha\}$
15: **else if** acc.current = acc.opt **then**
16: opt.alphas $\leftarrow$ opt.alphas $\cup \{\alpha\}$
17: **end if**
18: **end if**
19: **end for**
20: best.alpha = median(opt.alphas)
21: **Output** $FRIS^{best.alpha}(X, \mathscr{A} \cup \{d\})$

$\forall r \in \mathbb{R}, r < \infty$. Por sencillez, asumimos que $FRIS^{\infty}$ devuelve el conjunto de entrenamiento X entero. En la línea 5, se obtiene el vecino más cercano para cada instancia. A continuación, en la línea 6 se calcula el acierto usando el conjunto entero de entrenamiento. En cada ejecución del bucle, las instancias para las que $\alpha(x)$ excede el umbral actual α se eliminan del conjunto actual de instancias. Nótese que esto equivale a aplicar $FRIS^{\alpha}$. Si el acierto del conjunto de instancias actual es igual o mejor que el mejor acierto encontrado hasta ahora, se actualizan este último valor y la lista correspondiente de valores α optimales. Finalmente, el mejor valor de α se calcula como la mediana de todos los valores óptimos de α values, y se devuelve el subconjunto obtenido con $FRIS^{best.alpha}$. La razón por que consideramos los valores de α en orden decreciente es que así podemos implementar este paso de manera eficaz. Los subconjuntos de instancias son generados de manera decreciente. Cada vez que se eliminan instancias del subconjunto actual, solo recalculamos los vecinos más cercanos de las instancias afectadas por esta operación. El criterio de parada en la línea 10 asegura que se puede obtener un vecino más cercano para cada instancia; si solo hay una instancia x para la que $\alpha(x) = \alpha_p$, es decir, solo hay una instancia con menor valor de α, entonces en la línea 11 el vecino más cercano no se puede obtener para x: el único candidato es x pero esta instancia no se puede elegir como vecino más cercano de x en la estragía leave-one-out. Por eso, optamos por no considerar este subconjunto $\{x\}$.

Un posible inconveniente del método básico SPCAD es que las fórmulas (1) y (2) para $\alpha(x)$ están basadas en el operador máximo. Esto significa que cambios pequeños en los datos pueden alterar los cálculos de manera drástica, y por lo tanto la robustez de la selección de prototipos podría verse afectada. Por tanto, consideramos una generalización del enfoque básico usando operadores de agregación del tipo OWA [34].

Recordemos que, dada una serie de valores $a_1,\ldots,a_p \in \mathbb{R}$ y un vector de pesos $W = \langle w_1,\ldots,w_p\rangle$ tal que $\forall i \in 1,\ldots,p$: $w_i \in [0,1]$ y $\sum_{i=1}^{p} w_i = 1$, la agregación OWA de estos valores se obtiene como $[OWA_W(a_1,\ldots,a_p) = \sum_{i=1}^{p} w_i b_i$, donde $b_i = a_j$ si a_j es el i-ésimo mayor valor entre $a_1,\ldots,a_p$. Es decir, los valores se ordenan y después se extrae un promedio ponderado de estos valores.

El operador OWA se puede usar para relajar la idea del máximo, haciéndolo menos estricto. Consideremos un vector de pesos $W_{max} = \langle w_1,\ldots,w_p\rangle$ tal que $w_1 \geq w_2 \ldots \geq w_p$, donde valores mayores corresponden a pesos mayores. Se puede usar $OWA_{W_{max}}$ para calcular los $\alpha(x)$ de cada $x \in X$. Para $\mathscr{T} = \mathscr{T}_L$, la definición generalizada de $\alpha(x)$ se denota por:

$$\forall x \in X : \alpha^{OWA}(x) = \underbrace{OWA_{W_{max}} \frac{1}{\sum_{i=1}^{m} \delta_{a_i}(x,y)}}_{y \notin [x]_d},$$

mientras que para $\mathscr{T} = \mathscr{T}_M$ or $\mathscr{T} = \mathscr{T}_P$, podemos generalizar la definición de $\alpha(x)$ por:

$$\forall x \in X : \alpha^{OWA}(x) = \underbrace{OWA_{W_{max}} \underbrace{\frac{1}{OWA_{W_{max}} \delta_{a_i}(x,y)}}_{i=1,\ldots,m}}_{y \notin [x]_d}.$$

La ventaja del operador OWA_W es que todos los pesos pueden ser mayores de cero, lo que significa que todos valores pueden influir el resultado de agregación y se pueden resultados más estables. Un posible vector de pesos para una agregación OWA con el comportamiento del máximo puede obtenerse por:

$$\forall i \in 1,\ldots,p : w_i = \frac{2(p-i+1)}{p(p+1)}. \quad (3)$$

3. Resultados Experimentales

3.1. Configuración Experimental

En esta sección estudiamos el rendimiento de SPCAD. Seguimos la configuración experimental del estudio realizado en [15]. Empleamos 58 conjuntos de datos y sus particiones del repositorio KEEL[1] y comparamos SPCAD con los 21 métodos de selección de prototipos que tienen mejor desempeño entre los diferentes tipos de métodos de SP.

Usamos la estrategia de validación cruzada en 10 partes para evaluar el rendimiento de los clasificadores. Para cada parte (datos de test), reducimos las partes restantes (datos de entrenamiento) usando un método de SP determinado, y clasificamos los datos de test con kNN, empleando los datos de entrenamiento reducidos.

El rendimiento de SPCAD se compara usando 4 medidas:

- *Precisión (acc):* la tasa de instancias de test clasificadas correctamente
- *Kappa de Cohen (κ) [8]:* una medida adicional de precisión, basándose en la matriz de confusión, que compensa los éxitos obtenidos al azar, .
- *reducción de almacenamiento (red):* la fracción de instancias eliminadas del conjunto de entrenamiento
- *tiempo de ejecucíon (time):* el tiempo de ejecucíon en segundos del método SP. El tiempo de ejecución de la clasificación 1NN posterior no se considera.

Experimentos[2] muestran que SPCAD utilizando $\mathscr{T}_L$ como t-norma y el enfoque OWA con los pesos definidos como en la ecuación (3) obtienen los mejores resultados. Por lo tanto, empleamos esta configuración para los experimentos. Además, solo experimentamos con el clasificador 1NN.

3.2. Resultados

En la Tabla 1, se muestran los resultados medios para todos los métodos de SP. De dicha tabla podemos concluir que SPCAD funciona muy bien con respeto a exactitud y Cohen's kappa. El test de Wilcoxon [3] [32] confirma esta conclusión: comparando SPCAD con cada de los demás métodos SP con respeto a exactitud y kappa, SPCAD siempre los supera con un nivel de $\alpha = 0{,}1$ de significancia.

Pese a ello, usando el test de Wilcoxon para multiples comparaciones por pares, perdemos el control sobre la verdadera tasa de error, es decir la probabilidad de hacer uno o más discubrimientos falsos entro todas las hipótesis [11]. Por ello, también usamos el test de Friedman y el procedimiento post-hoc de Holm [3] [11], diseñado especificamente para comparar multiples hipótesis. Hemos aplicado este test solo para los 9 algoritmos con mejor desempeño con respeto a la precisión en test y kappa (AllKNN, CHC, GGA, HMNEI, MoCS, RMHC, RNG, SSMA, FRPS) ya que el test pierde poder si se comparan demasiados algoritmos [11].

[1] http://sci2s.ugr.es/keel/datasets.php

[2] No incluidos en este trabajo debido a restricciones de espacio. Se pueden consultar en: users.ugent.be/~nverbies

[3] Las estadísticas de este test se pueden consultar en: users.ugent.be/~nverbies

Test acc.		Kappa		Red.		Time	
SPCAD	**0.7920**	**SPCAD**	**0.5942**	CHC	0.9788	POP	0.0620
SSMA	0.7830	SSMA	0.5756	SSMA	0.9653	CNN	0.2757
RMHC	0.7777	RMHC	0.5686	MCNN	0.9373	MCNN	1.5016
RNG	0.7773	HMNEI	0.5644	GGA	0.9301	FCNN	2.3161
CHC	0.7766	CHC	0.5618	RNN	0.9284	MSS	2.5574
GGA	0.7733	RNG	0.5601	CCIS	0.9213	CCIS	4.1314
MoCS	0.7710	MoCS	0.5575	CPruner	0.9070	MoCS	5.1014
HMNEI	0.7679	GGA	0.5572	RMHC	0.9010	AllKNN	8.1177
AllKNN	0.7662	AllKNN	0.5374	DROP3	0.8452	MENN	12.1973
POP	0.7541	POP	0.5298	ICF	0.7452	**SPCAD**	**27.6379**
MENN	0.7520	MENN	0.5175	IB3	0.7197	ICF	30.4970
RNN	0.7519	MSS	0.5172	FCNN	0.6638	HMNEI	66.5790
MSS	0.7460	RNN	0.5100	CNN	0.6187	IB3	185.2876
FCNN	0.7361	FCNN	0.5073	Reconsistent	0.5886	CPruner	194.9527
IB3	0.7347	IB3	0.5041	HMNEI	0.5464	Reconsistent	534.1992
CNN	0.7336	CNN	0.5026	MSS	0.4750	RNG	616.2494
DROP3	0.7139	DROP3	0.4686	MENN	0.4507	SSMA	2084.4178
Reconsistent	0.7083	Reconsistent	0.4659	**SPCAD**	**0.3557**	CHC	2244.7731
CPruner	0.6971	MCNN	0.4407	AllKNN	0.3180	DROP3	2296.5474
CCIS	0.6820	ICF	0.4118	RNG	0.2081	RMHC	3962.0273
MCNN	0.6819	CCIS	0.4016	MoCS	0.1113	GGA	6965.3978
ICF	0.6773	CPruner	0.3894	POP	0.0767	RNN	8030.0853

Cuadro 1: Resultados medios de SPCAD y de los algoritmos del estado del arte de SP

Todos los métodos salvo RNG son significativamente peores que SPCAD con respeto a la precisión en test. Con respeto a kappa, SPCAD claramente supera a todos los demás algoritmos SP considerados.

La tasa de reducción de SPCAD es similar a la de los demás métodos de edición. El tiempo de ejecución de SPCAD es aceptable: SPCAD es mucho más rápido que los demás métodos de envoltura, pero más lento que algunos de los filtros. Esto es interesante: SPCAD es un método de envoltura, pero su naturaleza decremental permite una implementación eficaz, lo que resulta en un tiempo de ejecución bastante bajo.

4. Conclusión y Trabajo Futuro

En este artículo hemos presentado un nuevo método de selección de prototipos, SPCAD. Este método de preprocesamiento elimina instancias ruidosas e intenta mejorar el rendimiento del clasificador *k*NN. Hemos comparado SPCAD con 21 algoritmos del estado del arte de SP para 58 conjuntos de datos. Los resultados obtenidos muestran un rendimiento excelente: SPCAD supera a todos los algoritmos significativamente con respeto a precisión y kappa.

Para mejorar la tasa de reducción de SPCAD, planeamos combinarlo con métodos de condensación para desarrollar un algoritmo híbrido de selección de prototipos basado en conjuntos aproximados difusos. Además, dado que *k*NN es susceptible a los datos de alta dimensionalidad, planeamos abordar dicho problema combinando SPCAD con selección de características.

Referencias

[1] David W. Aha and Dennis Kibler. Instance-based learning algorithms. In *Machine Learning*, pages 37–66, 1991.

[2] F. Angiulli. Fast nearest neighbor condensation for large data sets classification. *IEEE Transactions on Knowledge and Data Engineering*, 19(11):1450–1464, 2007.

[3] R. Barandela, F. J. Ferri, and J. S. Sanchez. Decision boundary preserving prototype selection for nearest neighbor classification. *International Journal of Pattern Recognition and Artificial Intellgence*, 19:787–806, 2005.

[4] K. S. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft. When is nearest neighbor meaningful? In *7th International Conference of Database theory (ICDT)*, pages 217–235, 1999.

[5] Henry Brighton and Chris Mellish. Advances in instance selection for instance-based learning algorithms. *Data Mining and Knowledge Discovery*, 6:153–172, 2002.

[6] Carla E. Brodley. Recursive automatic bias selection for classifier construction. *Machine Learning*, 20:63–94, 1995.

[7] J. R. Cano, F. Herrera, and M. Lozano. Using evolutionary algorithms as instance selection for data reduction in kdd: an experimental study. *IEEE Transactions on Evolutionary Computation*, 7(6):561–575, 2003.

[8] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20:37–46, 1960.

[9] C. Cornelis, R. Jensen, G. Hurtado, and D. Slezak. Attribute selection with fuzzy decision reducts. *Information Sciences*, 180(2):209–224, 2010.

[10] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[11] Joaquín Derrac, Salvador García, Daniel Molina, and Francisco Herrera. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1(1):3–18, 2011.

[12] V. Susheela Devi and M. Narasimha Murty. An incremental prototype set building technique. *Pattern Recognition*, 35(2):505–513, 2002.

[13] D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17:191–209, 1990.

[14] Robert P. W. Duin and Pavel Paclík. Prototype selection for dissimilarity-based classifiers. *Pattern Recognition*, 39:189–208, 2006.

[15] S. García, J. Derrac, J.R. Cano, and F. Herrera. Prototype selection for nearest neighbor classification: Taxonomy and empirical study. *To appear in: IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[16] Salvador García, José Ramón Cano, and Francisco Herrera. A memetic algorithm for evolutionary prototype selection: A scaling up approach. *Pattern Recognition*, 41:2693–2709, 2008.

[17] G. Gates. The reduced nearest neighbor rule. *IEEE Transactions on Information Theory*, 18(3):431–433, 1972.

[18] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.

[19] P. E. Hart. The condensed nearest neighbor rule. *IEEE Transactions on Information Theory*, 18:515–516, 1968.

[20] Kazuo Hattori and Masahito Takahashi. A new edited k-nearest neighbor rule in the pattern classification problem. *Pattern Recognition*, 32:521–528, 2000.

[21] R. Jensen and C. Cornelis. Fuzzy-rough instance selection. In *Proceedings of the 19th International Conference on Fuzzy Systems (FUZZ-IEEE 2010)*, pages 1–7, 2010.

[22] Ludmila I. Kuncheva. Editing for the k-nearest neighbors rule by a genetic algorithm. *Pattern Recognition Letters*, 16(8):809–814, 1995.

[23] Ludmila I. Kuncheva and Lakhmi C. Jain. Nearest neighbor classifier: Simultaneous editing and feature selection. *Pattern Recognition Letters*, 20:1149–1156, 1999.

[24] M. T. Lozano, J. S. Sanchez, and F. Pla. *Using the geometrical distribution of prototypes for training set condensing*, volume 3040. 2003.

[25] Elena Marchiori. Hit miss networks with applications to instance selection. *Journal of Machine Learning Research*, 9:997–1017, 2008.

[26] Elena Marchiori. Class conditional nearest neighbor for large margin instance selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:364–370, 2010.

[27] Z. Pawlak. Rough sets. *International Journal of Computer Information Science*, 11:341–356, 1982.

[28] J.C. Riquelme, J.S. Aguilar-Ruiz, Jesus S. Aguilar-ruiz, and M. Toro. Finding representative patterns with ordered projections. *Pattern Recognition*, 36(4):1009–1018, 2003.

[29] J. S. Sánchez, F. Pla, and F. J. Ferri. Prototype selection for the nearest neighbour rule through proximity graphs. *Pattern Recognition Letters*, 18:507–513, 1997.

[30] David B. Skalak. Prototype and feature selection by sampling and random mutation hill climbing algorithms. In *Machine Learning: Proceedings of the Eleventh International Conference*, pages 293–301, 1994.

[31] I. Tomek. An experiment with the edited nearest-neighbor rule. *IEEE Transactions on Systems, Man and Cybernetics*, 6(6):448–452, 1976.

[32] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, (6):80–83, 1945.

[33] D. Randall Wilson and Tony R. Martinez. Reduction techniques for instance-based learning algorithms. *Machine Learning*, 38:257–286, 2000.

[34] Ronald R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.

[35] L.A. Zadeh. Fuzzy sets. *Information Control*, 8:338–353, 1965.

[36] Ke-Ping Zhao, Shui-Geng Zhou, Ji-Hong Guan, and Ao-Ying Zhou. C-pruner: an improved instance pruning algorithm. In *International Conference on Machine Learning and Cybernetics*, volume 1, pages 94–99, 2003.

CONSTRUCCIÓN DE RESÚMENES LINGÜÍSTICOS INFORMATIVOS SOBRE SERIES DE DATOS METEOROLÓGICOS: INFORMES CLIMÁTICOS DE TEMPERATURA

A. Ramos-Soto, **F. Díaz-Hermida**, **A. Bugarín**

Departamento de Electrónica e Computación
Centro de Investigación en Tecnoloxías da Información (CITIUS)
Universidade de Santiago de Compostela,
{alejandro.ramos, felix.diaz, alberto.bugarin.diz}@usc.es

Resumen

En este trabajo se describe un método para la construcción de resúmenes lingüísticos para datos de climatología. El método propuesto utiliza criterios generales como la intensificación o la excepción a un resumen general, particularizados para que los resúmenes obtenidos automáticamente resulten totalmente consistentes con los informes climatológicos mensuales de temperatura que publica el Servicio Gallego de Meteorología (Meteogalicia) [1]. En alguno de los casos se consigue una mejora cualitativa en la descripción frente a la realizada por los expertos, pero manteniendo en todo caso la sintaxis y el estilo utilizados por estos en los informes.

Keywords: sumarización lingüística de datos, cuantificación borrosa, informes climáticos de temperatura.

1 INTRODUCCIÓN

Hoy en día, el uso masivo de grandes cantidades de datos hace necesario disponer de métodos que nos permitan extraer una descripción adecuada de los mismos que los haga comprensibles para cualquier usuario (en particular para los no especializados). Una de las tareas de mayor interés en este contexto es la de *sumarización*, entendida como la *obtención de una descripción general, breve y precisa de un conjunto de datos* [2], normalmente de tipo numérico. Esta descripción puede emplear expresiones matemáticas (frecuentemente estadísticos), utilizar un lenguaje gráfico (cuadros de mando, gráficos de tendencias, ...) o realizarse mediante frases o expresiones. En este último caso estamos en el contexto de la *sumarización lingüística*, donde resulta clave que las expresiones que forman el sumario estén completamente adaptadas en su sintaxis y su semántica a las necesidades informativas y/o estilos de los usuarios. Existen diversas aproximaciones borrosas para la generación de resúmenes lingüísticos, que usualmente se construyen mediante proposiciones cuantificadas, para diferentes aplicaciones (descripción del flujo de ingresos en centros médicos [2] o informes de consumo doméstico [3]). En otros trabajos se permiten expresiones de mayor complejidad, que involucran relaciones entre distintos atributos ([4] en información económica) o, fuera ya del ámbito borroso, que utilizan reglas y estructuras gramaticales predefinidas formadas por varias sentencias ([6] en el ámbito de predicción meteorológica) o que contemplan expresiones con fusión o traducción de eventos ([7] en la descripción de señales fisiológicas).

En este trabajo se seleccionan algunos de los criterios generales definidos en [5] (tomados del ámbito de la lingüística y la comunicación) para su aplicación en la generación automática de los informes climáticos mensuales de temperatura publicados por Meteogalicia (Servicio Meteorológico Gallego, [1]). Los informes climáticos tienen un estilo y una sintaxis predefinidas (que combinan información técnica especializada con información de tipo mas genérico) con la que se describe de una forma simple y directa la situación general para cada mes de las temperaturas en Galicia, dando en primer lugar una información general sobre como ha sido la temperatura en dicho mes y seguidamente se añade alguna matización sobre los episodios que hayan resultado especialmente destacables de acuerdo con el criterio del experto.

2 MODELO DE CONSTRUCCIÓN DE RESÚMENES CLIMÁTICOS CON INCORPORACIÓN DE MATICES

El problema abordado en este trabajo consiste en la obtención de un método que genere resúmenes lingüísticos similares y consistentes con los que producen los expertos meteorólogos del Servicio Meteorológico de la Xunta de Galicia en sus informes climáticos mensuales. Dichos informes recogen una breve descripción de los hechos más

relevantes referidos a la temperatura, de forma general para toda Galicia, a partir de los datos de un conjunto representativo de estaciones meteorológicas ubicadas por toda la geografía gallega. La información promedio de dichas estaciones se compara con la media general de temperaturas de los últimos treinta años (temperatura de referencia denominada "media climática"). El informe destaca los episodios que resulten, a juicio del meteorólogo y de su experiencia, mas significativos. Así, algunos ejemplos pueden ser expresiones que van desde la simple "Las temperaturas fueron altas" hasta la inclusión de referencias precisas como "Las temperaturas fueron bajas, especialmente en la 2ª y 3ª semanas" o "Las temperaturas fueron normales, siendo destacables los episodios muy fríos entre el día 6 y el 10 y las temperaturas bajas de la última semana". Los siguen un estilo y una estructura claras, repetitivas y muy intuitivas que pretende transmitir una información para usuarios no especializados (público en general). Los sumarios están formados en primer lugar, por un **resumen general**, que describe globalmente todo el mes y compara la variación existente entre media de dicho mes y la media climática (por ejemplo, "Las temperaturas fueron superiores a lo habitual en esta época del año"). Con ello se proporciona una información básica del comportamiento del mes, poco precisa pero muy directamente comprensible. La valoración resultante corresponde siempre a una etiqueta lingüística perteneciente a un conjunto definido por el experto (muy frío, frío, normal, cálido, muy cálido). El resumen general suele ir acompañado de **matices**, que tratan de destacar situaciones altamente significativas e informativas para el usuario. Los matices se construyen habitualmente sobre referencias temporales en las que las etiquetas predominantes estén alejadas de la asociada al resumen general. Se dan, a la hora de considerar los matices, diversas situaciones:

- **Resumen general:** el resumen general está asociado a una etiqueta central ("las temperaturas han sido normales"). En este caso, las etiquetas a ambos extremos de la escala de variación resultarán relevantes en la misma medida, ya que se apartan por igual de la media. Un ejemplo muy común se encuentra en los meses de gran variabilidad, como en *"El mes de marzo ha sido normal en promedio, en el que la primera quincena las temperaturas han sido elevadas pero en la segunda muy bajas"*.

- **Matiz de excepción**: el resumen general está asociado a una etiqueta cercana a un extremo (frío, cálido), pero existen periodos en los que predominan etiquetas del extremo opuesto a la del resumen. El matiz informativo, en este caso, resulta ser de oposición (excepción), destacándose los episodios que contradicen en mayor o menor medida al resumen general: *"El mes de julio fue **cálido**, aunque destaca la primera semana con temperaturas **muy frías** para esta época del año"*.

- **Matiz de intensificación**: el resumen general está asociado a una etiqueta cercana a un extremo (frío, cálido), pero existen períodos en los que predominan etiquetas del mismo extremo que la del resumen. El matiz informativo, en este caso, resulta ser de refuerzo (intensificación) de la etiqueta asociada al resumen general: *"El mes de julio fue **cálido**, especialmente la última semana con temperaturas **extremadamente elevadas**"*.

Resulta esencial en este caso incorporar matices como los indicados anteriormente. Para llevar a cabo esta incorporación se han adaptado algunos de los criterios generales descritos en [5]. Por lo tanto, aunque la estrategia seguida se ha aplicado a un ámbito concreto, en realidad se trata de una aproximación general a la construcción de resúmenes, que extiende en esta línea este campo de investigación.

2.1 DESCRIPCIÓN DEL ALGORITMO DE GENERACIÓN DE RESÚMENES LINGÜÍSTICOS

Hemos optado por una aproximación borrosa, con un algoritmo que genera un resumen general y añade los matices más relevantes según una heurística basada en el criterio del experto, que busca mejorar la relevancia de la información que se presenta en el resumen final. Dicho proceso se encuentra dividido en tres módulos principales, que describimos seguidamente.

El módulo de **jerarquía temporal descendente** se encarga de obtener un resumen general para todo el mes, comparando la media del mes con la media climática, utilizando un conjunto de etiquetas nítidas definidas en base al conjunto de ejemplos disponible como intervalos numéricos precisos. La definición de dichas etiquetas se hace tomando como referencia el manual de estilo elaborado por los expertos para la realización de sus informes: [-10,-2) "muy frío', [-2, -0.5) "frío", [-0.5, 0.5) "normal", [0.5 , 2) "cálido", [2, 10) "muy cálido".

A continuación, se ejecuta recursivamente para intervalos temporales de menor longitud, pero significativos dentro del mes (quincenas y semanas), de modo que se desciende en la jerarquía temporal hasta la menor granularidad establecida para esta jerarquía por los expertos (semanal en este caso, como se aprecia en la Figura 1). Cada uno de los intervalos obtenidos son evaluados y etiquetados de tal forma que indican cómo ha sido dicho período respecto a la media climática.

Una vez que se han obtenido y evaluado estos períodos, se generan todas las combinaciones posibles constituidas por las evaluaciones de los intervalos temporales (matices). Cada una de dichas combinaciones es un candidato que debe ser considerado a la hora de acompañar al resumen mensual, por lo cual se construye un *ranking* que los clasifica según el nivel de información relevante que aportan respecto al general. Dicha información se obtiene evaluando una serie de criterios sobre las características que a continua-

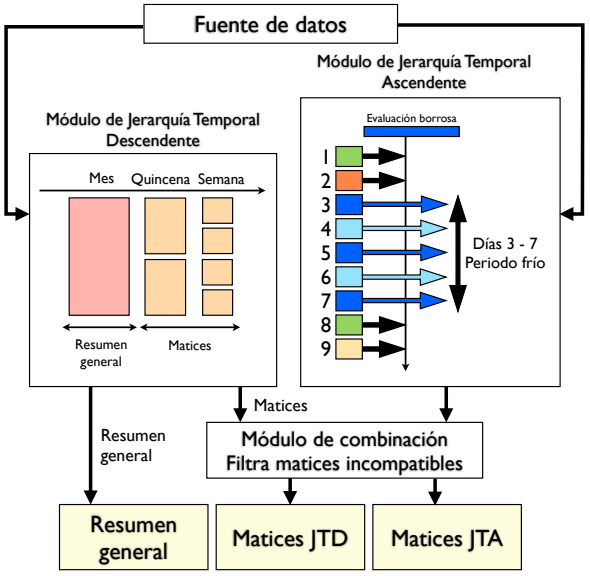


Figura 1: Esquema general del algoritmo de construcción de resúmenes.

ción se definen para cada combinación.

Dada una combinación C, compuesta por I matices (C_1,C_2,...,C_i,...,C_I), se definen las siguientes propiedades:

- **Etiqueta de referencia E_{C_i} y etiqueta asociada A_{C_i}.** Cada uno de los matices C_i tienen asociada una etiqueta de temperatura A_{C_i} (el matiz en sí), que se compara con la etiqueta correspondiente al periodo matizado E_{C_i}.
- **Longitud de la combinación l_C.** Número de matices que conforman la combinación.
- **Grado de cumplimiento del matiz p_{C_i}.** Mide, para un matiz C_i dado, la proporción de puntos en ese intervalo temporal que pertenecen a la etiqueta de temperatura E_{C_i} asociada a dicho matiz.

De entre los criterios de evaluación disponibles [5], se han seleccionado los siguientes:

- **Evaluación de longitud de la combinación F_l.** En general, se prefieren resúmenes que no sean largos, por lo que se penalizan longitudes excesivas (más de 2 matices). Para ello se define la función de evaluación de longitud $Fl(l_C)$ como su la pertenencia de l_C al número borroso trapezoidal definido en la figura 2a.

- **Grado de cumplimiento de la etiqueta F_p.** Puesto que las etiquetas asociadas a los intervalos temporales de la jerarquía descendente se calculan sobre la diferencia de medias, resulta importante conocer el porcentaje de días pertenecientes a ese periodo que realmente se corresponde con

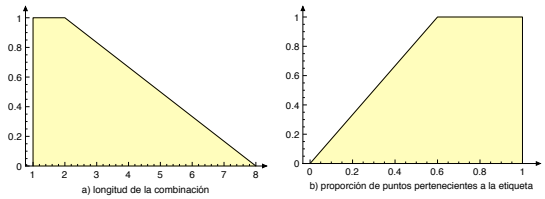


Figura 2: Funciones de evaluación de a) la longitud de la combinación, $Fl(l_C)$ y b) del cumplimiento de la etiqueta, $Fc(p_{C_i})$ para un periodo dado, C_i.

la etiqueta. Esto permite medir la variabilidad del intervalo, de tal modo que cuanto menor sea el porcentaje, más variable será el periodo y por tanto menos información aportará. Cada matiz C_i tiene asociados un porcentaje de puntos p_{C_i} que pertenecen a la etiqueta E_{C_i} correspondiente. Así pues, se define la función de evaluación de cumplimiento $Fc(p_{C_i})$ como la pertenencia de p_{C_i} al número borroso trapezoidal de la figura 2b.

La agregación de todos los matices se realiza con la t-norma producto para cada intervalo C_i:

$$Fp(C) = \prod_{i=1}^{I} Fc(p_{C_i}) \quad (1)$$

- **Grado de matización F_m.** Este criterio es el más novedoso de los tres, y debe atender distintos casos, para cada uno de los cuales es preciso definir una función de evaluación distinta. Por ejemplo, supongamos que en febrero las temperaturas fueron normales respecto a la media climática. Los matices interesantes serán aquellos que se encuentren a ambos extremos del conjunto de etiquetas. Sin embargo, si el mes ha sido muy cálido, una semana con frío dentro del mismo, aunque no sea extremo, resultará igualmente relevante. Por el contrario, las semanas muy cálidas no aportarán ninguna información y su relevancia será nula.

A partir de los ejemplos anteriores surge una idea simple, pero bastante ajustada a la definición de matiz. Una vez que se conoce la etiqueta sobre la cual se va a realizar la comparación, en términos generales, cuanto mayor sea la distancia entre esa etiqueta y la que va a ser evaluada, mayor será el grado de matización.

Sin embargo, es posible que se den situaciones en las que esto no tiene por qué ser exactamente así. Un ejemplo interesante que ilustra este caso sucede cuando el mes se encuentra en una etiqueta que no es la intermedia, pero tampoco la extrema. Por ejemplo, en un mes cálido, se podría afirmar que los períodos fríos serán relevantes (suponen un matiz de excepción) y que los muy fríos obtendrán la evaluación más alta. Sin embargo, existen también períodos muy cálidos, que son los que han arrastrado la media mensual hacia la etiqueta “cálido". Dichos períodos suponen un

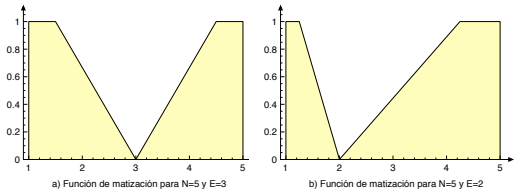


Figura 3: Representación gráfica de la función de matización, Fm, para a) 5 etiquetas, con la etiqueta 3 (temperatura normal) como referencia y b) para 5 etiquetas, con la etiqueta 2 (temperatura fría) como referencia.

matiz de intensificación del resumen general y presentan información relevante, a pesar de que la distancia respecto a la etiqueta mensual es mínima.

Estas ideas se han incluido en la definición de la función de matización, que consta de dos números borrosos trapezoidales que se construyen dinámicamente, en función del número de etiquetas total y la etiqueta de referencia para la comparación (Figura 3).

Su definición se formaliza del siguiente modo: sea N el número total de etiquetas, y E_{C_i} la etiqueta de referencia para un matiz C_i dado. Definimos los dos números borrosos trapezoidales, $T1$ y $T2$ como:

$$\begin{aligned} T1 &= trap[1,1,1+(E_{C_i}-1)/4,E_{C_i}] \\ T2 &= trap[E_{C_i},N-(N-E_{C_i})/4,N,N] \end{aligned} \tag{2}$$

Sea A_{C_i} la etiqueta perteneciente al intervalo C_i de la combinación C sobre la cual se aplica la función de matización Ft. Definimos $Ft(A_{C_i})$ como:

$$Ft(A_{C_i}) = max(T1(A_{C_i}),T2(A_{C_i})) \tag{3}$$

Al igual que en el criterio anterior, el producto de los $Ft(A_{C_i})$ de la combinación proporcionará la puntuación de dicha combinación en cuanto a matización se refiere. Definimos este criterio como:

$$Fm(C) = \prod_{i=1}^{I} Ft(A_{C_i}) \tag{4}$$

Finalmente, la agregación de los tres criterios se obtiene utilizando la t-norma producto:

$$Fg(C) = Fl(l_C) \cdot Fp(C) \cdot Fm(C) \tag{5}$$

El valor numérico que resulta de aplicar la función Fg a una combinación determina la posición de dicha combinación en el ranking global, que se ordena de mayor a menor. En caso de obtener resultados similares para muchas combinaciones se introduce un criterio adicional de filtrado, consistente en realizar una ordenación posterior en base a la

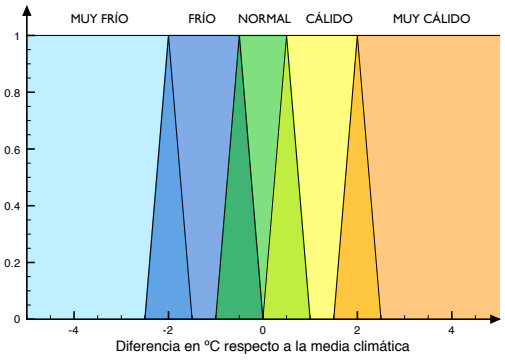


Figura 4: Conjunto borroso para las diferencias de temperatura.

longitud temporal de las combinaciones y su nivel de jerarquía, de modo que para varias combinaciones con la misma puntuación, tendrán preferencia aquellas que abarquen un mayor número de días.

El **módulo de jerarquía temporal ascendente** funciona de manera inversa al anterior módulo, partiendo de una granularidad temporal a nivel de día. Su objetivo es la obtención de períodos temporales más pequeños, pero igualmente informativos. Para ello, analiza el comportamiento de las temperaturas a lo largo del mes, identificando grupos de días consecutivos que pueden ser relevantes para su inclusión en el resumen completo y que no han sido detectados por el primer módulo. Para la obtención de estos intervalos se utiliza una evaluación borrosa de las diferencias de temperatura (Figura 4) que permite la inclusión de intervalos aproximados que relajan las restricciones de la definición precisa original, donde puedan aparecer días que no pertenecen completamente a la etiqueta bajo evaluación, pero que tampoco son suficientemente significativos para indicar un final de secuencia. Para ello se establecen unos umbrales de cumplimiento que exigen un elevado grado de verificación de las etiqueta. La Figura 1 muestra un ejemplo donde la evaluación borrosa de la etiqueta frío permite la consideración del período que va desde el día 11 al 15, incluyéndose dos días (12 y 14), que cumplen parcialmente dicha etiqueta.

Por último, el **módulo de combinación** se encarga de establecer un control sobre todos los intervalos temporales que resultan de la aplicación de los módulos de jerarquía temporal descendente y ascendente, de modo que sobre la selección de las cinco mejores combinaciones producidas por el primer módulo se comprueba la compatibilidad de los matices obtenidos por el segundo. La incorporación de dichos matices se da siempre que el grado de matización sea elevado y el resumen no resulte excesivamente largo. En caso de que no existan combinaciones válidas a nivel de referencias temporales, las comparaciones se realizan sobre

el resumen general directamente.

3 VALIDACIÓN

Los resúmenes generados por el método automático se validaron contrastándolos con los obtenidos por los meteorólogos del Servicio Meteorológico de la Xunta de Galicia [1], para los 20 meses que van de octubre de 2009 a mayo de 2011. Los datos proceden de un conjunto de nueve estaciones meteorológicas representativas de la geografía y climatología gallega. Los resúmenes con los que se ha validado el algoritmo son los informes climatológicos mensuales publicados oficialmente por Meteogalicia, y que son generados de forma manual por los expertos. Como ya se ha indicado, dichos informes, en lo que se refiere a la temperatura mensual, constan de un resumen general que valora la temperatura del mes frente a la media climática (valoración comparativa relativa frente al histórico del período objeto del informe) y una serie de matices que complementan la información aportada por el resumen general, destacando los episodios de mayor interés a criterio del experto. La validación se centró en varios aspectos: i) El resumen mensual general debe coincidir con el generado por el experto, ii) los resúmenes automáticos deberán obtener intervalos temporales compatibles con los originales y iii) siempre debe aparecer una combinación de sentencias similar a la generada por el experto.

Debemos considerar que aunque los resúmenes manuales siguen un estilo y una estructura sintáctica bien definida y repetitiva, la semántica de los mismos no sigue unos protocolos perfectamente establecidos, sino que provienen de la experiencia de los meteorólogos. Para ayudar a verificar los resultados y, al mismo tiempo, comprobar que los resúmenes originales sean coherentes con los datos, se han revisado adicionalmente las gráficas de temperatura.

La Tabla 1 recoge los resultados para los meses sobre los que se ha probado el método propuesto. Se distinguen entre los casos en los que i) la coincidencia entre los intervalos recogidos por el meteorólogo y el resumen generado es muy elevada, ii) cuando este último amplía el original con nuevos matices, iii) aquellos en los que pueden existir ciertas discrepancias pero el resultado sigue siendo razonable o incluso mejor que el del experto, y iv) los casos no válidos debido a que no coinciden en ningún aspecto con el resumen original ni presentan uno que represente correctamente los datos procesados.

Se puede afirmar que el método planteado obtiene en la mayoría de los casos resúmenes similares a los del experto. En algunos casos es capaz de precisarlos mejor y en otros la coincidencia no es plena, pero con la ayuda de las gráficas se puede comprobar que todos los resultados se ajustan a los criterios establecidos para la selección de los mejores resúmenes. No se ha observado ningún resultado anómalo.

	Núm. casos	% casos
Coincidentes	12	60
Compatibles (mejora respecto al resultado original)	4	20
Compatibles (difiere del original, pero resume bien los datos)	4	20
No válidos	0	0
Total	20	100 %

Tabla 1: Resultados de las pruebas efectuadas sobre los meses correspondientes al intervalo 10/2009-5/2011.

A título ilustrativo, presentamos a continuación dos ejemplos representativos de los casos de validación.

Diciembre de 2009

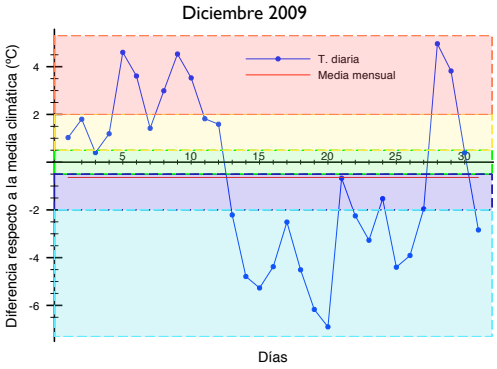


Figura 5: Diferencia de temperaturas respecto a la media climática para diciembre de 2009 en Galicia.

Resumen de Meteogalicia: *"Las temperaturas fueron bajas en general en toda la Comunidad. La media de las temperaturas quedó marcada por los períodos fríos que vivimos en la segunda y tercera semana del mes"*.

Resultados del algoritmo propuesto:

- Resumen general: Etiqueta 2 (Frío)
- Matices: La segunda quincena del mes ha sido muy fría. Periodo cálido del 1 al 4, y periodo muy cálido del 5 al 10.

Si se observa la gráfica, se puede comprobar que el intervalo temporal señalado por el meteorólogo es inexacto, ya que más que de segunda y tercera semana, el periodo de bajas temperaturas al que se refiere, ocupa la tercera y parte de la cuarta semana. Este matiz ha sido detectado por el primer módulo, que detecta como periodo muy frío la segunda quincena, acompañado de dos matices obtenidos

por el segundo módulo, unos días cálidos seguidos de otros muy cálidos.

Julio de 2010

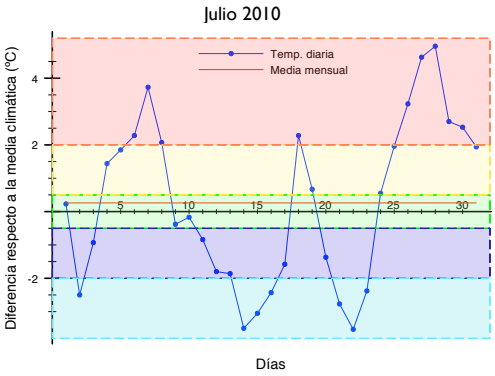


Figura 6: Diferencia de temperaturas respecto a la media climática para julio de 2010 en Galicia.

Resumen de Meteogalicia: *"Las temperaturas estuvieron en general en valores por encima de los normales para la época del año, impulsadas principalmente por los altos valores alcanzados en la última semana, donde se dieron temperaturas extremas"*.

Resultados del algoritmo propuesto:

- Resumen general: Etiqueta 3 (Normal)
- Matices: La última semana del mes ha sido muy cálida. Periodo muy frío del 14 al 16 y del 21 al 23.

El matiz correspondiente a la última semana, resaltado por el meteorólogo, ha sido obtenido como única combinación válida por el primer método. A dicho intervalo se le añaden dos cortos períodos muy fríos obtenidos por el segundo módulo. La única diferencia entre los dos resúmenes radica en el mensual, en el que el meteorólogo indica que los valores se encuentran por encima de la media climática, mientras que el algoritmo los considera normales, ya que la diferencia es muy pequeña. Tras estudiar todos los ejemplos, existen algunos en los que, a pesar de estar ligeramente por encima, la media se considera normal. En otros, como en este caso, sí se señala el signo de la comparación, aunque la diferencia sea mínima. Una posible solución, que sería siempre coherente, consistiría en indicar el signo de la diferencia (si la media está por encima o por debajo) y a continuación indicar el grado de la diferencia, que es lo que obtiene el algoritmo propuesto. En cualquier caso, pese a esta leve discrepancia, el resultado es totalmente coherente con el propuesto por el experto.

4 CONCLUSIONES

En este trabajo se han definido los matices de excepción e intensificación sobre un resumen general y se han aplicado a la generación de informes climáticos mensuales. La aplicación ha utilizado datos meteorológicos reales de temperatura, y los resultados de sumarización se han validado experimentalmente frente los informes publicados por los expertos del Servicio Gallego de Meteorología (Meteogalicia), dando lugar a resúmenes en general consistentes con los aportados por los expertos meteorólogos. En los casos donde no es así, se dan circunstancias de apreciación subjetiva (o errónea) del experto.

Como trabajo futuro se plantea la generalización de la arquitectura de la solución presentada, para la inclusión de otra información relevante -de este y otros ámbitos- como los perfiles temporales, así como el añadido de una interfaz de generación textual que mejore la expresividad final de los resúmenes obtenidos.

Agradecimientos

Este trabajo ha sido realizado en el marco de los proyectos TIN2008-00040 y TIN2011-29827-C02-02 (Ministerio de Ciencia e Innovación).

Referencias

[1] Web de Meteogalicia http://www.meteogalicia.es/observacion/informesclima/informesIndex.action. Accedido el 14 de noviembre de 2011.

[2] R. M. Castillo-Ortega, N. Marín, D. Sánchez. A Fuzzy Approach to the Linguistic Summarization of Time Series. Actas ESTYLF 2010, 267-272, 2010.

[3] A. van der Heide, G. Triviño. Automatic Generated Linguistic Summaries of Energy Consumption Data. Actas 9th ISDA Conference, 553-559, 2009.

[4] J. Kacprzyk. Computing With Words Is An Implementable Paradigm: Fuzzy Queries, Linguistic Data Summaries, and Natural-Language Generation. IEEE Trans. Fuzzy Systems, 18, 3, 451-472, 2010.

[5] F. Díaz-Hermida, A. Bugarín. Semi-fuzzy Quantifiers as a Tool for Building Linguistic Summaries of Data Patterns. Actas 2011 IEEE Symposium on Foundations of Computational Intelligence, 45-52.

[6] Eli Goldberg, Norbert Driedger, Richard I. Kittredge. Using Natural-Language Processing to Produce Weather Forecasts. IEEE Expert, 9, 45-53, 1994.

[7] F. Portet, E. Reiter, A. Gatt et al. Automatic Generation of Textual Summaries from Neonatal Intensive Care Data. Artificial Intelligence, 173, 7-8, 789-816, 2009.

Finding motifs in DNA sequences

David Gutiérrez-Avilés[1], **Francisco Martínez-Álvarez**[2], **Cristina Rubio-Escudero**[1], **José C. Riquelme**[1]

[1]*Department of Languages and Systems, University of Seville, Spain*, {davgutavi, crubioescudero, riquelme}@us.es
[2]*Department of Computer Science, Pablo de Olavide University of Seville, Spain*, fmaralv@upo.es

Abstract

Many approaches are currently devoted to find DNA motifs in nucleotide sequences. However, this task remains a challenge due to the difficulties found to deeply understand gene regulatory mechanisms, especially when analyzing binding sites in DNA. These sites or specific nucleotide sequences are known to be responsible for transcription processes. This work aims at providing an updated overview on strategies developed to discover meaningful motifs in DNA-related sequences, and, in particular, their attempts to find out relevant binding sites. From all existing approaches, this work is focused on dictionary, ensemble, and artificial intelligence-based algorithms since they represent the classical and the leading ones, respectively.

Keywords: motifs discovery, DNA sequences, regulatory mechanisms, binding sites.

1 Introduction

Genes are the main biological units of heredity and can be understood as a string of bases or sequence of chemicals whose apparently random combination encodes the hereditary information as well as genetic traits (individual characteristics). In particular, genes are composed of four different bases: Adenine (A), thymine (T), cytosine (C) and guanine (G).

In DNA helices it is usual to find particular sequences of bases, called binding sites, devoted to regulate transcriptions factors or factors that regulate gene expression by activating or inhibiting transcription machinery. In other words, from all the sequence of bases, these short specific sequences (typically from five to twenty pair long, versus 220 million existing base pairs long) are bound by more than one DNA-binding protein complexes.

Therefore, DNA motifs are meaningful base sequence patterns that identify binding sites responsible for transcription factors, and are known to appear in different genes and more than once within a gene. Finding motifs in DNA is a challenging task that many researchers are fielding nowadays insofar as their discovery may lead to major understanding of evolutive processes in living organisms.

Since Stormo [33] reviewed strategies to find motifs with computer algorithms, a large amount of algorithms have been developed. A major classification of these algorithms is done according to the type of DNA data sequence used to find motifs. Although there is no universal consensus on how to divide algorithms based on the input data, three major groups are generally considered [7]:

1. The algorithms that use information from coregulated genes from a single genome.
2. The algorithms that use information of a single gene in multiple species.
3. The algorithms that use information from phylogenetic footprints.

Despite the large amount of works found and grouped in classes 2) (see [28, 34]) and 3) (see [1, 12]), this paper is focused on providing a general overview on current algorithms that make use of the information that promoter sequences of coregulated genes generate. Actually, these methods can also be subdivided into multiple strategies, but this work only examines those based on dictionaries, ensembles and artificial intelligence-based techniques, as they represent the classical and the leading ones, respectively.

The rest of the paper is structured as follows. Section II discusses the most relevant works recently published related to dictionary-based algorithms. On the other hand, Section III presents ensemble algorithms used to find motifs in DNA.

As for Section IV, it presents the latest AI works in the DNA motifs discovery field. Finally, Section V provides a brief summary of strengths and weaknesses of the reviewed strategies.

2 Dictionary-based algorithms

These are enumerative algorithms which, in contrast to heuristic methods, exhaustively cover the space of all possible motifs for a specific motif model description. The methodology progressively considers over-represented words, from short to long. The over-representativeness of a long word is computed as the weighted average of the short words in the current dictionary which could be part of the long word. Although this methodology is, in essence, word counting based, it can filter out many spurious motifs which are over-represented owing to their overlapping with some real motifs. Therefore, it has a higher accuracy than a pure word counting method. However, computing the over-representativeness of a longer word by concatenating shorter ones is problematic, and may miss substantial over-represented motifs.

In this sense, van Helden et al. [36] developed a motif finding algorithm. The program counts all nucleotide occurrences within the sequence, and estimates their statistical significance. An essential prerequisite is that the system has to be calibrated to take into account the uneven nucleotide representation in the genome. Although conceptually simple, the algorithm proved efficient for extracting motifs from most of the yeast *(Saccharomyces cerevisiae)* regulatory families analyzed. These motifs had been previously found by laboratory experimental analysis. Furthermore, putative new regulatory sites were predicted within upstream regions of coregulated genes. However, its range of detection is limited to relatively simple patterns that include short motifs with highly conserved cores.

Later, van Helden et al. [37] extended their method to find transcription factors forming a dimer, with each unit binding to a similar small element, accounting for the symmetry of the site. The fixed spacing in the DNA site is due to the existence of a linker domain in the transcription factor, separating the DNA-binding and dimerization domains. These are called spaced dyad motifs, for the detection of spaced pairs shared by a set of upstream regions. The method is based on a systematic counting of pairs of short words separated by a fixed distance (space dyads) followed by a calculation of their statistical significance. Because the spacer can be different for distinct motifs, the spacer length is systematically varied between 0 and 16. The significance of this type of motif can be computed based on the combined score of the two conserved parts in the input data or based on the estimated complete dyad frequency from a background dataset. There is a big drawback in van Helden et al. [36] approach: there are no variations allowed within an nucleotide.

This problem was addressed by Tompa [35] with a proposal of an exact dictionary-based method. Tompa took into account both the absolute number of occurrences and the background distribution and created a table that, for each sequence s of a given length, records the number n of sequences containing an occurrence of s. There is a fixed number of substitutions sb allowed for the occurrences. The existence of a motif is calculated based on the probability of having n occurrences in a random sequence according to the background distribution. Thus, Tompa proposed an efficient algorithm to estimate the probability that a single random sequence contains at least one occurrence of the sequence s from a set of background sequences based on a Markov chain.

Brazma et al. [2] used a dictionary-based approach that searches exhaustively for an a priori unknown regular expression-type patterns that are over-represented in a given set of sequences. This proposal is capable to discover various subclasses of regular expression type patterns of unlimited length common to as few as ten sequences from thousands. It was applied in two cases, (1) discovery of patterns in the complete set of > 6000 sequences taken upstream of the putative yeast genes and (2) discovery of patterns in the regions upstream of the genes with similar expression profiles. Among the highest rating patterns, most had matches to known motifs in yeast.

Sagot [32] introduced a dictionary-based approach for motif finding that is based on the representation of a set of sequences with a suffix tree. Vanet et al. [38] used suffix trees to search for single motifs in whole genomes of bacteria. Marsan and Sagot [27] extended this method to search for combinations of motifs. Representation of upstream sequences as suffix trees gave a large number of possible combinations, however, the implementation was still efficient.

Bussemaker [4] presented MobyDick, suitable for discovering multiple motifs from a large collection of sequences. This approach formalizes how one would proceed to decipher a *text* consisting of a long string of letters written in an unknown language in which words are not delineated. The algorithm is based on a statistical mechanics model that segments the string probabilistically into *words* and concurrently builds a *dictionary* of these words. MobyDick can simultaneously find hundreds of different motifs, each of them present in only a small subset of the sequences, e.g., between 10 and 100 copies within the 6000 upstream regions in the yeast genome. The algorithm does not need an external reference dataset to calibrate probabilities and finds the optimal lengths of motifs automatically. They illustrated and validated the approach by segmenting a scrambled English novel, by extracting regulatory motifs from the entire yeast genome, and by analyzing data gen-

erated from a few DNA microarray experiments.

Wang [39] approached the problem of motif finding from the perspective of steganography [40]. They thought of the sequences as if they formed a stegoscript in which functional transcription factor binding motifs were secret messages embedded in a text of background sequences.They developed WordSpy, a dictionary based motif finding algorithm, integrating a word counting method and a statistical model. The word counting method is used to examine every possible word and the statistical model to capture over-represented motifs and background words in the given sequences. The algorithm presents some advantages. First, it does not require a background sequence, because it is capable of modeling background words based on the steganographic approach to the problem. This is an important feature for applications where a true background sequence model is hard to determine. Second, WordSpy measures the over-representativeness of a word relative to that of all the other words modeled by the statistical model, resulting in an accurate measure of the over-representativeness. This feature helps to identify motifs of exact length. Third, the algorithm can incorporate gene expression profiling information to separate biologically significant motifs from spurious ones. Fourth, WordSpy is a discriminative motif finding algorithm. It can directly take as input two sets of sequences and find motifs that are over-represented in one set of sequences but not in the other set. Finally, the algorithm can conduct a whole genome analysis on the motifs that discovers the fidelity of the motifs to the given sequences.

3 Ensemble algorithms

Many motif finders have been proposed using different approaches. They have shown to be effective for discovering motifs in small living organisms, such as yeast [10]. However, their effectiveness remains unproven when dealing with huge DNA sequences belonging to more complex living forms.

In an attempt to solve this problem, Burset and Guigó [3] presented the idea of combining the outputs of several gene finding algorithms. Each algorithm typically covers only a small subset of the known binding sites, with relatively little overlap between the algorithms. They analyzed 9 motif finding programs with 570 DNA sequences. The dataset contained 2649 exons, and 174 of them where predicted by all programs and only 33 of them where not predicted by any of them. It is therefore advised to combine the results from multiple motif discovery tools, ideally covering a range of motif descriptions and search algorithms.

Harbison et al. [8] observed that different motif finders have different strengths. They successfully identified more binding sites by combining results of six motif finders compared to using only single finder. In fact, the benchmark datasets from Tompa et al. [10] also support this. By simply taking the union of all binding sites predicted by 10 selected motif finders, the sensitivity can be increased by more than double over each selected motif finder. However, the union of all predicted sites could contain a lot of noise therefore decreasing specificity. It is not trivial to distinguish the real binding sites from the noise.

Liu et al. presented EVIGAN [23] (EVidence Integration for Genome ANnotation using a Network). EVIGAN employs a dynamic Bayes net (DBN), a type of probabilistic graphical model that can accommodate multiple (possibly incomplete) gene predictions and other lines of evidence, yielding consensus gene models that maximize the probability of the evidence provided. The DBN model supports a wide variety of evidence types, including computational gene predictions, sequence homology search results, EST alignments and splice site predictions and it is easily extensible to incorporate other evidence types, such as proteomics hits, predicted domain architecture, SAGE tags, or Affymetrix tiling array data. EVIGAN's annotation process simulates an idealized human curator: different evidence sources are compared, those that tend to agree in particular contexts are assigned higher confidence and a consensus model is then created that reflects those confidence estimates. EVIGAN can produce a single consensus gene model or an ordered list of the n-best gene models, along with associated posterior probabilities for each. They applied EVIGAN to three large-scale datasets: The ENCODE regions of the human genome [30], and the genomes of *Plasmodium vivax* and *Arabidopsis thaliana* [9]. These experiments demonstrate that for all three species, EVIGAN achieves better performance than any individual data source used as evidence.

Rubio-Escudero et al. [31] formulated the motif finding as a classification problem. It was interpreted as a decision between which section of a sequence is protein coding and which is not. The methodology uses a multiobjective approach to extract the best methodsï¿½ aggregations by maximizing the specificity and sensitivity of their predictions individually. It was applied to the EGASP sets from the ENCODE Genome Annotation Assessment Project (EGASP) [11]. These datasets contain manually curated fragments of the human genome originating from the ENCODE project [30]. This dataset was selected by the EGASP assessment because the genes encoded in these regions were not used to train any particular gene predictor. Therefore, it is not a biased dataset. The aggregation of the results from various methods is accomplished using the union and intersection operators [13]. The methodology obtained successful results and consistently outperformed even the best individual approach and, in some cases, produced dramatic improvements in sensitivity and specificity. Moreover, they observed that even the worst methods contributed to the aggregation with more accurate programs.

4 Artificial Intelligence-based techniques

This section explores the latest AI-based works published in the field of DNA motifs discovery. In particular, this section reviews approaches based on evolutionary techniques, self-organized maps, clustering and support-vector machines techniques.

Thus, a new multi-objective genetic algorithm (MO-GA) was introduced in [41] for the dyad motif discovery issue. In particular, the authors focused on optimizing three features: The sum of pairs, the number of matches and the information content. They also proposed new genetic operators to carry out such a task. Another GA-based approach was described in [24]. However, this time, the authors preferred to adopt a mechanism to regulate the concentration so that both the population diversity and vaccine mechanism are maintained in order to inhibit degeneracy during the evolutive process. Also in 2010, a GA was developed in [22]. This algorithm used a stochastic optimization technique based on particle swarm optimization (PSO). In particular, they proposed a modification of the standard PSO algorithm to adapt it to the discrete values that DNA sequences exhibit. The authors claim that the approach is especially useful when gaps are present in the motifs.

The application of self-organizing maps (SOM) can also be found in the literature. Hence, a SOM-based clustering algorithm was presented in [21], in which the authors extracted binding sites in DNA sequences. The main novelty of this work was to consider two different types signals in DNA sequences, showing that treating them separably better results can be achieved. On the contrary, three self-organizing neural networks were presented in [25] to find short motifs. Another SOM-based technique called SOMIX was introduced in [20] to discover binding sites in a set of regulatory regions. The tool proposed a intra-node soft competitive procedure in each node model to achieve maximum discrimination of motif from background signals, by weighting two different models: position specific scoring matrix and Markov chain. As it happened in [21] and [25], this method was inserted in another SOM-based approach, called SOMBRERO [26], that constructed models for motifs that were structurally similar.

The use of clustering techniques is also a usually strategy among researchers in this area. Thus, a hierarchical model with variable number of clusters was described in [15]. In particular, they used the Gibbs sampling strategy to allow width variation for each of the motifs. Moreover, a tool called Matlign based on hierarchical clustering was presented in [17]. The authors claimed that the tool was capable of post-processing large collections of DNA sequence motifs and of providing a non-redundant set of motifs, which could be further associated to known regulatory elements. Also, the well-known Fuzzy C-means (FCM) algorithm was applied in [18] to identify motifs in some particular regions of DNA sequences. The authors also tested K-means and Expectation-Maximization algorithms, showing that the fuzzy solution outperformed all others. The use of the K-means, and in particular an improved version, has been also explored in [5]. Thus, based on a previous enhancement by Zhong et al. [42] to overcome the random initialization problem associated to the original K-means version, the authors proposed two granular computing models that use FCM to split the dataset into smaller ones. Once divided, they applied their own K-means clustering algorithm version to every set to extract meaningful knowledge, reducing thus time costs.

There are also some relevant works that made use of support-vector machines (SVM) techniques. Thus, an approach that used one-class SVM algorithms to recognize transcription factor binding sites was proposed in [16]. Its main feature lied on the assumption that there exist correlations between transcription factors. The use of SVM combined with evolutionary processes can be found in [19]. This time the authors developed a method to predict binding proteins in DNA sequences. Thus, they created several SVM modules that were successfully combined with position specific scoring matrix (PSSM) profiles, a sort of evolutionary information. Pavesi and Valentini [29] formalized the problem of predicting genes' functional information as a classification problem, by using SVM with non-linear kernels. The training of such SVMs were carried out by means of both some particular DNA motifs and statistical procedures. Finally, another SVM classifier can be found in [6], in which the authors used stochastic grammar rules to find regulatory DNA sequences that were indeed evaluated by means of SVM.

5 Conclusions

Gene expression regulatory mechanisms are widely studied among biologists and computer scientists. Particularly, most of their efforts are directed towards analyzing protein generation in the so-called binding sites. The discovery of these particular sequence of bases (or motifs) has generated numerous works insofar as they provide meaningful information on evolutive processes.

Many different strategies and subsequent approaches have been published. Consequently, experts need to have a piece of advice when selecting one or another algorithm in order to find motifs the best possible way. Although all of them are limited in what they can find [14], and it has been a challenging task to conduct studies on performance comparisons of motif finding tools, the scientific community agrees in labeling ensemble algorithms as the most effective ones owing to their capability of retrieving results from cooperative different methods.

In contrast, algorithms based on dictionaries have proven to be useful when analyzing small organisms but insufficient

in big organisms. However, their inherent simplicity makes them as popular as widespread, and many experts continue conducting research on this topic nowadays.

Despite dictionary-based and ensembles approaches provides the researcher with reasonably good results, the discovery of DNA motifs has to deal with enormous amounts of data, being difficult to mine them with classical methodologies. Therefore artificial intelligence techniques have turned into necessary tools to speed up the full analysis of such data, as the other ones usually are able to partially face the complexities associated to such a problem.

None of these algorithms claim to be the *panacea*, and they are not indeed. However, it is an undeniable fact that the combination of all their strengths are leading to important discoveries that are helping to better understand transcription mechanisms in genes and therefore in human beings.

Acknowledgements

The research has been partially supported by the Spanish Ministry of Science and Technology under project TIN2007-68084-C-00, and Junta de Andalucía under project P07-TIC-02611.

References

[1] T. L. Bailey, Mikael Bodén, T. Whitington, and P. Machanick. The value of position-specific priors in motif discovery using MEME. *BMC Bioinformatics*, 11(179), 2010.

[2] A. Brazma, I. Jonassen, J. Vilo, and E. Ukkonen. Predicting gene regulatory elements in silico on a genomic scale. *Genome Research*, 8:1202–1215, 2000.

[3] M. Burset and R. Guigó. Evaluation of gene structure prediction programs. *Genomics*, 34:353–357, 1996.

[4] H. J. Bussemaker, H. Li, and E. D. Siggia. From the cover building a dictionary for genomes: Identification of presumptive regulatory sites by statistical analysis. *Proceedings of the National Academy of Science of the USA*, 97(18):10096–10100, 2000.

[5] B. Chen, S. Pellicer, R. Harrison, P. C. Tai, and Y. Pan. Novel efficient granular computing models for protein sequence motifs and structure information discovery. *International Journal of Computational Biology and Drug Design*, 2(2):169–196, 2009.

[6] R. Damasevicius. Structural analysis of regulatory DNA sequences using grammar inference and support vector machine. *Neurocomputing*, 73(4-6):633–638, 2010.

[7] M. K. Das and H. K. Dai. A survey of DNA motif finding algorithms. *BMC Bioinformatics*, 8(Suppl 7):S21, 2007.

[8] C. Harbison et al. Transcription regulatory code of a eukaryotic genome. *Nature*, 431:99–104, 2004.

[9] J. E. Allen et al. Computational gene prediction using multiple sources of gene evidence. *Genome Research*, 14, 2004.

[10] M. Tompa et al. Assessing computational tools for the discovery of transcription factor binding sites. *Nature Biotechnology*, 23(1):137–144, 2005.

[11] R. Guigó et al. EGASP: the human ENCODE Genome Annotation Assessment Project. *Genome Biology*, 7(1):S2, 2006.

[12] V. Storms et al. The effect of orthology and coregulation on detecting regulatory motifs. *PloS ONE*, 5(2):e8938, 2010.

[13] P. Halmos. *Naive set theory*. Princeton, NJ: D. Van Nostrand Company, 1960.

[14] J. Hu and D. Kihara. Limitations and potentials of current motif discovery algorithms. *Nucleic Acids Research*, 33:4899–4913, 2005.

[15] S. T. Jensen and J. S. Liu. Bayesian clustering of transcription factor binding motifs. *Journal of the American Statistical Association*, 103(481):188–200, 2008.

[16] B. Jiang, M. Q. Zhang, and X. Zhang. OSCAR: One-class SVM for accurate recognition of cis-elements. *Bioinformatics*, 23(21):2823–2828, 2007.

[17] M. Kankainen and A. Löytynoja. MATLIGN: a motif clustering, comparison and matching tool. *BMC Bioinformatics*, 8(8):189, 2007.

[18] M. Karabulut and T. Ibrikci. Fuzzy C-Means based DNA motif discovery. *Lecture Notes in Computer Science*, 5226:189–195, 2008.

[19] M. Kumar, M. M. Gromiha, and G. P. S. Raghava. Identification of DNA-binding proteins using support vector machines and evolutionary profiles. *BMC Bioinformatics*, 8:463, 2007.

[20] N. K. Lee and D. Wang. SOMIX: Motifs discovery in gene regulatory sequences using self-organizing maps. *Lecture Notes in Computer Science*, 6444:242–249, 2010.

[21] N. K. Lee and D. Wang. SOMEA: self-organizing map based extraction algorithm for DNA motif identification with heterogeneous model. *BMC Bioinformatics*, 12(Suppl 1):S16, 2011.

[22] C. Lei and J. Ruan. Finding gapped motifs by a novel evolutionary algorithm. *Lecture Notes in Computer Science*, 6023:50–61, 2010.

[23] Q. Liu, A. J. Mackey, D. S. Roos, and F. C. N. Pereira. Evigan: A hidden variable model for integrating gene evidence for eukaryotic gene prediction. *Bioinformatics*, 24(5):597–605, 2008.

[24] J. W. Luo and T. Wang. Motif discovery using an immune genetic algorithm. *Journal of Theoretical Biology*, 264(2):319–325, 2010.

[25] S. Mahony, P. V. Benosa, T. J. Smith, and Aaron Golden. Self-organizing neural networks to support the discovery of DNA-binding motifs. *Neural Networks*, 19:950–962, 2006.

[26] S. Mahony, D. Hendrix, A. Golden, T. J. Smith, and D. S. Rokhsar. Transcription factor binding site identification using the self-organizing map. *Bioinformatics*, 21(9):1807–1814, 2005.

[27] L. Marsan and M. F. Sagot. Algorithms for extracting structured motifs using a suffix tree with an application to promoter and regulatory site consensus identification. *Journal of Molecular Biology*, 287:345–362, 2000.

[28] M. J. Palumbo and L. A. Newberg. Phyloscan: Locating transcription-regulating binding sites in mixed aligned and unaligned sequence data. *Nucleic Acids Research*, 38(2):W268–W274, 2010.

[29] G. Pavesi and G. Valentini. Classification of co-expressed genes from DNA regulatory regions. *Information Fusion*, 10(3):233–241, 2009.

[30] ENCODE project consortium. The ENCODE (ENCyclopedia Of DNA Elements) project. *Science*, 306:636–640, 2004.

[31] C. Rubio-Escudero, R. Romero-Záliz, I. Zwir, and C. del Val. Optimization of multiclassifiers for computational biology: Application to gene finding and gene expression. *Theoretical Chemical Accounts*, 125(3-6):599–611, 2010.

[32] M. Sagot. Spelling approximate repeated or common motifs using a suffix tree. *Lecture Notes in Computer Science*, 1380:111–127, 1998.

[33] G. D. Stormo. DNA binding sites: Representation and discovery. *Bioinformatics*, 16(1):16–23, 2000.

[34] D. Straume, R. F. Johansen, M. Bjørås, I. F. Nes, and D. B. Diep. DNA binding kinetics of two response regulators, PlnC and PlnD, from the bacteriocin regulon of Lactobacillus plantarum c11. *BMC Biochemistry*, 10(17), 2009.

[35] M. Tompa. An exact method for finding short motifs in sequences, with application to the ribosome binding site problem. In *Proceedings of the Seventh International Conference on Intelligent Systems on Molecular Biology*, pages 262–271, 1999.

[36] J. van Helden, B. Andre, and J. Collado-Vides. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *Journal of Molecular Biology*, 281:827–842, 1998.

[37] J. van Helden, A. F. Rios, and J. Collado-Vides. Discovering regulatory elements in non-coding sequences by analysis of spaced dyads. *Nucleic Acids Research*, 28:1808–1818, 2000.

[38] A. Vanet, L. Marsan, A. Labigne, and M. F. Sagot. Inferring regulatory elements from a whole genome. an analysis of Helicobacter pylori σ^{80} family of promoter signals. *Journal of Molecular Biology*, 297:335–353, 2000.

[39] G. Wang, T. Yu, and W. Zhang. WordSpy: Identifying transcription factor binding motifs by building a dictionary and learning a grammar. *Nucleic Acids Research*, 33:412–416, 2005.

[40] P. Wayner. *Disappearing Cryptography*. Morgan Kaufmann Publishers, 2002.

[41] F. Zare-Mirakabad, H. Ahrabian, M. Sadeghi, S. Hashemifar, A. Nowzari-Dalini, and B. Goliaei. Genetic algorithm for dyad pattern finding in DNA sequences. *Genes and Genetic Systems*, 84(1):81–93, 2009.

[42] W. Zhong, G. Altun, R. Harrison, P. C. Tai, and Y. Pan. Improved kmeans clustering algorithm for exploring local protein sequence motifs representing common structural property. *IEEE Transactions on Nanobioscience*, 4(3):255–265, 2005.

PREDICTING GENE NETWORK KINETICS FROM *CIS*-ACTING REGULATORY FEATURES

Oscar Harari [1], Coral del Val[1], Igor Zwir [1,2]
[1]*Department Computer Science and Artificial Intelligence, University of Granada, Granada, Spain,*
{harari, delval, igor}@decsai.ugr.es
[2]*Howard Hughes Medical Institute, Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA*

ABSTRACT
Genetic and genomic approaches have been successfully employed to assign genes to distinct regulatory networks. The strength of the connections in these networks must be specified to define the kinetics of a group of genes, but the uncertainty concerning the connections between genes, the ambiguity inherent to the biological processes, and the impossibility of experimentally determining the underlying biological properties only allow a rough prediction of gene interaction. Here we describe a framework that examines promoter sequences and identifies those *cis*-acting features that define transcriptional network motifs. Then, we employ an iterative process, based on Ordinary Differential Equations, to learn a network architecture that appropriately integrates these motifs into a full structure. The application of this method to the two component systems PhoP/PhoQ and the PmrA/PmrB in *Salmonella* enterica uncovered novel mechanisms that enable the inter-connection of these networks. The predictions were experimentally verified.

Keywords: Genetic Networks, *cis*-acting features, Kinetic Models, Clustering

1. INTRODUCTION

Gene expression is determined by interactions among regulatory proteins, called transcription factors, and RNA polymerase(s), as well as the interactions of the transacting factors and RNA Polymerase with *cis*-acting DNA sequences in the promoters of regulated genes [1]. Computational tools that look for these *cis*-elements in genome sequences and genome-wide gene expression patterns (usually in the form of microarray data) provide the raw material for the characterization and understanding of transcription regulation of target genes. Recurrent patterns of interactions among these features define network motifs, which are elementary building blocks [2]. Few works have been devoted to integrate those blocks into a more complex genetic network [3-5].

In this work, we present a framework to infer gene network topologies based on genome sequences, and previous knowledge obtained by experimental assays. First, we enhance the discovery of the network motifs by providing a computational approach to improve the sensitivity while detecting *cis*-acting elements in promoter regions, including the analysis of transcription factor binding sites and RNA polymerase binding sites. Second, we connect network motifs and incorporate knowledge form the literature to conform complete network kinetic models. These allow the automatically test of hypotheses about the network motif integration, and select the most probable one [6]. We also consider different qualitative measures such as the *realism* of the topologies, the *flexibility* to reproduce the distinct behaviors under the distinct stimuli and their *robustness* to preserve functional characteristics when their parameters change (*e.g.*, initial concentrations, degradation rates).

We applied our method to analyze the expression of genes controlled by the PhoP/PhoQ regulatory system of *Salmonella* enterica serovar Typhimurium [7]. This specie has a cross-link between PhoP/PhoQ and PmrA/PmrB two component systems, enabling the simulation of a variety of network motifs. Measurements of time-dependent mRNA levels validated that our predictions could describe distinct kinetics even within a same network motif.

1.1. PhoP/PhoQ and PmrA/PmrB two component systems

The PhoP/PhoQ two-component system constitutes a master regulator in *Salmonella enterica*, regulating the transcription of more than 3% of the genes in response of a low extra cellular Mg^{2+}. Some of the genes regulated by PhoP/PhoQ two-component system are also regulated by the PmrA/PmrB two-component system, which is related to the polymyxin B antibiotic inducted resistance; and resistance to cell death mediated by Fe^{3+} among others. Thus these target genes respond independently to two signals: high level of extra cellular Fe^{3+}, sensed by the PmrB protein; and low levels of Mg^{2+}, sensed by the PhoQ protein.

1.2. Modeling genetic networks

We employ continuous models that determine the level of gene expression and relationships among them. These allow capturing biological properties that can be experimentally observed. Ordinary Differential Equations (ODE's) are

good approximations that calculate the difference of concentration of species (i.e. RNA, proteins) along the time. Statical ODE's [6] model the systems when they reach their steady state (i.e. the system has reached an equilibrium in which the difference of concentrations of species in function of time is equal to zero). Dynamic models [5] enable the simulation of the gene expression behavior before their reach their equilibrium.

2. RESULTS

2.1. *Cis*-regulatory features devoted to infer PhoP regulatory network

We investigated four types of *cis*-acting promoter features by extracting the maximal amount of useful information from datasets and then creating models, which in turn allowed the inference of the PhoP regulon. We incorporated into our analysis and learning process the PhoP regulated genes of *Escherichia coli*, bacterium specie closely related to *Salmonella*, which has been widely studied. It has been showed that *phoP* gene could complement a *Salmonella* phoP mutant [7].

2.1.1. PhoP binding site patterns

We decomposed set of binding site sequences corresponding to PhoP into four patterns and then combined them to increased the sensitivity to weak sites without losing specificity (Fig. 1A) [8]. This allowed the recovery of promoters, such as that corresponding to the *E. coli hdeA* gene or the *Salmonella pmrD*, that had not been detected by the single position weight matrix model [9] despite being footprinted by the PhoP protein [10].

2.1.2. RNA polymerase binding site patterns and location

We identified six patterns among PhoP-regulated promoters of *E. coli* and *Salmonella* that combine promoter class and distance between the PhoP box and the RNA polymerase site [8]. For example, *ugtL* and *pagC* promoters share the PhoP box but differ in the distance of the PhoP box to the RNA polymerase binding site [11] (Fig. 1B). The RNA polymerase site feature was evaluated and obtained an 82% sensitivity and 95% specificity for detecting RNA polymerase sites [8].

2.1.3. *Activated/ repressed promoters*

We determined that the location of binding sites functioning in activation is different from that corresponding to sites functioning in repression [8] (Fig. 1C). For example, we identified a PhoP binding site at a relative distance to the RNA polymerase consistent with repression in the promoter region of the *hilA* gene, which encodes a master regulator of *Salmonella* invasion and had been known to be under transcriptional repression by the PhoP/PhoQ system [7]. Several promoters, including those of the *Salmonella pipD* and *nmpC* genes, were classified as candidates for being both activated and repressed, because the distance between the predicted transcription start site and the PhoP box is consistent with either activation or repression.

2.1.4. Binding sites for other transcription factors

We analyzed the intergenic regions of the *E. coli* and *Salmonella* genomes for the presence of binding sites for 54 transcription factors [8,12]. We then investigated the co-occurrence of 24 sites with the binding site of the PhoP protein in an effort to uncover different types of independent or orchestrated co-regulation of PhoP and other TFs (Fig. 1D). By analyzing both the binding site quality and the location of transcription factor binding sites, we increase the chances of identifying co-regulated promoters.). For example, the *Salmonella pmrD*, *ugd* and *yrbL* promoters and the *E. coli yrbL* promoter harbor PhoP- and PmrA-binding sites, consistent with the experimentally verified regulation by both the PhoP and PmrA proteins [13].

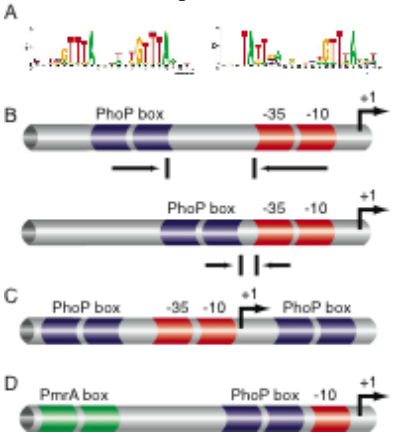


Figure 1. *Cis*-features identified in pomoter regions of PhoP regulated genes. (A) The PhoP protein recognizes a binding site motif consisting of a hexameric direct repeat separated by 5 bp, but distinguishes between different patterns with different specificities (i.e. *phoP* on the left and *pmrD* on the right). **(B)** PhoP-regulated promoters differ in the RNA polymerase sites. The PhoP-activated *ugtL* and *pagC* genes exhibit a class I sigma 70 promoter, but differ in the distance between the PhoP box and the RNA polymerase site. **(C)** The *mgtC* promoter harbors a PhoP binding box upstream of the RNA Polymerase binding site, positioned in a typical activation location. It also harbors a PhoP binding box downstream of the RNA Polymerase binding site, in a relative distance usual employed by PhoP to repress expression. **(D)** The *Salmonella pmrD* promoter harbors experimentally verified PhoP- and PmrA-binding sites.

2.2. PhoP network motifs uncovered by the *cis*-regulatory features

Complex biological systems are often modeled as networks. Network motifs are understandable patterns of connections

that occur significantly more often than pure chance or random networks [2], and alleviate the complexity of the study of transcriptional regulatory networks. One of the most elemental network motif is the autoregulation, whereas a gene is controlled by its own protein product (Fig. 2A). We identified a PhoP binding site upstream and close to a RNAP binding site indicating a positive autoregulation. Indeed. PhoP autoregulation is critical for *Salmonella* virulence [14]. Mice inoculated with wild-type *Salmonella* died, where mice inoculated with mutant *phoP* promoter survived. [14]. Other network motifs were identified in the same fashion (Fig. 2).

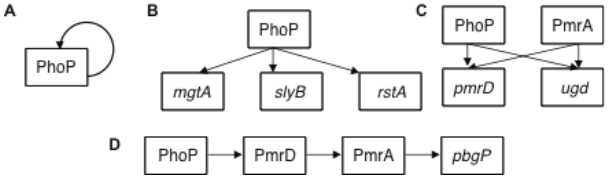


Figure 2. **The PhoP/PhoQ system employs a variety of network motifs to regulate gene transcription**. **(A)** Autorregulation: the *phoP* gene is regulated by its own product the PhoP protein. **(B)** In the simple-input module, PhoP as a single transcription factor regulates a set of genes. **(C)** In the bi-fan module, a set of genes (i.e. *pmrD* and *ugd*) are each regulated by a combination of transcription factors (i.e. PhoP and PmrA). **(D)** In the chained motif, genes are regulated in an ordered cascade.

Chained modules, allows transducing a signal (Fig. 2D). The cross-link between PhoP/PhoQ and PmrA/PmrB two-component systems is mediated by the *pmrD* gene, which resulting protein PmrD can bind the PmrA protein probably in a posttranscriptional or posttranslational fashion. As a result some of the genes governed by PmrA, including *pbgP*, are expressed even without the presence of the inducting signal of this two-component system. Curiously, the repression of *pmrD* by PmrA results in a negative feedback that closes the regulatory loop (Fig. 3). Although, this system has been widely studied [13], the exact mechanisms that defines the system dynamics is still controversial.

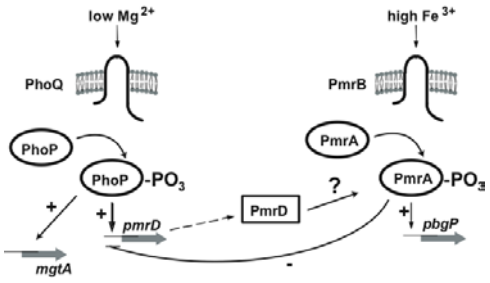


Figure 3: The PhoP/PhoQ-PmrA/PmrB functional scheme in *Salmonella enterica serovar Typhimurium*. The PhoQ protein senses low Mg^{2+} and the PmrB protein high Fe^{3+} concentrations from the environment and both proteins phosphorylate their cognate response regulators PhoP and PmrA, respectively. Although each of these proteins control the expressions of their target genes in response to their own signal, an alternative cross-talk suggest that some genes regulated by the PmrA protein can be regulated by PhoP in low Mg^{2+} conditions via the PmrD protein. Indeed, a transcriptional negative feedback has been detected in the *pmrD* gene.

2.3. Integrating network motifs into a genetic network

The initial model included the activation of *phoP, mgtA* and *pmrD* genes by phosporylated-PhoP, representing the positive autoregulation and simple network motifs respectively. Identically, we modeled the activation of *pmrA* and *pbgP* genes by phosporylated-PmrA. We translated the architecture into a system of ODE's, by employing the Ingeneue library [5], which allows simulating the dynamic behavior of the network architecture. The constraints imposed to this model only reflected the expression of *mgtA and pmrD* under low Mg^{2+} and *pbgP* on high Fe^{3+}. As expected, we observed the model could satisfy every input/output pattern with high probabilities, and could obtain high probability values (p=0.83) to satisfy all (*i.e.*, AND operation) of the constraints simultaneously.

The next proposed architecture incorporated the "backward" connection between PhoP/PhoQ and PmrA/PmrB systems as a first attempt to connect both network motifs in the whole network. Binding sites evidence and ChIP experiments showed that PmrA represses the expression of *prmD* gene [10]. Consequently, we specified that *pmrD* should be repressed under low Mg^{2+} and high Fe^{3+} environment concentrations (p=0.81).

Several frustrating attempts to integrate the network motifs into a whole network model derived the search towards relaxing all the constraints imposed to the system (i.e., values 1 in the tables). After an exhaustive search, we found that relaxing the expected output for the *pmrD* gene in the first constraint under low Mg^{2+} and high Fe^{3+} (Fig. 4; Table 1) allowed us to find a set of solutions with high probability that satisfy all constraints simultaneously (p=0.83). Moreover, this is the only architecture that could obtain such probabilities. The analysis of these results suggests that the *pmrD* gene can alternate between activated and repressing states according to the concentrations of its activation by the PhoP protein and repression by the PmrA protein. In addition, the state of this gene depends on the time when of occurrence of these events. We conclude that the uncertainty about the state of the *pmrD* gene does not constraint the connection of both network motifs when low Mg^{2+} and low Fe^{3+} occur (Table 1, constraint 2) and that, in that case, the PmrD protein protects the phosporylated-PmrA form the dephosphorylating activity of PmrB allowing the activation of the *pbgP* gene. Overall, the finding of sets of parameters that concurrently satisfy all constraints of the last architecture with high probability permits the interaction of all network motifs, including

simple, autoregulated, chained, and bi-fan, into a model of the whole regulatory network.

2.3.1. Learning parameters of the model

We tested two inference methods to estimate the parameters of the network by executing the native random walk already codified in the Ingeneue package [5], and compare the results obtained by a genetic algorithm (GA) implemented by ourselves.

Table 1: Patterns of input/output and constraints for the final PhoP/PhoQ-PrmA/PrmB architecture*

Input		Output			
Mg^{2+}	Fe^{3+}	*mgtA*	*pmrD*	*pbgP*	Probability
1	1	1	-	1	0,87
1	0	1	1	**1**	0,90
0	1	0	0	1	0,89
0	0	0	0	0	0,98
AND					0.83

* both values are acceptable. AND indicates probability of solutions satisfying all the constraints simultaneously

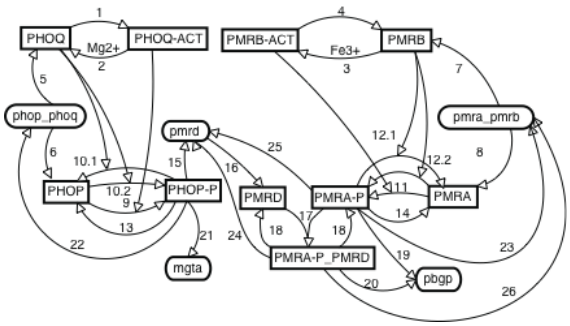


Figure 4: PhoP/PhoQ-PmrA/PmrB architecture. The species interact as follows: 1/2- Low/High Mg^{2+} level favors the PHOQ-ACT(ivated)/PHOQ state in equilibrium. 3/4- High/Low Fe^{3+} level favors PMRB-ACT(ivated)/PMRB state in equilibrium. 5/6- phop_phoq is translated into PHOQ/PHOP proteins. 7/8- pmra_pmrb is translated into PMRB/PMRA proteins. 9- PHOP is phosphorilated (PHOP-P) by PHOQ-ACT kinase activity. 10.1- PHOP-P is desphosphorilated to PHOP by PHOQ phosphatase activity. 10.2- PHOP is phosphorilated to PHOP-P by PHOQ kinase activity. 11- PMRA is phosphorilated to PMRA-P by PMRB-ACT kinase activity. 12.1- PMRA-P is desphosphorilated to PMRA by PMRB phosphatase activity. 12.2- PMRA is phosphorilated to PMRA-P by PMRB kinase activity. 13/14- PHOP-P/PMRA-P is spontaneous desphosphorilated to PHOP/PMRA. 15- PHOP-P activates the *pmrD* transcription. 16- *pmrD* is translated into PMRD. 17- PMRD binds PMRA-P (constituting PMRD_PMRA-P) which activates *pbgP* and represses *pmrD* genes, but it is not affected by the phosphatase activity of PMRB-ACT. 18- PMRA-P_PMRD unbinds into PMRD and PMRA-P. 19/20- PMRA-P/ PMRA-P_PMRD activates the transcription of *pbgP* gene. 21/22- PHOP-P activates the transcription of *mgta*/*phoP_phoQ*. 23- PMRA-P activates the transcription of *pmrA_pmrB*. 24/25- PMRA-P_PMRD/PMRA-P represses the transcription of *pmrD*. 26- PMRA-P_PMRD activates the transcription of *pmrA_pmrB*.

The solutions obtained were scored by a function that evaluates if the predicted concentrations of distinguished species match the expected ones (the lower, the better). The GA was executed using different configurations (i.e. population size, number of generations) and observed that both the population size and the maximum number of executions independently improve the quality of the results (Table 2). Furthermore, we compare the solutions obtained by the GA to the solutions obtained by the random walk approach, obtaining a score difference above 0.20 (Table 3).

Table 2: Evaluation of the performance of the GA.

Pop. size	Nr. Generations	Evaluations	Best score	Generation
50	100	5,000	0.1914	20
200	100	20,000	0.0522	9
50	250	12,500	0.0473	22

Table 3: Performance comparison (Random walk vs. GA)

Population size	Evaluations	Best score
Random Walk	100,000	>0.25
GA	1,100/12,500*	0.0473

* The GA obtained the best score after 1,100 evaluations. Heuristics like stall time can decrease the number of evaluations by indicating possible algorithm's stop condition.

2.3.2. Validating the prediction of simulated species

The mRNA levels, product of the transcription of our genes were experimentally measured with an interval of 15 minutes six times [14]. We interpolated the activity signal, and calculated the Pearson's coefficient of correlation (*c*) to the predictions of the model. We observed that PhoP, the distinguished specie that represents the positive autoregulation network motif, showed a high correlation (c=0.97), and exhibited a high level of expression. Similarly, the remaining network motifs showed a correlation of: c=0.95 for *mgtA* gene which represents the simple module; c=0.88 for *pmrD*, which exemplifies the bi-fan module; and c=0.92 for *pbgP* (chained module). These results reflect a highly correlated kinetics between our predictions and the experimentally observed values (Fig. 5).

2.3.3. Robustness of the model

Our analysis of the robustness of the network architecture for the PhoP/PhoQ-PrmA/PrmB system shows a tolerance of different magnitude order for distinct set of parameters. Indeed, we analyzed a random feasible solution for the final network architecture and found that some parameters (*e.g.*, Hill coefficient for *mgtA* - nu_phop_mgta) could take the entire biological meaningful range, and that only 3 of the 68 parameters could accept less than 25% of their entire biological significant range. We repeated this analysis for 10 additional random feasible solutions, grouped the parameters according to their types (*e.g.*, Phosphorylation, Dimerization, etc.) and calculated the average of the percentage of range acceptance (Fig. 6). We found high average values what showed the robustness quality of our final architecture.

2.3.4. Predicting simple input network motif regulated genes by scanning ranges of feasible solutions

Genes embedded in a same network motif can show differential expression. Thus, we hypothesize about the different kinetic behavior that genes co-regulated by PhoP might exhibit by scanning the parameters related to the *mgtA* specie.

We simulated the previously learned system, ranging the Hill coefficient for *mgtA* (*i.e.*, nu_mgtA took the values of 1, 5 , 10) while also changing the half-maximum rate of activation (i.e., k_mgtA was assigned with 0.01, 0.025, 0.05, 0.075, and 0.1); the remaining parameters were not modified (i.e, H_PhoP_P=10; H_mgtA=20). We obtained 15 dynamics and observed that these could produce different kynetics. We clustered the patterns by applying the hierarchical method (Fig. 7), and found three groups that exhibit distinct kinetics. Cluster 1 groups patterns that show smooth peaks of expression that tends to decay fast. Cluster 2 characterizes patterns that also exhibit smooth peaks, but in contrast to the previous one, these transcriptions decay slowly. Finally, patterns grouped by cluster 3 show a high rate of transcription after minute 20.

We employed experimentally measured mRNA levels to evaluate the kinetics of PhoP regulated genes [14], and calculated the correlation (c) between these observation and the patterns predicted (Fig. 8). Our analysis showed that pattern 7 (cluster 3) predicted the dynamics of genes with early rise time and high level of transcription (*i.e. mig-14* c=0.76); pattern 5 (cluster 2) correlates to genes with a late rise time and low level of expression (*i.e. pagC* c=0.71)); and finally that pattern 8 (cluster 1) correlates to kinetic behavior of *pagD* gene (*i.e.*, c=0.73).

1. CONCLUDING REMARKS

In this work, we proposed a framework to cope to the difficulties that arise at every inference stage. We provided tools to improve the sensitivity of detecting the *cis*-acting elements that determine the interaction among genes and their products, which in turn conform the whole set of network motifs instances. Employing an incremental approach, we integrated into a unique architecture the autoregulation, simple, bi-fan and chained modules, without compromising neither the *flexibility* nor the *robustness* of the final architecture. The predictions produced by the entire network for each of the network motifs correlated to the experimentally observed ones. We proved that the entire network preserves and recovers the network motifs kinetics, resulting in an adequate approach that solves the difficulties that arise when connecting previously identified network motifs. Moreover, simulating the dynamics of genes belonging to any of these modules independently of the reaming ones would allow the evaluation of the *realism* and *flexibility* of the proposed architectures, but would produce a cumbersome study of the *robustness* qualitative measures (*e.g.*, initial concentration of governing genes is a common parameter).

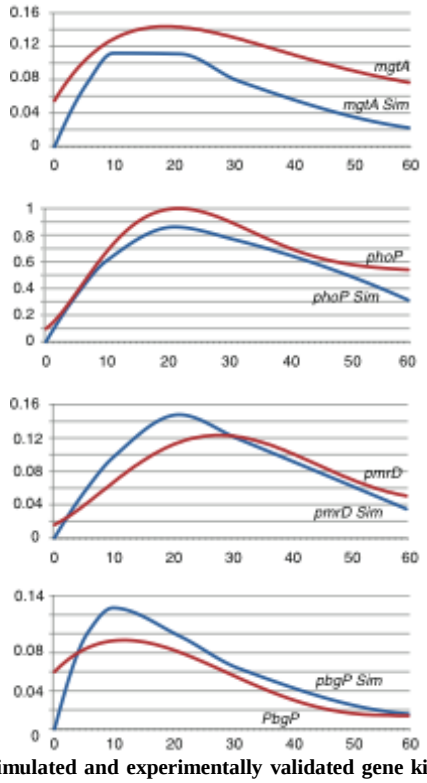


Figure 5: Simulated and experimentally validated gene kinetics. This charts reflects the high correlation between the predicted behavior and the experimentally obtained results (*i.e.*, mRNA expression quantified by real-time PCR).

2. MATERIALS AND METHODS

Tables and supplemental figures are available online at http://gps-tools2.its.yale.edu/Sim/Appendix.pdf

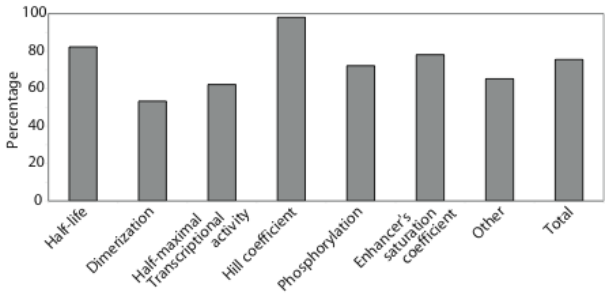


Figure 6: Robustness analysis. This chart shows the percentage of fulfillment for the biological meaningful range for parameters. The parameters are grouped by their type; the values represent the obtained average for 10 random solutions that satisfied the constraints).

ACKNOWLEDGMENT

This research was supported, in part, by Howard Hughes Medical Institute, and by The Spanish Ministry of Science and Technology grant TIN2009-13950 and GREIB grant to I.Z.

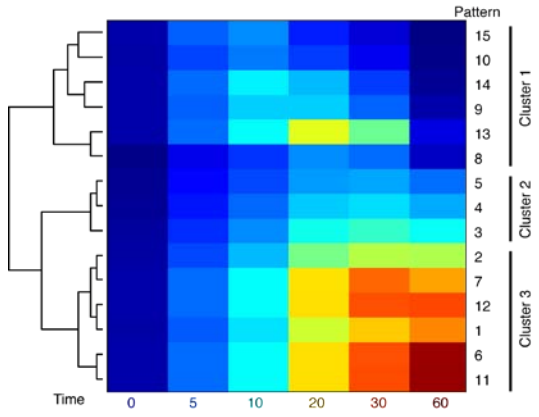


Figure 7. Identifying different kinetics obtained by scanning parameters of feasible solutions for simple network motif. The rows correspond to patterns obtained by different configuration; and the columns to the values (blue: low; red: high) obtained during the simulated time. The dendrogram (left panel) indicates the similarity (based on the correlation) of the patterns, revealing three characteristic clusters: low and short levels (cluster 1); low but longer peaks (cluster 2); and high level (cluster 3).

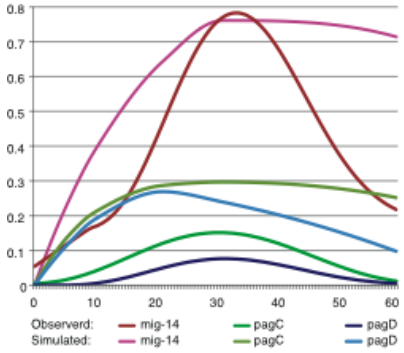


Figure 8. Simple network motif simulated dynamics vs. real-time PCR measured mRNA levels. Levels are normalized to the maximum one observed for PhoP gene (Fig. 5).

REFERENCES

1. Kærn, M., *Regulatory dynamics in engineered gene networks.*, in *4th International Systems Biology Conference*. 2003: Washington University, St. Louis.
2. Alon, U., *An introduction to System Biology*. Mathematical and Computational Biology Series, ed. C. Hall/CRC. 2007, London: CRC Press, Taylor & Francis Group.
3. Alon, U., *Biological networks: the tinkerer as an engineer.* Science, 2003. **301**(5641): p. 1866-7.
4. Shen-Orr, S.S., et al., *Network motifs in the transcriptional regulation network of Escherichia coli.* Nat Genet, 2002. **31**(1): p. 64-8.
5. Meir, E., et al., *Ingeneue: a versatile tool for reconstituting genetic networks, with examples from the segment polarity network.* J Exp Zool, 2002. **294**(3): p. 216-51.
6. von Dassow, G., et al., *The segment polarity network is a robust developmental module.* Nature, 2000. **406**(6792): p. 188-92.
7. Groisman, E.A., *The pleiotropic two-component regulatory system PhoP-PhoQ.* J Bacteriol, 2001. **183**(6): p. 1835-42.
8. Harari, O., et al., *Identifying promoter features of co-regulated genes with similar network motifs.* BMC Bioinformatics, 2009. **100 Suppl IEEE**.
9. Stormo, G.D., *DNA binding sites: representation and discovery.* Bioinformatics, 2000. **16**(1): p. 16-23.
10. Kato, A., T. Latifi, and E.A. Groisman, *Closing the loop: the PmrA/PmrB two-component system negatively controls expression of its posttranscriptional activator PmrD.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4706-11.
11. Barnard, A., A. Wolfe, and S. Busby, *Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes.* Curr Opin Microbiol, 2004. **7**(2): p. 102-8.
12. Salgado, H., et al., *RegulonDB (version 4.0): transcriptional regulation, operon organization and growth conditions in Escherichia coli K-12.* Nucleic Acids Res, 2004. **32**(Database issue): p. D303-6.
13. Zwir, I., et al., *Dissecting the PhoP regulatory network of Escherichia coli and Salmonella enterica.* Proc Natl Acad Sci U S A, 2005. **102**(8): p. 2862-7.
14. Shin, D., et al., *A positive feedback loop promotes transcription surge that jump-starts Salmonella virulence circuit.* Science, 2006. **314**(5805): p. 1607-9.

PERFILES COMPARATIVOS DE TRANSCRIPCIÓN DE ESPECIES SILVESTRES DE *SOLANUM* EN CONDICIONES DE SEQUÍA

M.F. Lucca [1], E. Hopp [1], R. Romero-Zaliz [2]

[1] *Instituto de Biotecnología, CICVyA - INTA, Buenos Aires, Argentina*, {flucca,ehopp}@balcarce.inta.gov.ar

[2] *Depto. de Ciencias de la Computación e I. A., UGR, Granada, España*, rocio@decsai.ugr.es

Resumen

La papa (*Solanum tuberosum*) es el tercer cultivo más importante y, a la vez, la dicotiledónea de más relevancia en la alimentación humana. La escasez de agua es la mayor limitación de la producción de cultivos. En respuesta a la sequía, la planta activa estrategias que implican cambios morfológicos, fisiológicos y bioquímicos; y, con frecuencia, rasgos poligénicos controlados por genes aditivos y sinérgicos. En este trabajo nos centramos en el análisis de los perfiles de transcripción en genotipos de especies silvestres de *Solanum* que presentan un comportamiento diferencial cuando hay escasez de agua al realizar hibridaciones de cDNA microarrays. El análisis bioinformático de datos de expresión de especies silvestres de *Solanum* mostraron diferencias significativas en los perfiles de transcripción. Estas diferencias entre las especies silvestres de papa sugieren nuevas formas de explotar la variabilidad genética del germoplasma silvestre para el mejoramiento de los cultivos.

Palabras Clave: Microarrays; clustering; bioinformática; papa.

1 INTRODUCCIÓN

La papa es un miembro clave de la familia *Solanaceae* y también es la dicotiledónea más importante en la alimentación humana, lo que constituye un alimento ideal en los países en desarrollo y especialmente en zonas con altitudes superiores a 2.000 metros, en el que hay un número limitado cultivos y donde el hombre y los animales necesitan un mayor aporte calórico.

El género *Solanum* está constituido por numerosas especies y 5500 diferentes variedades [7]. El hábitat de estas especies es muy variable, creciendo a diferentes alturas desde el nivel del mar hasta más de 4000 metros y en un amplio rango de temperaturas, fotoperíodo y suministro de agua. Además, los genotipos silvestres adaptados a diferentes condiciones ecológicas van desde las tierras altas, los trópicos a las regiones desérticas [13].

La base genética de las papas cultivadas de los Andes es muy estrecha, lo que dificulta el desarrollo de nuevas variedades de mayor rendimiento o mayor resistencia a los principales estreses bióticos y abióticos en este cultivo. El crecimiento y desarrollo de las plantas depende de factores genéticos y ambientales [10]. Utilizamos el concepto de "estrés" cuando uno o más de estos factores están en exceso o deficiencia, por tanto, el desarrollo y crecimiento de las plantas se ve afectado. Los estreses abióticos son la principal causa de bajo rendimiento en todo el mundo, reportando pérdidas de más del 50% en los principales cultivos [4][3]. La adaptación de las plantas al estrés implica la activación de varias respuestas fisiológicas, metabólicas y del desarrollo, siendo, con frecuencia, caracteres poligénicos controlados por genes aditivos y sinérgicos [2].

Los mecanismos en que los genes actúan en respuesta al déficit hídrico puede implicar diferentes caminos dentro de la célula, sufriendo cambios en su expresión genética. Estos cambios pueden ser regulados directamente por condiciones de estrés hídrico, estreses secundarios y/o respuestas a lesiones. Por otra parte, diferentes subconjuntos de transcripciones de ARNm se expresan en diferentes momentos durante el período de estrés hídrico. A pesar de los avances en biología molecular, los cambios en la expresión de genes inducidos por señales de transducción complejas aún no están establecidas claramente [5].

Con el fin de analizar la expresión de genes hemos utilizado microarrays ADNc de papa. Esta tecnología permite la medición simultánea de los niveles de expresión de miles de genes representados en el microarray [6]. A diferencia de la biología molecular tradicional, el uso exitoso de la técnica de microarrays requiere la aplicación de estadística e informática para el diseño de las matrices y los experimentos, y también para analizar y gestionar los datos. Para navegar por esta gran cantidad de información necesitamos herramientas bioinformáticas. La bioinformática es un campo en rápido desarrollo, cuya naturaleza multidisciplinar sirve para mejorar el conocimiento de los procesos biológicos, ya que implica el uso de técnicas diferentes, incluyendo las matemáticas, informática, estadística, inteligencia artificial, química y bioquímica para resolver problemas biológicos, por lo general, en el nivel molecular.

Una de las metodologías más conocida aplicacadas en bioinformática es la denominada *clustering* o agrupamiento de datos. Su use es frecuente en el análisis de datos de expresión de genes. Esta metodología permite la partición de un grupo de genes en una serie en grupos, de modo que los datos de cada uno participa en algún proceso biológico de interés, como ser, la falta de agua.

En este trabajo, se aplica clustering a un conjunto de datos de expresión génica con el análisis estadístico de microarrays para descubrir nuevos conocimientos de diferentes especies de *Solanum* silvestres. Esta información puede ser usada para mantener el metabolismo organismo y limitar el daño en períodos prolongados de estrés hídrico, ya que los genotipos silvestres de *Solanum* son una rica fuente de nuevos genes para su introducción en las especies cultivadas para mejorar la tolerancia al estrés.

2 MATERIALES Y MÉTODOS

2.1. TRATAMIENTO (SEQUÍA)

Para inducir escasez de agua, plantas pertenecientes a las especias *Solanum cardiophyllum* (CPH18325) y *S. venturii* (VNT8239) de cinco semanas de edad fueron ordenadas aleatoriamente en dos grupos (20-25 plantas por grupo). Un grupo de plantas fue sometido a condiciones de sequía (retención de agua) y el segundo grupo fue regado con normalidad. Todas las plantas fueron regadas 72hs antes de iniciar el tratamiento. La humedad de las plantas fue seguido por una cámara de presión Scholander y con la determinación del contenido relativo de agua (RWC) [12]. El tejido de las hojas de tres-cuatro plantas por grupo se recogieron en cuatro puntos de tiempo: 0, 5, 7 y 10 días después del inicio del tratamiento, luego se combinaron, se congelaron en nitrógeno líquido y se almacenaron a -80ºC.

2.2. EXTRACCIÓN DE ARN

Se extrajo el ARN del tejido de las hojas siguiendo el método de Trizol (Invitrogen). La integridad del ARN se verificó en gel de agarosa y la concentración se ajustó a 1µg/ml por precipitación con etanol y re-suspensión.

2.3. MICROARRAYS

Hemos diseñado los ensayos de microarrays utilizando la variedad de papa 10K cDNA versión 4 proporcionada por el Instituto de Investigación Genómica (TIGR). Un total de 15.264 cDNAs fueron vistos por duplicado en los microarrays (22.824 elementos (spots) 32.448 puntos en total).

Hemos seguido el estándar MIAME, Información Mínima sobre un Experimento de Microarrays), para nuestos experimentos. Con *n* especies, *t* puntos de tiempo y *r* repeticiones, se utilizaron 32 microarrays (2 especies, 4 puntos de tiempo; 4 repeticiones (dye-swap)).

2.4. PREPARACIÓN E HIBRIDACIÓN

La síntesis de ADNc se llevó a cabo mediante una reacción de etiquetado directo de la transcripción inversa siguiendo los protocolos de TIGR [15]. La hibridación de microarrays de vidro de papa se realizó de acuerdo a los protocolos de TIGR [16].

2.5. ANÁLISIS DE DATOS

La digitalización de los microarrays se realizó en un escáner Genpix (Axon Instruments, EE.UU.). Se ha utilizado el paquete de Bioconductor *Limma* para leer y analizar todos los microarrays. Se ha utilizado el método de *mínimos* para la corrección de fondo conjuntamente con la normalización por defecto dentro de cada matriz (*printtiploess*) y la normalización *aquantile* entre los microarrays. Los puntos duplicados fueron promediados. Se calculó un modelo lineal y todos los *p-values* fueron ajustados por el método empírico de Bayes para reducir al mínimo la tasa de falsos positivos. Solamente aquellos resultados con un p-value ajustado inferior a 0.003 se retuvieron para el posterior análisis.

2.6. CLUSTERING C-MEANS DIFUSO

El método *c-means* difuso (FCM) es un método de clustering que permite que un dato pertenezca a dos o más grupos. Este método (desarrollado por Dunn en 1973 y mejorado por Bezdek en 1981) se utiliza con frecuencia en el reconocimiento de patrones. Se basa en la minimización de la función objetivo cuya fórmula se muestra a continuación, con respecto a U, una partición difusa del conjunto de datos, y a V, una serie de prototipos de K:

$$J_m(U,V) = \sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m} \left\| X_j - V_i \right\|^2, \quad 1 \le m \le \infty \quad (1)$$

donde m es cualquier número real > 1, u_{ij} es el grado de pertenencia de X_j en el grupo i, X_j es el j-ésimo dato d-dimensional de los datos medidos, V_i es el centro d-dimensión del clustering y ||*|| es una norma que expresa de la similitud entre los datos medidos y el centro [1].

2.7. CLUSTERING DE LOS DATOS DE EXPRESIÓN

Los genes diferencialmente expresados se agruparon mediante el método c-means difuso utilizando la medida de distancia euclídea [1][11]. En lugar de agrupar los valores de expresión por sí mismos, se calculó la diferencia entre los puntos de tiempo contiguos y se usaron estos valores para realizar el clustering. Al utilizar estos nuevos valores, cambios en la expresión con el mismo signo y fuerza relativamente similares se agrupan juntos. Con esta técnica se simula un estudio de correlación sencillo. El clustering difuso se utiliza para proporcionar la posibilidad de un gen que pertenecen a más de un grupo, permi-

tiendo así que el experto para tomar la decisión final de qué grupo debe ser asignado cada gen.

3 ANÁLISIS DE RESULTADOS

Se analizaron todos los microarrays de ambas especies y se extrajo una lista de genes expresados diferencialmente. Los diez primeros genes diferencialmente expresados de cada especie se muestran en las Tablas I y II.

Se ha aplicado c-means difuso a la lista de genes diferencialmente expresados para obtener 10 grupos para cada especie, donde todos los genes fueron asignadas a un solo grupo, seleccionando el grupo de mayor pertenencia difusa (Figuras 1 y 2).

A partir de estos grupos podemos ver que hay varios genes con una expresión muy similar en ambas especies. A continuación, se estudiaron estos genes y también aquellos que tienen diferencias sustanciales. En este último grupo encontramos los resultados más interesantes. En las figuras 3, 4, 5 y 6 (las líneas verdes representan el control y las rojas el tratamiento (sequía)) se muestran algunos genes con un comportamiento diferente entre las Especies I y II.

El análisis comparativo del transcriptoma de CPH18325 y VNT8239 mostraron la puesta en práctica de diferentes estrategias para hacer frente al estrés, con las diferencias en el número de genes como en su cinética durante el tratamiento. En ambas especies la expresión diferencial de genes en respuesta al estrés validan el ensayo (Tablas 1 y 2), e. g., Metallothioneins 1, Ferritin y Dehydrins actúan para mantener la homeostasis celular. Los genes se comportaron en grupos en ambas especies en respuesta al estrés pero con distintos perfiles de expresión diferencial temporal (Figuras 1 y 2). En VNT8239, al analizar la expresión diferencial de algunos genes observamos patrones de expresión contrastantes, lo que sugiere una regulación común sincrónica, como se ejemplifica en el non-specific putative lipid transfer protein StnsLTP que contrasta con el Photosystem I reaction center subunit X psak y con la proteína Ethylene-responsive late embryogenesis-like que contrasta con la methyl esterase Pectin, lo que confirma la diversidad genética de las respuestas en el germoplasma de las especies silvestres de *Solanum* a la sequía. Mientras CPH18325 muestran poca expresión diferencial de genes (235 genes) asociados con el metabolismo de la pared celular, la transcripción, el metabolismo de proteínas, la señalización y los genes de respuesta a factores bióticos y abióticos, VNT8239 muestra en todo momento una expresión diferencial enorme que comprende de 4.133 genes. Del total de genes diferencialmente expresados las dos especies comparten la expresión de 151 genes.

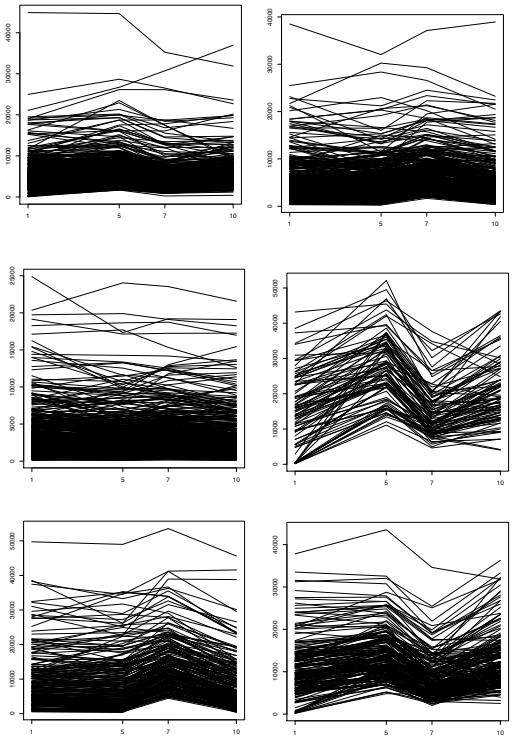
Algunos de estos genes han demostrado ser diferencialmente regulados por el estrés en ambas especies, como el Ammonium transporter 1 (STMEP30), aunque con una dinámica diferente en cada especie (Figura 3). Otros genes mostraron estar inducidos en ambas especias como ser Dehydrin Tas14, con una fuerte y temprana en CPH18325 (Figura 4). El precursor Polyphenol oxidase B (Figura 5) es un ejemplo de contraste de expresión en el grupo de genes comunes.

Aunque el análisis de los genes comunes y su dinámica nos ayuda a entender la respuesta diferencial a la sequía, las más interesantes son los genes específicos asociados con las especies más resistentes, *S. cardiophyllum*, tales como la Heat Shock Protein 90, que aparece como un gen candidato asociado con estrés hídrico en papa (Figura 6).

Estas diferencias de expresión entre las especies silvestres de papa sugieren nuevas formas de explotar la variabilidad genética del germoplasma de América para el mejoramiento del cultivo de la papa.

4 DISCUSIÓN

S. cardiophyllum 18325 ha demostrado ser más resistentes a la sequía que *S. venturii* 8239, con la RWC de 87 y 67% respectivamente. Los ambientes nativos de las especias CPH18325 y VNT8239 son muy diferentes. Se ha reportado que *S. cardiophyllum* proviene de Méjico, desarrollado con una vegetación arbustos secos, en el borde de los viejos campos y los cultivos [8], mientas que *S. venturii* fue descrito por Hawkes y Hjerting [9] y Spooner y Clausen [14] en el noroeste de Argentina, especialmente en laderas húmedas, incluyendo bosques de *Polylepis* y *Alnus*, o a lo largo de los campos de cultivo, caminos y terrazas de los ríos a una altitud entre 1900-3000 metros.



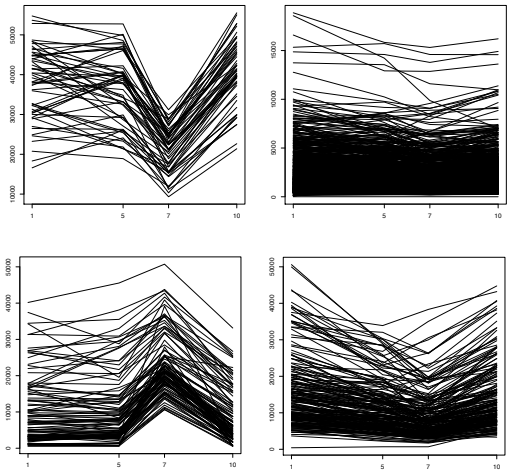


Figura 1: Clustering c-means difuso sobre los genes diferencialmente expresados de la Especie I (CPH18325). El eje x corresponde a los puntos de tiempo mientras que el eje y corresponde al nivel de expresión.

Watkinson et al [17] sugiere que la adaptación y aclimatación a la sequía implica la participación coordinada de las redes de regulación y, en algunos genes, relacionados con la señalización, la transcripción de genes y caminos metabólicos implicados en la distribución de recursos, la defensa y protección del estrés.

Especies silvestres de *Solanum* CPH18325 y VNT8239 han mostrado diferentes respuestas a la sequía. Estas especies provienen de zonas geográficas muy diferentes, lo que sugiere que las diferencias en las respuestas al de estrés reflejan la adaptación a su ambiente nativo a través de estrategias específicas de la expresión génica diferencial. La información obtenida en estas estrategias de adaptación pueden ser utilizados en programas de mejoramiento de la papa para tomar ventaja de la diversidad genética y el desarrollo de variedades resistentes a la sequía. Desde el *S. cardiophyllum* 18325 variadas en la expresión de sólo 235 genes y los genes 57 eran únicos, son genes candidatos interesantes a ser evaluados como nuevos genes de resistencia a la sequía.

Desde el punto de vista computacional, la agrupación de los genes con comportamiento similar ha dado una visión interesante y ayudado en el estudio de especies silvestres de *Solanum*. Se podrían haber utilizado métodos alternativos de clustering para agrupar genes con comportamiento similiar y tenemos la intención de ponerlos a prueba en un trabajo futuro para mejorar nuestros resultados preliminares.

Agradecimientos

Agradecemos al Dr. Enrique Ritter y al Instituto Neiker-Tecnalia por el germoplasma silvestre de *Solanum* utilizado en este estudio y las instalaciones donde se llevó a cabo los tratamientos de sequía.

R. Romero-Zaliz es apoyada por el Ministerio español de Educación y Ciencia (TIN2009-13950), Junta de Andalucía (TIC-02788) y el programa GENIL (PYR-2010-28).

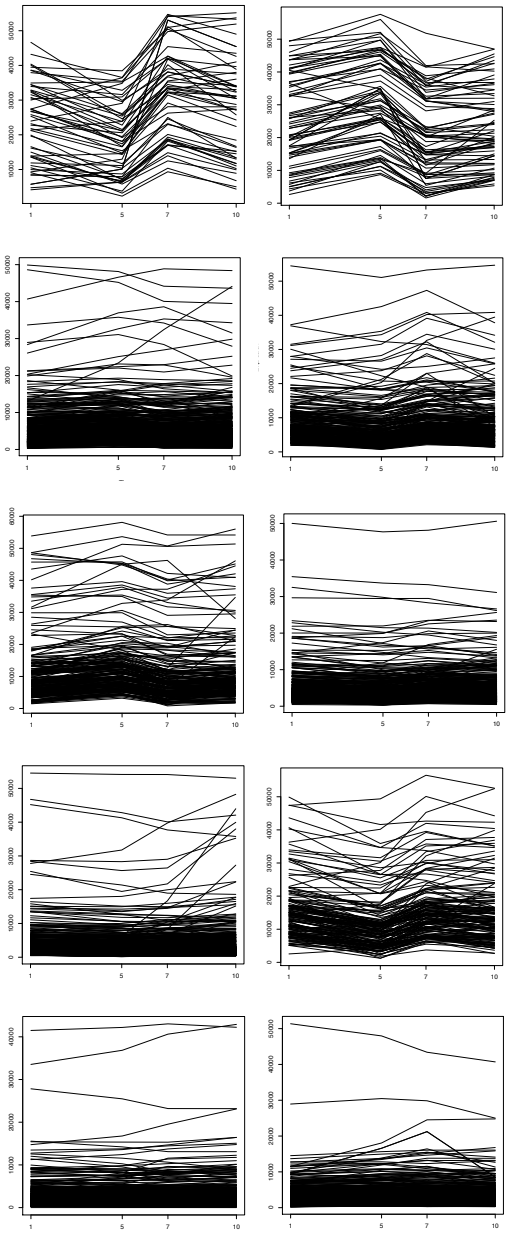


Figura 2: Clustering c-means difuso sobre los genes diferencialmente expresados de la Especie II (VNT8239). El eje x corresponde a los puntos de tiempo mientras que el eje y corresponde al nivel de expresión.

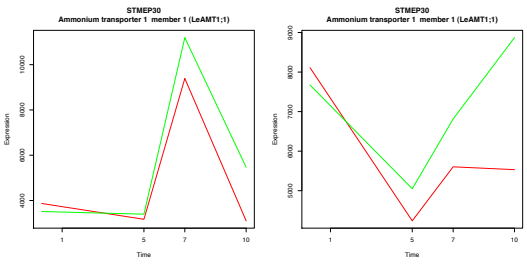


Figura 3: Lado izquierdo: STMEP30 en la Especie I (CPH18325). Lado derecho: STMEP30 en la Especie II (VNT8239).

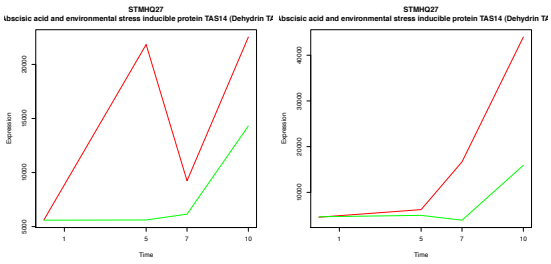


Figura 4: Lado izquierdo: STMHQ27 en la Especie I (CPH18325). Lado derecho: STMHQ27 en la Especie II (VNT8239).

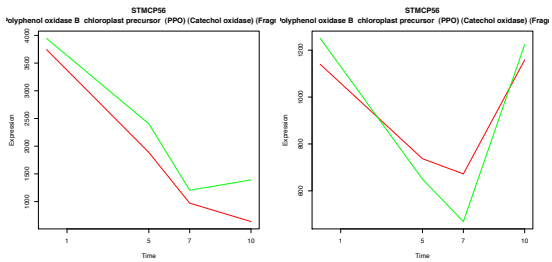


Figura 5: Lado izquierdo: STMCP56 en la Especie I (CPH18325). Lado derecho: STMCP56 en la Especie II (VNT8239).

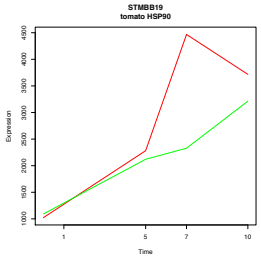


Figura 6: STMBB19 en la Especie I (CPH18325).

Referencias

[1] Bezdek, J. 1981. *Pattern recognition with fuzzy objective function algorithms*. Plenum Press.

[2] Bohnert, H.J. et al. 1995. Adaptations to Environmental Stresses. *The Plant cell*. 7, 7 (Jul. 1995), 1099-1111.

[3] Boyer, J.S. 1982. Plant Productivity and Environment. *Science (New York, N.Y.)*. 218, 4571 (Oct. 1982), 443-448.

[4] Bray, E. A., Bailey-Serres, J., and Weretilnyk, E. 2000. Responses to abiotic stresses. *Biochemistry and Molecular Biology of Plants*. B.B. Buchanan et al., eds. American Society of Plant Physiologists. 1158-1203.

[5] Bray, E.A. 2002. Classification of Genes Differentially Expressed during Water-deficit Stress in Arabidopsis thaliana: an Analysis using Microarray and Differential Expression Data. *Annals of Botany*. 89, 7 (Jun. 2002), 803-811.

[6] Bryan, G.J. and Hein, I. 2008. Genomic resources and tools for gene function analysis in potato. *International journal of plant genomics*. 2008, (Jan. 2008), 216513.

[7] FAO Statistics 2008: *http://www.potato2008.org/es/index.html*.

[8] Hawkes, J. 1990. *The potato : evolution, biodiversity and genetic resources*. Belhaven.

[9] Hawkes, J.G. and Hjerting, J.P. 1960. Some wild potato species from Argentina. *Phyton, Vicente Lopez*. 9, (1960), 140-46.

[10] Levitt, J. 1980. *Responses of Plants to Environmental Stresses (Physiological Ecology): Chilling, freezing, and high temperature stresses*. Academic Pr.

[11] Nock, R. and Nielsen, F. 2006. On weighting clustering. *IEEE transactions on pattern analysis and machine intelligence*. 28, 8 (Aug. 2006), 1223-35.

[12] Reigosa Roger, M.J. et al. 2003. *Handbook of Plant Ecophysiology Techniques*. Kluwer Academic Publishers.

[13] Rousselle, P. et al. 1999. *La Patata*. Mundi-Prensa Libros.

[14] Spooner, D.M. and Clausen, A.M. Wild potato (*Solanum* sect. Petota) germplasm collecting expedition to Argentina in 1990, and status of Argentinian potato germplasm resources.

[15] TIGR Protocol http://jcvi.org/potato/images/SGED_SOP_5.1.1.pdf.

[16] TIGR Protocol http://jcvi.org/potato/images/SGED_SOP_6.1.1.pdf.

[17] Watkinson, J. et al. 2006. Accessions of Solanum tuberosum ssp. andigena Show Differences in Photosynthetic Recovery After Drought Stress as Reflected in Gene Expression Profiles. *Plant Science*. 171, (2006).

Tabla 1: Diez mejores genes diferencialmente expresados para la especie I (CPH18325).

NOMBRE	NOMBRE DEL CLON	TIEMPO 5	TIEMPO 7	TIEMPO 10	P-VALUE AJUSTADO
peptidase M20/M25/M40 family protein similar to acetylornithine deacetylase	STMHR62	0.20	0.16	-0.96	1.90e-06
Metallothionein-like protein type 2 B	STMGS26	0.32	0.55	0.401	3.30e-06
Avr9/Cf-9 rapidly elicited protein 284	STMGP26	-1,31	-0.43	-0.21	1.41e-05
GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3	STMCV95	-0.09	-0.55	-1,78	2.62e-05
1-deoxyxylulose 5-phosphate synthase	STMCY65	-0.07	0.11	-0.86	2.62e-05
GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3	STMDB20	-0.14	-0.54	-1,70	2.62e-05
WIZZ	STMEP07	-1,04	-0.32	-0.74	2.62e-05
expressed protein	STMGJ17	-1,16	-0.10	-0.15	2.62e-05
null	STMEP27	0.45	-0.03	-0.60	2.62e-05
Ferritin 1 (Fragment)	STMGY57	1.34	0.06	-0.30	2.71e-05

Tabla 2: Diez mejores genes diferencialmente expresados para la especie II (VNT8239).

NOMBRE	NOMBRE DEL CLON	TIEMPO 5	TIEMPO 7	TIEMPO 10	P-VALUE AJUSTADO
Jasmonic acid 2	STMIY82	0.17	1.57	0.56	1.47E-02
dehydrodolichyl diphosphate synthase putative / DEDOL-PP synthase putative similar	STMIR06	0.11	0.41	1.30	3.90E-03
expressed protein	STMDO28	0.19	0.88	0.52	7.70E-03
2-oxoglutarate-dependent dioxygenase	STMGO18	-0.47	0.73	-0.88	9.09E-02
hypothetical protein	STMCZ51	0.08	0.73	-0.62	1.39E-01
Abscisic acid and environmental stress inducible protein TAS14 (Dehydrin TAS14)	STMHQ27	0.30	2.10	1.47	1.63E-01
24K germin like protein precursor	STMCX35	-0.03	-1.46	0.37	1.71E-01
17.6 kD class I small heat shock protein (Type I small heat shock protein 17.6 kDa isoform)	STMGB34	0.30	0.96	0.72	2.25E-01
TATA-binding protein-associated factor TAFII55 family protein contains Pfam profile	STMHN79	0.29	0.92	0.64	3.08E-01
Putative non-specific lipid transfer pro-tein StnsLTP	STMGQ20	0.19	1.87	0.57	3.39E-01

IDENTIFICACIÓN DE PERFILES DE PROMOTORES UTILIZANDO UN ENFOQUE MULTI-OBJETIVO BASADO EN RESTRICCIONES

R. Romero-Zaliz [1], J. Arnedo-Fdez [1], I. Zwir [1,2], C. del Val [1]

[1] *Depto. de Ciencias de la Computación e I.A., UGR, Granada, España*,
{rocio,arnedo,delval,igor}@decsai.ugr.es
[2] *Washington University, St. Louis, USA*

Resumen

La regulación de la expresión génica es un fenómeno complejo y dinámico esencial para todas las funciones biológicas. Las instrucciones necesarias para la expresión de los genes están codificados en *cis*-elementos reguladores que trabajan juntos e interactúan con la ARN polimerasa para conferir los patrones espaciales y temporales de la transcripción. Por lo tanto, la identificación de estos elementos es actualmente un área activa de investigación en el análisis computacional de secuencias reguladoras. Sin embargo, el problema es difícil, ya que las interacciones combinatorias entre los factores que regulan pueden ser muy complejas. Aquí presentamos un servidor web que identifica módulos de *cis*-elementos dado un conjunto de sitios de unión de factores de transcripción y sitios de ARN polimerasa para un grupo de genes.

Palabras Clave: Multi-objetivo, módulos *cis*-elementos; regulación genética.

1 INTRODUCCIÓN

La regulación de la expresión génica es un fenómeno complejo y dinámico esencial para todas las funciones biológicas. Las instrucciones necesarias para la expresión de los genes están codificadas en un conjunto de *cis*-elementos o módulos (CRM), en la región *upstream* de un gen, también conocido como la región intergénica (IGR) o simplemente promotor. Un CRM consiste en un conjunto de factores de transcripción (TF) que trabajan juntos [7] e interactuan con la ARN polimerasa (ARN pol) para conferir ciertos patrones espaciales y temporales de la transcripción. Por lo general, un CRM oscila entre unos pocos cientos a unos pocos miles de pares de bases, varios factores de transcripción se unen a ella, y cada uno de estos factores de transcripción pueden tener múltiples sitios de unión (TFBS). Además, pueden estar influidos por la presencia de elementos repetitivos u otros elementos cercanos dentro de la misma IGR como ARN no codificantes no anotados en bacterias. Un estudio reciente ha confirmado la importancia de las funciones de los CRM, y reveló cómo los cambios sutiles de la disposición original de los elementos del módulo puede afectar su función.

Por lo tanto, la identificación de CRM es, actualmente, un área activa de investigación en el análisis computacional de secuencias reguladoras [2].

Hoy en día hay muchos métodos disponibles para el análisis de las funciones reguladoras de promotores que predicen TFBS (TRANSFAC [12][16], JASPAR [13], RegTrans [11]), sitios de unión ribosomal (RBS) o sitios de reconocimiento de ARN polimerasa [5]. Por otra parte, las nuevas tecnologías de secuenciación (next generation sequencing (NGS)), tales como ChIP-Seq, que combina la inmuno-precipitación de cromatina (ChIP) con la secuenciación masiva y paralela de ADN, es capaz de asignar con precisión sitios de unión para cualquier proteína de interés. En cuanto a otros *cis* y *trans* elementos reguladores que pudieran interactuar con la ARN polimerasa y los TFBS, se encuentran los pequeños ARNs en bacterias y riboswitches, que se predicen con un alto número de falsos positivos, lo que hace difícil la selección de muestras para pruebas de laboratorio. En definitiva, esta gran cantidad de datos junto con el creciente número de genomas disponibles hace necesario un filtro a través de la generación de hipótesis a fin de hacer viables las pruebas experimentales.

Nuestro buscador de perfiles de promotores es capaz de identificar conjuntos de IGRs descritos por un conjunto de características comunes, por lo tanto la identificación de genes que son controlados por su expresión en las mismas condiciones. Sin embargo, el problema es difícil, ya que las interacciones combinatorias entre los factores que regulan pueden ser muy complejas, y también porque la gran mayoría de las relaciones espacio-temporales entre los factores de transcripción y otras características se desconocen.

Aquí les presentamos una herramienta que identifica CRMs dado un conjunto de sitios de unión de factores de transcripción y otras características para un grupo de genes. Este proyecto tiene en cuenta las limitaciones espaciales en la disposición de *cis*-elementos como la posi-

ción relativa, distancia y orientación, y hace frente a más de un sitio de unión para un determinado TF en la región promotora. El servidor web infiere estos módulos utilizando técnicas multi-objetivo y multi-modal [8]. Nuestro algoritmo es capaz de identificar no sólo las soluciones más obvias (es decir, las más frecuentes), sino también aquellas que no son localmente dominadas y que representan a los menos frecuentes, pero relevantes.

2 METODOLOGÍA

2.1. DEFINICIONES

El aprendizaje por reglas de asociación [1] es un método muy popular y bien documentado para descubrir relaciones interesantes entre las variables en grandes bases de datos. Esta tarea se ha llevado a cabo mediante una modificación del algoritmo Apriori [4] sobre la implementación de Eclat desarrollada por C. Borgelt [3], donde los conjuntos de elementos candidatos se cuentan en una sola pasada del algoritmo y no se generan sobre la marcha. La implementación utilizada se basa en la idea de organizar los itemsets (conjuntos de características) en un tipo especial de árbol de prefijo, que no sólo nos permite almacenar de manera eficiente haciendo poco uso de memoria, sino también es compatible con el procesamiento de transacciones, así como la generación de reglas.

Para obtener el mejor conjunto de CRMs, utilizaremos un enfoque evolutivo multi-objetivo: el algoritmo NSGA-II. Este algoritmo ha demostrado ser uno de los algoritmos más eficientes para la optimización multi-objetivo en una serie de problemas de referencia [9]. A todos los individuos no dominados por cualquier otra solución, se le asigna el número de frente de Pareto 1. Todas las soluciones que son sólo dominadas por los individuos en el frente número 1 se incluirán en el frente número 2 y así sucesivamente. La selección se realiza utilizando un torneo entre dos soluciones. La solución con el menor número de frente se selecciona si las dos soluciones son de diferentes frentes. La solución con la distancia más alta se selecciona si son del mismo frente, es decir, un mayor fitness se asigna a soluciones situadas en una parte escasamente poblada del frente. De N padres se generan N nuevas soluciones (hijos) en cada iteración. Ambos padres e hijos compiten entre sí para su inclusión en la siguiente iteración.

2.2. ALGORITMO

La metodología desarrollada sigue una serie de pasos que se detallan a continuación. El primer paso es la generación de reglas de asociación, lo que implica el descubrimiento de los iremsets que aparecen con frecuencia juntos en el conjunto de cis-elementos en el archivo de entrada. El método Apriori se puede interpretar como un problema de búsqueda con una única función objetivo: la cantidad de elementos que pertenecen a un itemset o *soporte*. Con el fin de hacer frente al problema multi-objetivo, hemos introducido un nuevo objetivo a maximizar: la *complejidad*. El soporte es dado por el número de genes que tienen los mismos cis-elementos, mientras que la complejidad se calcula por el número de diferentes cis-elementos pertenecientes al conjunto. Extraemos los conjuntos de elementos de una base de datos de transacciones que maximizan tanto el soporte como la complejidad mediante un post-procesamiento de la información generado por el algoritmo Apriori. A medida que el número de características y genes (transacciones) crece, el número de posibles soluciones crece exponencialmente. Pasado cierto número de características y combinaciones de genes, la aplicación Eclat es incapaz de producir resultados y se detiene por un problema de memoria.

Para estos casos, hemos desarrollado un método heurístico. Utiliza un enfoque evolutivo para el problema de encontrar los mejores conjuntos de elementos en una base de datos de transacciones. Se propone un algoritmo genético (GA) con un enfoque multi-objetivo, basado en el algoritmo NSGA-II. El AG propuesto instancia el algoritmo NSGA-II utilizando un cromosoma sencillo compuesto por una lista de las estructuras de módulos (por ejemplo, <(AP2A, 0,3, D, 0), (SP1, 0.6, D, 1), (SP1, 0.9, D , 110), (SP1, 0.5, D, 25), (SP1, 0.7, D, 32)>). Una estructura de módulo se compone de cinco elementos: nombre, puntuación, orientación, patrón y distancia (por ejemplo, el nombre de: CAAT, puntuación: 0,7, orientación: D, distancia: 30: (CAAT, 0.76, D, 30)). Algunas de las características de los cromosomas no se pueden utilizar en algunos casos, por ejemplo, cuando se ignora la hebra de la cadena, entonces la orientación también será ignorada. Se utiliza un sistema de cruce de un punto, es decir, los cromosomas de los padres se cortan en algún punto elegido al azar (entre las estructuras de módulo) y se produce el consiguiente intercambio de subcromosomas. Diferentes operadores de mutación se han utilizado: añadir, borrar y modificar. Un operador de mutación *añadir* simplemente extiende el cromosoma mediante la adición de una nueva característica al principio o al final de los cromosomas. El operador de mutación *borrar* selecciona cualquier módulo de la estructura del cromosoma y lo elimina. Por último, un operador de mutación *modificar* selecciona una de las características de la estructura del módulo y la cambia. Si la característica *nombre* es elegida, entonces un nombre al azar entre el conjunto de nombres válidos es utilizado para sustituirlo en la estructura del módulo; si la característica *orientación* es elegida, entonces se cambia la orientación de D a R y viceversa. La característica de *puntuación* no se utiliza en el proceso de evaluación, por lo tanto no es utilizado por los operadores de mutación. Finalmente, si la característica distancia es elegida, un pequeño número entero se suma o resta a los valores existentes. Una vez más, las características hebra y distancia sólo se utilizan cuando el usuario los haya seleccionado en los parámetros originales.

3 LA HERRAMIENTA

La herramienta creada con esta metodología se ha aplicado en un servidor web: http://gps-tools2.its.yale.edu/modulos/modules.html.

3.1. ENTRADA

A pesar de que hay un gran número de algoritmos desarrollados para descubrir motivos de ADN, no hay estándares para medir su validez. La mayoría de los algoritmos tienen un mejor desempeño en organismos inferiores, incluyendo la levadura, en comparación con otros organismos superiores [6]. Por esta razón, se deja al usuario la selección de uno u otro algoritmo para descubrir motivos de ADN o del uso de bases de datos disponibles. Estos algoritmos pueden utilizar la enumeración combinatoria, modelos probabilísticos (Stormo, Gibbs, AlineACE, ANN-Spec, MEME), programación matemática, redes neuronales y/o los algoritmos genéticos (EC, GAME), mientras que las bases de datos pueden ser TRANSFAC, REGULON DB, JASPAR, etc, así como los programas o fuentes para extraer información sobre elementos repetitivos o riboswitches en la misma región.

El servidor requiere como entrada un archivo, proporcionado por el usuario en formato csv (separado por comas). Cada registro del archivo de entrada contiene la información sobre las características presentes en los promotores. El archivo contiene los siguientes campos:

- Nombre de la secuencia: identificador de región promotora (por ejemplo, E_coli_K12_flhDC).
- Tipo de característica: a elección, por ejemplo, "dnapat" para los patrones de ADN.
- Nombre de la característica: cis-elemento identificador (por ejemplo, rpod18).
- Hebra: un carácter 'D' para la hebra directa o 'R' para la hebra inversa.
- Inicio de la característica: la posición inicial del cis-elemento.
- Final de la característica: la última posición del cis-elemento.
- Patrón: patrón de ADN (por ejemplo, TTGACA). Opcional.
- Puntuación: puntuación otorgada por el programa de predicción de cis-elementos o por el curador (por ejemplo, 0.652860653). Opcional.

Como parámetros de entrada el usuario puede seleccionar que las restricciones a aplicar al proceso de búsqueda de los perfiles promotores. Es importante mencionar que la búsqueda de un módulo de cis-elementos se compone de la búsqueda de dos propiedades: un conjunto de señales, y las relaciones espacio-temporales (restricciones) entre este conjunto de señales. Por desgracia, a excepción de un pequeño número de objetivos específicos, bien caracterizados, las interacciones, la gran mayoría de las relaciones espacio-temporales (restricciones) entre los factores de transcripción son desconocidos. Con el fin de seleccionar el tipo de restricción que podrían tener un impacto en la expresión de una transcripción nos centramos en el grupo de trabajo anterior sobre los motivos de ARN polimerasa [5], la predicción de pequeños ARNs en bacterias [14] y la co-expresión y ubicación de los factores de transcripción en bacterias y humanos [15, 17]. Estas publicaciones plantear la importancia de las distribuciones de distancia (cerca, medio, y lejano) entre la ARN polimerasa y los sitios de unión de factores de transcripción en la activación y represión de los promotores. Por tanto, las restricciones seleccionadas fueron:

- Orden: El orden tiene en cuenta no sólo la presencia de una característica, sino también el orden relativo de aparición de todas las características dadas en un promotor.
- Distancia: Si el usuario ha pedido que se tenga en cuenta la distancia entre las características dentro de los módulos, entonces se pueden especificar esta distancia como estricta o difusa. Al utilizar distancias estrictas se busca que todas las características estén a una misma distancia entre ellas. Si se piden distancias difusas entonces se buscaran características que estén a una distancia de + o - un cierto valor dado. Las distancias difusas pueden ir en el rango entre 1 y 50 con un valor predeterminado de 1.
- Orientación: si se especifica, entonces se diferencias grupos de características en una u otra hebra relativa al marco de lectura abierto.
- Número máximo de elementos: es el número máximo de diferentes cis-elementos (por ejemplo, TFBS, sitios ARN pol) por módulo que desea buscar. Este valor debe pertenecer al intervalo [1-100] con un valor predeterminado de 5. Cuanto más alto sea el valor, mas lenta será la ejecución.

Hay dos versiones disponibles: una exhaustiva, para archivos de entrada pequeños (bacterias), y una heurística para grandes conjuntos de datos (eucariotas). En caso de seleccionar la versión heurística hay algunas opciones adicionales:

- Tamaño de la población: Los algoritmos evolutivos se basan en una población de representaciones abstractas (denominadas cromosomas) de las soluciones candidatas (llamados individuos) a un problema de optimización, y evolucionar hacia mejores soluciones. Mayor tamaño de la población se traducirá en un rendimiento más lento, pero en mejores resultados. A medida que el tamaño de la población aumenta, el número de evaluaciones realizadas también debe aumentar. El usuario puede cambiar el tamaño de la población en el rango [10-1000]. Este campo es obligatorio y el valor por defecto es de 200.

- Número de evaluaciones: por lo general, una población inicial de soluciones candidatas generadas aleatoriamente comprende la primera generación. Durante cada generación sucesiva, una proporción de la población existente está seleccionada para crear una nueva generación. Una función de fitness se aplica a las soluciones candidatas y los descendientes posteriores para cuantificar la optimización de una solución (es decir, un cromosoma) en un EA para que ese cromosoma en particular puede ser clasificado en contra de todos los otros cromosomas. Cada una de estas evaluaciones de la función de costes se puede utilizar para determinar cuando parar una ejecución EA. El usuario puede especificar el número máximo de evaluaciones de la EA, donde los valores oscilan entre 1 y 99999. Este campo es obligatorio y el valor por defecto es 200. Como una regla sencilla, el número de evaluaciones debe ser un múltiplo del tamaño de la población. Este número será aproximadamente el número de generaciones para llevar a cabo.

3.2. SALIDA

La salida consta de una salida intermedia con dos enlaces correspondientes a dos archivos `features.txt` y `modules.txt`. El usuario puede guardarlos en su disco local para la inspección visual de los promotores individuales o los módulos que obtenga. El archivo cuenta con cuenta con las once columnas del archivo de entrada, más:

- ColorRGB: un color específico para cada tipo de característica.
- Comienzo de la IGR: la posición de inicio de la región intergénica.
- Final de la IGR: la posición final de la región intergénica.

El archivo `modules.txt` incluye la siguiente información para cada módulo:

- Nombre del módulo: identificador de módulo (por ejemplo, Module1).
- Elementos: los identificadores de región promotora de estos elementos en el módulo.
- Tipo de programas: a elección, por ejemplo, "dnapat" para los patrones de ADN.
- Nombre de la función: *cis*-elemento (por ejemplo, rpod18).
- Hebra: un carácter 'D' para la hebra directa o 'R' para la hebra reversa.
- Comienzo de la característica: la posición inicial del *cis*-elemento.
- Final de la característica: la última posición del *cis*-elemento.

```
#Date:Thu Jun 16 01:32:36 CDT 2011 #In-
put:/tmp/eDDEqMyKZN.input #Modu-
les:presence_of_features #Co-
lumns:Module_name;Feature_type;Feature_name;Stran
d;Start;End START #Module#:1 Ele-
ments:Y_pestis_CO92;Y_pestis_KIM_flhDC Module1
          dnapat    3_ygix    D         1
          15 Module1          dnapat    CpxR      D
Module2   dnapat    PhoP_3    D         95
          112 Module2         dnapat    rpod18    D
Module3   dnapat    rpod18    D         145
          151 Module3 dnapat    SoxS      D
```

Figura 1: Ejemplo de salida del archivo `module.txt`, cada módulo contiene el nombre de la región promotora/IGR que pertenecen a la misma. En las siguientes líneas se especifican las características comunes.

El panel gráfico contienen una descripción detallada de todas los *cis*-elementos en un CRM. El usuario puede explorar por separado cada módulo. Todas las vistas se pueden guardar como un gráfico o en texto en diferentes formatos como: txt, csv, png, html, jpg. También se puede hacer zoom hasta el nivel de pares de bases.

Hay varios tutoriales disponibles junto con archivos de prueba como ejemplo. Los tutoriales explican en detalle los parámetros que se pueden modificar y cubren los siguientes temas de ayuda: entrada de datos, máximo número de características para ser considerados en un módulo, la distancia, la orientación, etc. Los resultados también se pueden recibir por correo electrónico, donde en caso de error, se muestra un mensaje legible.

4 RESULTADOS

El enfoque aquí descrito ha sido utilizado con éxito para identificar a los promotores regulados por PhoP que tienen más de un sitio de unión para la PhoP TF y compartir una orientación atípica y distancias entre de la caja PhoP/ sitio de la ARN polimerasa (Figura 2) [5], lo que ha llevado a inferir el control de la transcripción de PhoP sobre genes *acid resistance* en *Salmonella typhimurium* [10, 17]. Recientemente se ha demostrado que estos módulos están relacionados con diferencias en la velocidad de la expresión (submitted, Zwir et al., 2011).

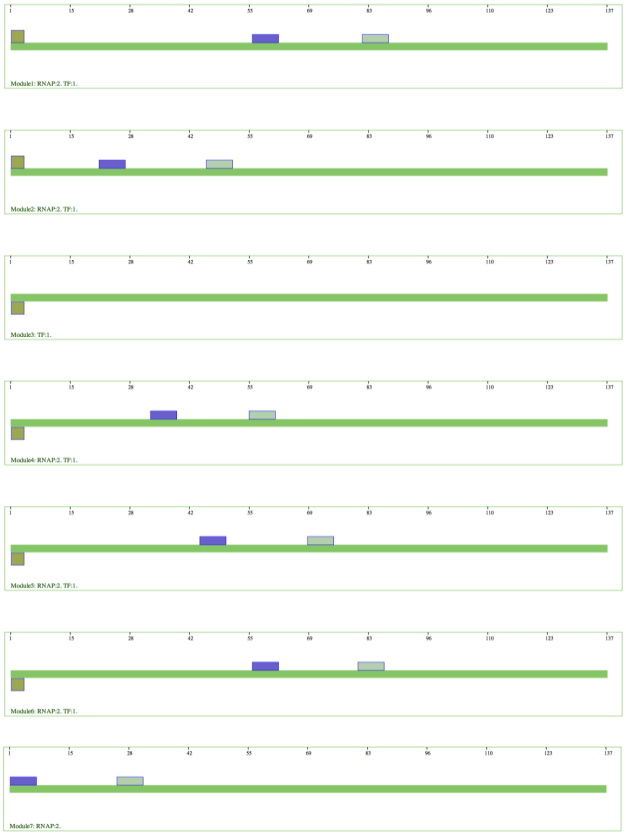


Figura 2. Ejemplo de la salida gráfica.

5 DISCUSIÓN

El desarrollo de la herramienta que aquí se presenta ha sido impulsada por los usuarios desde el principio. El enfoque aquí descrito ha sido utilizado con éxito para identificar a los promotores regulados PhoP que tienen más de un sitio de unión para la PhoP TF y comparten una orientación atípica y distancias de la caja PhoP / sitio de la ARN polimerasa. Esto ha permitido inferir el control de la transcripción de PhoP sobre los genes *acid resistance* en *Salmonella typhimurium*. El servidor ha estado activo durante los últimos 2 años y el número de secuencias utilizadas por los usuarios de variado mucho desde 10 hasta 500.

Existe una serie de herramientas para la búsqueda de CRMs específicos basado en un conjunto de genes co-regulados, pero se centran en el descubrimiento de nuevas señales y sitios de unión. Sin embargo, nuestro enfoque sólo predice CRMs de regiones para las que el usuario ha proporcionado un conjunto definido de señales. Sólo hay un par de métodos de evaluación para CRMs ya que es muy difícil de realizar porque ha muy pocas regiones genómicas donde estamos seguros de todos los elementos de regulación. Por lo tanto, es difícil estimar con precisión la tasa de falsos positivos.

Uno de nuestros puntos fuertes es la posibilidad de contar con restricciones de control, mientras que en las herramientas de cualquier otro cambio sutil en la combinación o el orden de las señales siempre pueden dar resultados diferentes. La otra fuerza es que nuestro algoritmo es capaz de identificar no sólo las soluciones más obvias (es decir, las más frecuentes), sino también aquellas que no son dominadas localmente y que representan a las menos frecuentes, pero relevantes.

Las restricciones seleccionadas son un conjunto muy reducido debido al hecho de que la gran mayoría de las relaciones espacio-temporales entre los factores de transcripción son desconocidas. La información de la distancia, la orientación y la conservación, están lejos de ser suficiente para identificar un módulo funcional. Para responder a esta petición, el análisis de nuevas estrategias y métodos de predicción que integran la información de la secuencia y la firma de la cromatina puede ser un importante paso adelante.

Agradecimientos

Agracedemos al Ministerio español de Educación y Ciencia (TIN2009-13950) y a la Junta de Andalucía (TIC-02788), por su financiación. R. Romero-Zaliz también quiere agradecer al programa GENIL (PYR-2010-28).

Referencias

[1] Agrawal, R. et al. 1993. Mining Association Rules Between Sets of Items in Large Databases. *ACM SIGMOD Record* (1993), 207–216.

[2] Berman, B.P. et al. 2002. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the Drosophila genome. *Proceedings of the National Academy of Sciences of the United States of America*. 99, 2 (2002), 757.

[3] Borgelt, C. Efficient Implementations of Apriori and Eclat.

[4] Borgelt, C. and Kruse, R. 2002. Induction of association rules: Apriori implementation. *Compstat: Proceedings in Computational Statistics: 15th Symposium Held in Berlin, Germany* (2002), 395.

[5] Cotik, V. et al. 2005. A hybrid promoter analysis methodology for prokaryotic genomes. *Fuzzy Sets and Systems*. 152, 1 (May. 2005), 83-102.

[6] Das, M.K. and Dai, H.-K. 2007. A survey of DNA motif finding algorithms. *BMC bioinformatics*. 8 Suppl 7, (Jan. 2007), S21.

[7] Davidson, E.H. 2001. *Genomic Regulatory Systems: Development and Evolution*. Academic Press, San Diegelements. Nat. Genet.

[8] Deb, K. 2001. *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons.

[9] Deb, K. et al. 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*. 6, 2 (2002), 182-197.

[10] Harari, O. et al. 2009. Identifying promoter features of co-regulated genes with similar network motifs. *BMC bioinformatics*. 10 Suppl 4, (Jan. 2009), S1.

[11] Kazakov, A.E. et al. 2007. RegTransBase--a database of regulatory sequences and interactions in a wide range of prokaryotic genomes. *Nucleic acids research*. 35, Database issue (Jan. 2007), D407-12.

[12] Markstein, M. and Levine, M. 2002. Decoding cis-regulatory DNAs in the Drosophila genome. *Current opinion in genetics & development*. 12, 5 (Oct. 2002), 601–606.

[13] Sandelin, A. et al. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic acids research*. 32, Database issue (Jan. 2004), D91-4.

[14] del Val, C. et al. 2007. Identification of differentially expressed small non-coding RNAs in the legume endosymbiont Sinorhizobium meliloti by comparative genomics. *Molecular microbiology*. 66, 5 (Dec. 2007), 1080-91.

[15] Val, C. et al. 2009. PromoterSweep: a tool for identification of transcription factor binding sites. *Theoretical Chemistry Accounts*. 125, 3-6 (Oct. 2009), 583-591.

[16] Wingender, E. et al. 2000. TRANSFAC: an integrated system for gene expression regulation. *Nucleic acids research*. 28, 1 (Jan. 2000), 316-9.

[17] Zwir, I. et al. 2005. Dissecting the PhoP regulatory network of Escherichia coli and Salmonella enterica. *Proceedings of the National Academy of Sciences of the United States of America*. 102, 8 (Feb. 2005), 2862-7.

A time series gene expression data analysis pipeline with a fuzzy clustering method to assess the human adenovirus infection process

P. Yankilevich[1], P. R. Barrero[2], I. Zwir[3]

[1] Bio Sidus S.A., Constitución 4234 (C1254ABX) Buenos Aires, Argentina.
[2] Virology Laboratory Dr. Ricardo Gutiérrez Children's Hospital, Buenos Aires, Argentina.
[3] Department of Molecular Biology, Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, MO,USA and with the Department Computer Science and Artificial Intelligence University of Granada, Spain.

Abstract **— Technical improvements in high-throughput gene expression experiments are making possible to obtain high quality time series whole-genome expression data sets. This valuable source of information can describe the unfolding biological processes during the development stages, the cell cycle or the immune response of an organism. In order to fully explore this type of data we developed an integrated time series gene expression analysis pipeline. The resulting method detects differentially expressed genes, cluster co-expressed genes, unveil hidden gene expression patterns, identify over represented biological function categories and infer gene regulatory networks. Some of the methods integrated in our pipeline are an empirical Bayes model, a noise robust fuzzy clustering and graphical Gaussian model. The use of this pipeline to analyze the human adenovirus infection process allowed us to discover new insights and hypothesis. No previous exhausted explorations including features as fuzzy clustering or regulatory network inference have been used on this biological phenomena data before.**

I. INTRODUCTION

Within the genetics field there is a growing interest in using time series gene expression experiments to study different biological systems and to describe and infer the interactions among gene products [1, 2, 3]. Gene expression is a temporal process that synthesizes different proteins with different functions according the organism requirements. By measuring this expression activity we are able to identify the genes which are express or repress under a particular condition and study its interdependency structure. We are particularly interested on the analysis and elucidation of the genetic expression programs launched by the cells when confronted with new conditions such stress, treatment or infection [4].

Most time series expression experiments are still been analyzed with methods developed originally for static data where a single snapshot of the gene expression in different samples is measured [5]. Another important difference between these two types of data is that while static data from a sample population are assumed to be independent identically distributed, time series data exhibit a strong autocorrelation between successive points so it have to be analyze differently [1]. Even though several new algorithms specifically targeting this type of temporal data were presented recently, it is extremely important to approach the analysis in an integrated manner in order to fully comprehend the whole picture of the biological system under study. The best known integral solutions are Ingenuity® [6] and PathwayStudio® [7] but are not freely available. In this paper we present a set of open source methods for gene expression data analysis implemented as an integrated pipeline to analyze time series expression experiments and infer gene association networks. The pipeline identifies differentially expressed genes by an empirical Bayes model and the most characteristic expression patterns by a noise robust fuzzy clustering method. Functional annotations of the differentially expressed genes are obtained and regulation network of these genes is also inferred based on a graphical Gaussian model. These are essential computational tasks in gene expression data analysis to better understand how a phenomenon such as stress, a drug or a virus may be affecting an organism.

We applied the proposed pipeline to analyze a time series human gene expression of the cells infection process by adenovirus. This was an *in vitro* experiment performed at the Virology Laboratory in the Dr. Ricardo Gutiérrez Children's Hospital, Buenos Aires, Argentina. The experiment is MIAME compliant [8] and its detailed analysis is being submitted elsewhere [9] (Barrero *et al.*, manuscript in preparation).

II. MATERIALS AND METHODS

The pipeline here presented was completely developed as an R computing environment and statistical programming language [10] script using different statistical and graphical packages included in Bioconductor [11]. Bioconductor is widely used open source and open development software project for the analysis and comprehension of genomic data.

Methodology pipeline implementation scheme

Starting from genetic expression data of an organism as an input, we have constructed an analytical methodology integrating Bioconductor packages as the building blocks of our pipeline to study its underling biological phenomenon. The interconnection of the proposed working packages implemented as an R script is presented in Figure 1.

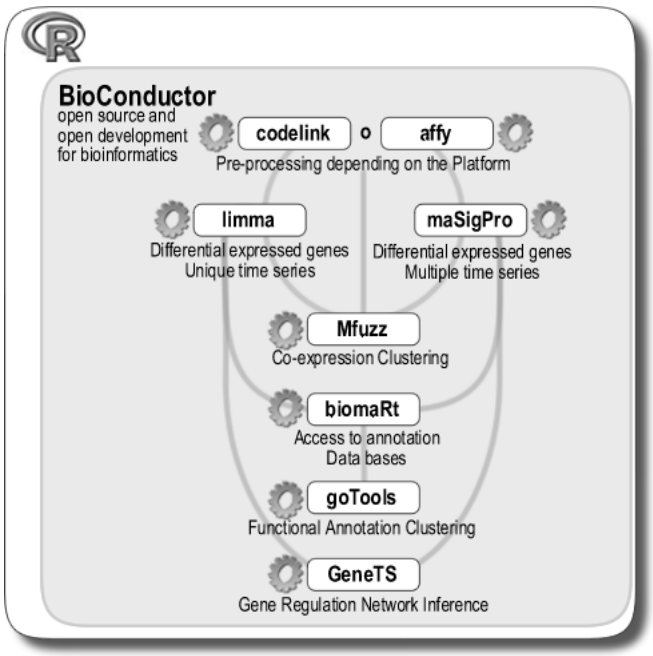


Fig. 1. Scheme of the Bioconductor packages interconnection that builds up the Time Series Gene Expression Analysis Pipeline.

A. Experimental design for the adenovirus infection process on human cells

Human cultured cells (Hep-2) were infected with human adenovirus type 7, and RNA samples were collected at 0, 1, 2, 8 y 24 hours post-infection (hpi). These samples were amplified, labeled and hybridized to CodeLink UniSet™ Human 20K microarrays which were then scanned to finally obtain the gene expression values. A pre-analysis of the expression patterns performed with the fuzzy clustering method showed the samples taken at 1 hpi behave very different the rest (Fig. 3), resulting in the design of the following comparisons to identify the most relevant genes:

- Early differentially expressed genes during infection
  1 hpi – 0 hpi
- Late differentially expressed genes during infection
  (2 hpi + 8 hpi + 24 hpi) / 3 – 0 hpi

B. Identification of differentially expressed genes

Raw expression data was imported into R and then preprocess to filter spots flagged as bad, performed background correction and normalized. Data normalization is required before they can be compared due to small differences in RNA quantities of the samples and fluctuations generated by the technique; the intensity levels may vary from one replicate to the other due to effects which are unrelated to the genes. An empirical Bayes moderated t-test was adjusted to the data in order to identify differentially expressed genes. The empirical Bayes procedure provides an effective framework for studying the relative changes in gene expression for a large number of genes. It uses a simple nonparametric mixture prior to model the population of affected and unaffected genes, thereby avoiding parametric assumptions about gene expression [12]. Probabilities of being differentially expressed (P values) associated with the t-test were adjusted due to false discovery rate (fdr) by multiple testing using the Benjamini and Hochberg method [13]. As a result we obtained the fold change and the P value of differential expression for each gene on every comparison of interest. The genes were then ranked in order of evidence for differential expression and those with an adjusted P value < 0.05 were considered differentially expressed. These methods are all included in the Limma (Linear Models for Microarray Data) package [14] included in Bioconductor.

C. Fuzzy co-expression and Functional clustering

Most clustering algorithms produce hard partitions of the data where each gene is assigned to a single cluster. However, this is not the case for time series gene expression data, where gene expression clusters frequently overlap. This limitation is overcome using a fuzzy unsupervised noise-robust soft clustering method based on co-expression, which correctly assess the expression patterns hidden on the time series data. The method included in the Mfuzz package [15] implements a fuzzy c-means algorithm based on the minimization of a weighted square error function [16] that allows the assignation of a gene to multiple clusters and indicates how well the corresponding clusters represent its genes. Additionally, this soft clustering method avoids an *a priori* pre-filtering of genes preventing the exclusion of biologically relevant genes from the data analysis. This makes it more noise robust than hard clustering algorithms, which are often highly sensitive to noise. As this flexible method let us obtain a membership value of each gene to its belonging cluster and vice versa, by defining different cluster minimum membership m values and analyzing the structure of the clusters obtained we are able to find the unstable or not significant clusters and ignore them until we unveil the appropriate number of clusters that best represent the expression patterns unfold by the biological process.

Previous functional annotation knowledge associated with each gene was integrated into the analysis using the biomaRt package [17] which access via web to different public annotation databases. This information allowed us to perform a functional annotation clustering based on the Gene Ontology (GO) using the goTools package and compare the functional categories overrepresented by the genes differentially expressed at different stages of the biological phenomenon under study.

D. Inference of gene regulatory networks

Finally, the inference of the gene association networks of the most relevant genes was performed with the GeneTS package (now called GeneNet) [18]. This package implements a Graphical Gaussian Model (GGM), where the genes (nodes) are related according they dynamic partial

correlation coefficient. The method apply bootstrap to obtain better estimates while working with small-sample experiments and use an empirical Bayesian fdr model, as used to identified differential expressed genes, to identify the partial correlations (edges) which are significantly different from zero. This method also deals with irregular sampling, unequal temporal spacing of time points and more importantly, its estimations takes the small-sample problem into account.

III. MAIN RESULTS

By applying this pipeline we were able to identify most of the human genes which are affected by the adenovirus infection process. The vast majority of the differentially expressed genes (over 5000 genes) obtained with Limma package were early expressed at 1 hpi while very few genes showed that variation at the late stages of the infection (2, 8 and 24 hpi). These findings were also confirmed by the Mfuzz soft clustering results, showing that the most populated expression patterns were the one with an up-regulated peak at 1hpi and the one with a down-regulated peak at 1hpi (Fig. 2).

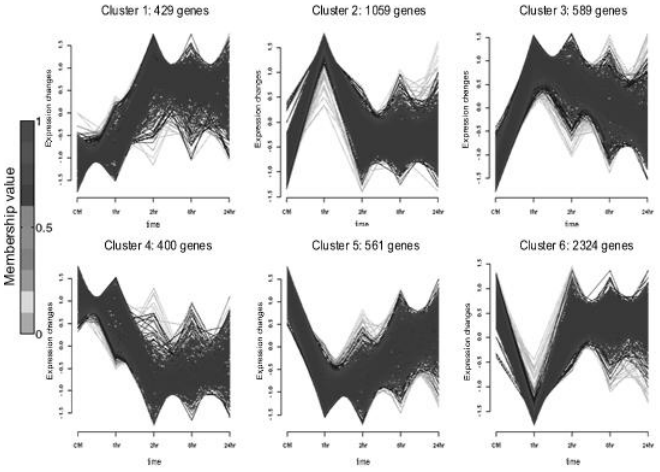


Fig. 2. Differential expression patterns obtained by the soft clustering for human adenovirus infection process.

The functional annotation clustering on the differential expressed genes allowed us to obtain the over-expressed functional categories resulting from the infection process. After discriminating among early and late differentially expressed genes we obtained and compared the over-expressed functional categories from each group with the biomaRt and goTools packages. Figure 3 shows one of these comparisons including the functional categories related to immune response only. These are very useful graphs to help the virologists understand the mechanisms of infection.

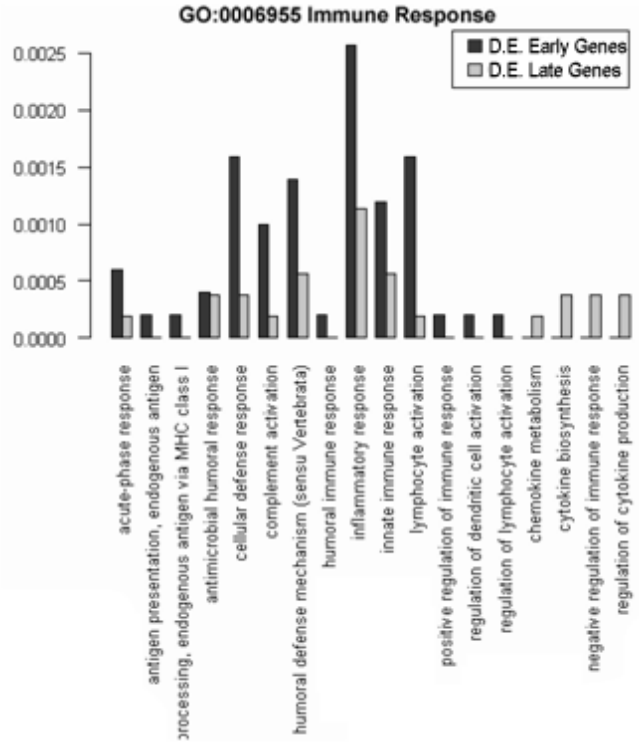


Fig. 3. Early vs. Late Differentially Expressed (D. E.) genes distribution over the Immune Response GO term classes.

The final stage of the pipeline allows us to infer different gene regulation networks associated to the adenovirus infection process. For example, figure 4 is a sub-network of the most significant connections among the human early differentially expressed genes during the infection process.

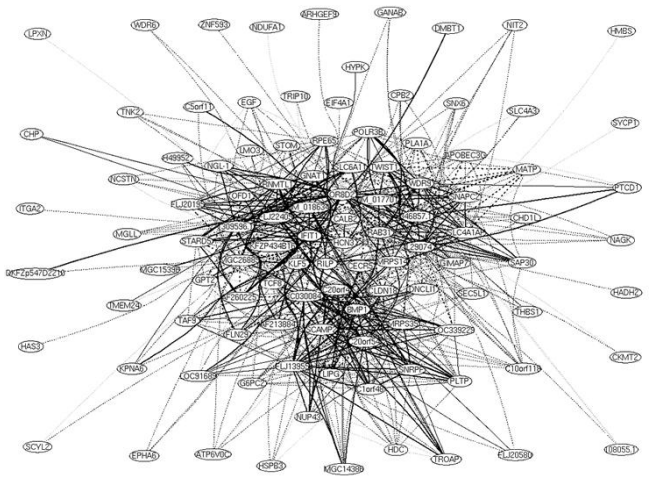


Fig. 4. Gene regulation sub-network of the early differential expressed genes.

IV. CONCLUSION

Motivated by the lack of freely available integral tools for the analysis of temporal gene expression experiments, this paper proposes an integrated time series gene expression analysis pipeline implemented with widely used open source

and open development software projects for the analysis of high-throughput experiments such as Bioconductor and R.

Analysis of temporal expression data with a single tool that integrates multiple state of the art methods for time series analysis really allow us a better understanding of the whole picture of a biological system under study. From the methods in our pipeline the fuzzy clustering is a key building block to understand the expression patterns of the differentially expressed genes, as well as to make sense of the gene regulation network inferred from them.

The new insights discovered with this method on the adenovirus infection process on human cells provide new hypothesis not implicated in previous studies on the possible action of the virus.

This pipeline offers a unique tool to examine gene expression time series data uncovering it most valuable features and integrating data as a whole for a better and more comprehensive exploration and visualization. It also provides a flexible and free open source platform that can be used with multiple data sources and acts as a guided tour into the vast and difficult task of gene expression data analysis.

ACKNOWLEDGMENT

The authors wish to thank Alicia Mistchenko from Virology Laboratory, Dr. Ricardo Gutiérrez Children's Hospital and personal from BioSidus S.A. for their helpful discussion in this study.

REFERENCES

[1] Bar-Joseph Z. **Analyzing time series gene expression data** Bioinformatics. 2004 Nov 1;20(16):2493-503.

[2] Gregory W. Carter **Inferring network interactions within a cell** Brief Bioinform 2005 6: 380-389

[3] Trey Ideker, Owen Ozier, Benno Schwikowski, and Andrew F. Siegel **Discovering regulatory and signalling circuits in molecular interaction networks** Bioinformatics 2002 18: S233-S240

[4] Nau GJ, Richmond JF, Schlesinger A, Jennings EG, Lander ES, Young RA. **Human macrophage activation programs induced by bacterial pathogens**. Proc Natl Acad Sci U S A. 2002 Feb 5;99(3):1503-8.

[5] Friedman,N., Linial,M., Nachman,I. and Peer,D. **Using Bayesian network to analyze expression data**. J. Comput. Biol. 2000, 7, 601–620.

[6] http://www.ingenuity.com

[7] http://www.ariadnegenomics.com

[8] Brazma, A., P. Hingamp, et al. **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data**. Nat Genet 2001 29(4): 365-71.

[9] Barrero PR, Yankilevich P, Diez RA, Mistchenko AS. **Bioarray analysis of differential cellular gene expression after adenovirus 7h infection**. Rev Argent Microbiol. 2006 (38):32.

[10] http://www.r-project.org

[11] Dudoit, S., Gentleman, R.C. & Quackenbush, J. **Open source software for the analysis of microarray** data. Biotechniques Suppl, 2003, 45-51.

[12] Efron, B., Tibshirani, R., Storey, J.D., and Tusher, V. **Empirical Bayes analysis of a microarray experiment.** Journal of the American Statistical Association, 2001, 96, 1151-1160.

[13] Hochberg, Y. & Benjamini, Y. **More powerful procedures for multiple significance testing**. Stat Med 1990, 9, 811-818.

[14] Gordon K Smyth. **Limma: linear models for microarray data**, pages 397–420. Springer, New York, 2005.

[15] Futschik ME, Carlisle B. **Noise-robust soft clustering of gene expression time-course data**. J Bioinform Comput Biol. 2005 Aug;3(4):965-88.

[16] Bezdak JC, **Pattern Recognition with Fuzzy Objective Function Algorithms**, Plenum Press, New York, 1981.

[17] Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., Huber, W. **BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis** Bioinformatics, 2005, 21, 3439–3440.

[18] Schäfer, J. and Strimmer, K. **An empirical Bayes approach to inferring large-scale gene association networks**, Bioinformatics, 2005 21, 754–764.

A graphical time series analysis tool to assess synchrony: Application on neuronal activity of Parkinson's disease

Patricio Yankilevich[1], Luis Riquelme[2] and Mario Gustavo Murer[2]

[1] *Departamento de Ciencias de la Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón I, Ciudad Universitaria (1428), Buenos Aires, Argentina.* Yankilevich@gmail.com

[2] *Departamento de Fisiología, Facultad de Medicina, Universidad de Buenos Aires, Paraguay 2155, Buenos Aires (1121), Argentina.*

Abstract

In order to better understand Parkinson disease, the electrophysiological changes of neurons in the Substantia Nigra Pars Reticulata (SNpr) on rat experimental models were characterized. Previous work by Tseng et al. demonstrated that periodic discharges occur in a synchronize manner in neurons affected by the disease. Motivated by a problem raised from neurophysiologists about the difficulty to systematically assess neuronal behavior and detect synchrony between firing patterns of different neurons affected by Parkinson disease, we developed a computational tool that makes a qualitative and quantitative assessment of the neural activity signals by integrating classical and novel graphical statistical methods corresponding to the non-linear time series analysis. The analyzed signals were recorded simultaneously in vivo from the SNpr on rat models induced to get Parkinson by 6-OHDA toxin. The graphical statistical methods implemented are: autocorrelation diagram, crossed correlation diagram, inter-spikes intervals (ISI) histogram, spike raster-grams and real time dynamic graphs to compare inter-spikes intervals. In addition, and innovating, we include the variability diagrams, which are a graphical method to represent the existing higher order local correlation among the successive values of a time series. The application of these methods on Parkinson neuronal time series detected an inhibitory synchrony among neurons, a negative correlation pattern where some neurons are inhibited while others exited. This inhibitory synchrony pattern was not clearly identified in previous studies. The resulting software is able to easily detect the existence of underlying patterns (determinism) and synchrony in the neuronal dynamics.

Keywords: Neurophysiology, Synchrony, Parkinson's disease, Neuronal time series.

1 INTRODUCTION

Although the physiological behaviour of a neuron is well understood, it remains unknown the different activities and patterns of interaction among them. There are 10^{11} neurons in the brain and each neuron can have a big number of connections, receiving information from about 1000 different cells. The neuron production of an action potential depends on the activity of hundreds of exciting and inhibiting synapses acting simultaneously on it. The central nervous system of humans and animals is known to suffer a progressive damage and loss of neurons along lifetime. Diseases like Parkinson or Alzheimer increase this loss and damage of neurons in certain specific areas of the human brain (Lozano and Lang 1998). The Parkinson neurological affection increase with the age; the progressive degeneration of the dopaminic neurons in the Substantia Nigra Pars Reticulata (SNpr) and the decrease of the dopamine levels are followed by the classical clinical signs of hyperkinesia, tremble in rest, stiffness, motor and posture dysfunctions. These symptoms are motional due to the relation of the SNpr with the motor systems (Hornykiewicz 1966). In order to better understand Parkinson disease we characterized the electrophysiological changes of neurons in the SNpr in rat experimental models induced to get Parkinson by 6-OHDA toxin. This assessment showed a high percentage (~40%) of the neurons were firing in bursts, where in animal control neurons, with spontaneous discharges, the bursting was detected only in up to 5% of the cells. On a later work, we found that the affected neurons showed a periodic activity on the bursting frequency, discharging approximately once per second. Furthermore, the electrophysiological recording analysis on pairs of neurons demonstrated that these periodic discharges occur in a synchronize manner in the neurons affected by the Parkinson disease (Tseng el al. 2001). In order to fulfilled the lack of a freely available robust computational tool to further analyze this signals, we developed the here

presented software that integrates classic and novel graphical statistical methods corresponding to the non-linear time series analysis to enhance our vision of the neuron discharges, reveal underlying activity patterns in its firing structure and to detect synchrony on pairs of neurons.

2 MATERIALS AND METHODS

2.1. DATA ACQUISITION

The firing process of the neuron, where a spike signal is created, is one of the most common processes present in nature and is called process of integrate and fire. Input information is accumulate by a non linear unit, and when this value reach a certain threshold, the unit 'discharge' emitting an action potential that is shown as a spike in the signal register.

In our analysis, the data is obtained with a measuring technique called extracellular single unit recording (in vivo). This technique is similar to an electroencephalogram but affecting a single neuron only. All the data signals were recorded inserting a micropipette at the SNpr from rats with induced Parkinson. The stored digital records of neural activity were first retrieved and then off-line analyzed. Every signal recording was made in pairs (on two different neurons) and simultaneously, inserting two micropipettes in different zones of the SNpr. Recordings last for approximately 3 minutes of duration. A detailed explanation of neuronal signal recording procedure can be found in Murer et al. 1997.

2.2. CONSTRUCTION OF THE NEURONAL TIME SERIES

Neural systems, as most of the other natural systems, are dynamic and evolve in time with a non linear behaviour. To analyze them, usually all we have is a set of measures of a single variable of the system that had been taken through time, named time series. In our case, the time series are made of measures of time intervals (milliseconds) between the discharges of the single neuron (events of high activity), as shown in Fig. 1. The time intervals are denoted r_i and represent the time passed in between neuron discharges. It is very important to note that the existent correlation in the resulting time series it corresponds with the real correlation of the neuronal process, thereby it can be interpreted in real time terms (Teich et al. 1997).

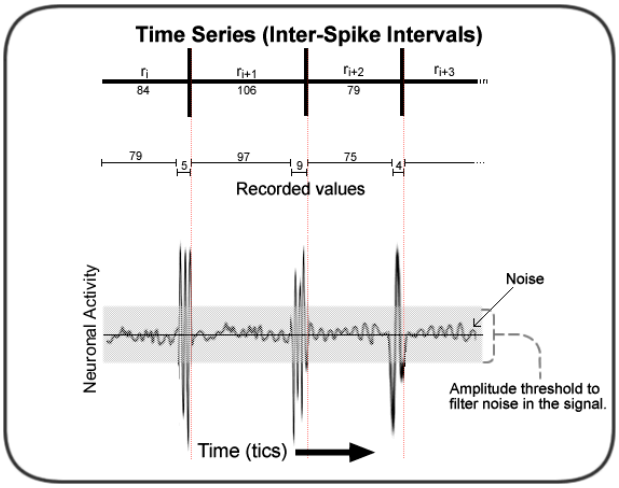


Figure 1: Time series creation process from a neuronal activity signal.

2.3. GRAPHICAL STATISTICAL METHODS IMPLEMENTED FOR TIME SERIES INSPECTION

Although there are several classification and analysis methods, the first thing to do with an unexplored data set is to observe it in many possible ways. The visual inspection can reveal symmetries or guide us to a better representation and interpretation (Kantz and Schreiber 1997). The following graphical statistical methods to study neuronal activity are included in the software. These methods were completely implemented in Java under the object oriented programming paradigm.

2.3.1. Spike raster-gram

The spike raster-gram graph let us have an idea of the activity state of the neuron under study by showing its firing pattern. The graph is made of a horizontal axis that represents time (the axis is fragmented in order to visualize a bigger time interval on the screen) and a set of perpendicular vertical lines, called spikes, that represent the firing of an impulse or action potential. Figure 2 shows the different neuronal activity states represented by a spike raster-gram together with its histogram and autocorrelogram which are later explained. As it was mention in the introduction, spontaneous discharges (regular firing pattern) is the observed in 95% of the neurons in the control specimens, while in specimens with Parkinson disease the states of excitation and inhibition (burst firing pattern) are present in 40% of its neurons.

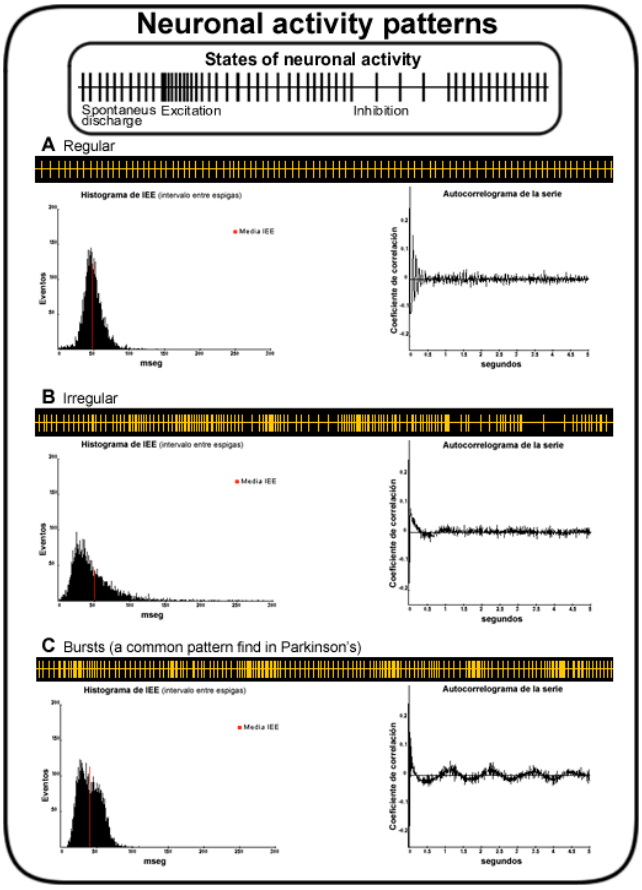


Figure 2: Most common neuronal activity patterns registered from the SNpr of experimental rat controls and with Parkinson. (A) Regular discharges are typical on control rat neurons, where the intervals of the between spikes distribution is almost symmetrical, (B) Irregular patterns are typical of injured rat neurons and (C) Burst firing pattern is common in neurons affected by Parkinson disease, where an oscillating frequency is clearly shown. This oscillation is somehow transferred to the motor activity, producing the classic tremor caused by the disease.

The spike raster-grams generated from a pair of neurons discharges simultaneously recorded to be compared are presented in Figure 3. Knowing that the amplitude of discharges of a neuron remains almost constant, the software obtains the amount of information that the neuron has transmitted by measuring its firing rate. The first 1000 milliseconds (a second) of the series are marked in red. The time series duration and the number of spikes are also displayed.

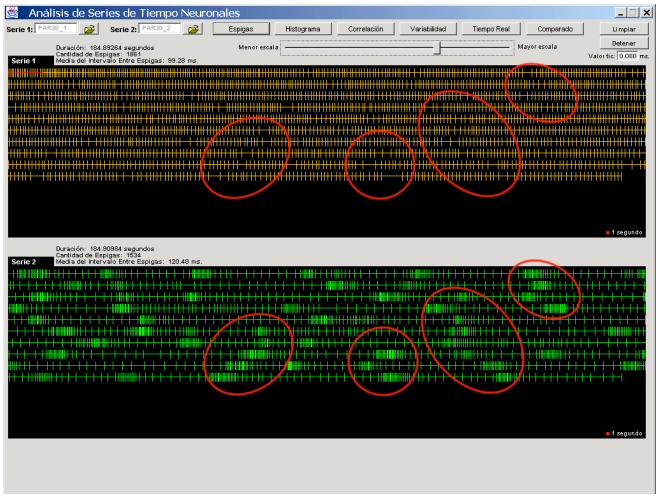


Figure 3: Spike raster-grams snapshot for a pair of neuronal time series. Both these recordings were taken simultaneously. This particular pair shows an inhibitory synchrony, a negative correlation pattern where one neuron is inhibited while the other excited. This type of synchronization is easily detected by this software.

2.3.2. Inter Spike Intervals Histogram

This statistical graph gives information about the way in which a neuron discharges. This method estimate the function of probability distribution of the inter spike interval values over the whole time series. The resulting histogram is a discrete function that represents the duration of each possible interval (in milliseconds) against the number of occurrences of that interval in the whole series, as shown in Figure 4. The histogram of a control neuron, with a regular firing rate, shows an almost Gaussian probability distribution; and its maximum (peak) indicates the mean interval. On the other hand, a neuron with a burst firing pattern usually has a two peak histogram, where one of the peaks represents the mean interval in between bursts and the other is the mean interval while burst discharging. In this case, the mean interval of the overall series is somewhere in between both peaks.

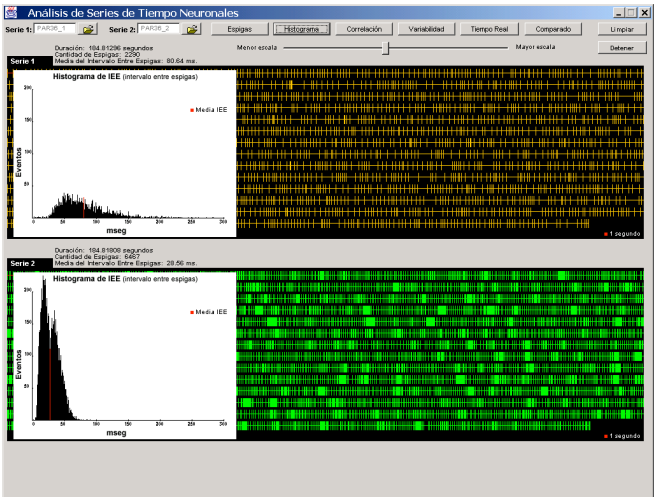


Figure 4: Inter spike intervals histograms snapshot. The histograms of this pair of neurons show different patterns of firing, the one above is more regular with a Gaussian related distribution, while the neuron below discharges in bursts (a common pattern found in Parkinson) showing a double peak histogram.

2.3.3. Autocorrelation diagram

While analyzing time series we must have in mind that the data may have an internal structure (correlation, bias, fluctuation, etc.) that have to be reveal. The autocorrelation graph of a signal gives us information about the periodicity of the system over the time. This graph unveils the presence of oscillations in the firing frequency, allowing us to perceive the rhythmic activity of the neuron. This is done by comparing the original values of signal *s(n)* against its correspondent values fixed on successive delay intervals *k* that occur later in time *s(n+k)*. A good understanding of the signal source is necessary in order to choose the correct duration of delay *k* to apply, and present the oscillatory rhythm at a scale that is better exhibit. While working with neuronal activity, the autocorrelation for high frequencies should be in the order of the milliseconds, while low frequencies correlations occur in the frame of seconds. Both autocorrelation diagrams of each series, for high frequency and low frequency, are generated by the tool to detect frequency rhythms at different scales (Fig. 5). Given a time series of $x_1,x_2...,x_N$, the autocorrelation function for the interval *k*, which studies how system evolved *k* events later, is defined by the following equation

$$r_k = \frac{\sum_{i=1}^{N-k}(x_i-\bar{x})(x_{i+k}-\bar{x})}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

where $\bar{x}$ is the mean of the time series values. The autocorrelation is a coefficient between -1 and 1. By plotting the autocorrelation coefficient values of the series *s(n)* versus *s(n+k)* fixed in successive intervals *k* occurring later in the time, we obtain an autocorrelation diagram, where the observed points distribution tells us about the regularity of the system.

2.3.4. Crossed correlation diagram

This is a similar method to the previous one but it lets us compare the discharges of two different neurons. In this method one of the series is taken as 'reference' and the other, the 'target' one, is displaced and compared against the reference series using different delay intervals *k* positive and negative. This graph gives a notion of dependence (or independence) between both time series. The cross correlation diagram compares two different time series of $x_1,x_2...,x_N$, and $y_1,y_2...,y_{N2}$, belonging to two different neurons discharges. The function to calculate the crossed correlation coefficient is defined as

$$c_k = \frac{\sum_{i=1}^{N-k}(x_i-\bar{x})(y_{i+k}-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2 \sum_{i=1}^{N2}(y_i-\bar{y})^2}}$$

where *N* and *N2* represent the length of each series. The method to generate the crossed correlation diagram selects the first series *x* as reference and plots the crossed correlation coefficient values of the series *x(n)* versus *y(n+k)* for different successive positive and negative intervals *k* (see Fig. 5).

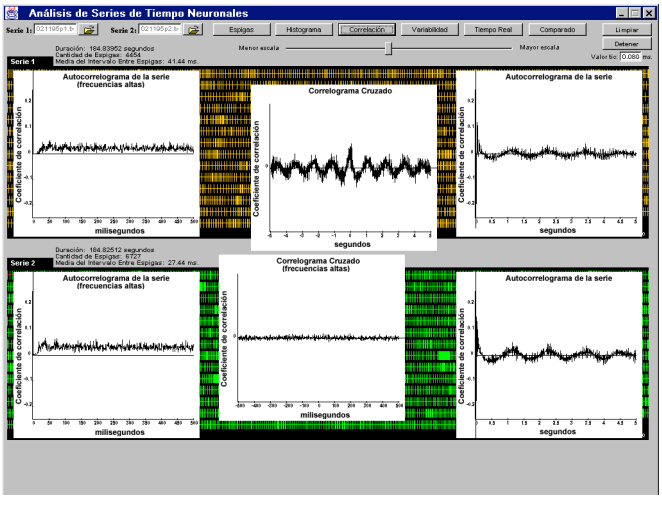


Figure 5: Autocorrelation and Cross correlation diagrams snapshot. Two autocorrelation diagrams are shown for each time series. High frequency autocorrelation diagrams appear on the left and low frequency autocorrelation

diagrams on the right. Both neurons present the same oscillation pattern at the low frequencies, making the synchrony of this pair evident. The generated crossed correlation diagrams are in the centre of the screen, the one at the top show a dependency at the lower frequencies, while at high frequencies, centre bottom, there is no dependency present. The synchrony of this pair of neurons at a low frequency it is confirmed by these diagrams.

All the graphical methods presented up to this point correspond to classic time series analysis.

2.3.5. Variability Diagrams

This graphical method let us study the existence of hidden and not easily detectable correlations in the data set. The variability diagrams represent the existent correlation between successive values of the time series by making different 2D diagrams, correlating 3, 4, 5, etc… successive events of the series. Every value of the series r_i can be related with the next by plotting r_{i+1} as a function of r_i. If there is a correlation among the successive values, a pattern will emerge, but if the correlation is not present only a cloud of dots will be seeing. By correlating just two successive values $X_i^{(0)} = r_i$ and $X_{i+1}^{(0)} = r_{i+1}$ we obtain the 'zero order variability diagram' also known as a first return map. The 'first order variability diagram' is generated by correlating three successive values in the following way $X_i^{(1)} = r_{i+1} - r_i$ and $X_{i+1}^{(1)} = r_{i+2} - r_{i+1}$ and let us perceive the acceleration or deceleration of the system, which in our case are changes in the firing rate of the neuron. For example, if $X_i^{(1)} > 0$ and $X_{i+1}^{(1)} > 0$, then we are in the presence of three successive intervals such as, in each event, the value of *r* is increased. If *r* represents time intervals, it means that the system is going through a deceleration and the spot corresponding to this coordinate in the first order variability diagram will be plotted in the first quadrant (▚). On the contrary, if $X_i^{(1)} < 0$ and $X_{i+1}^{(1)} < 0$ it would be indicating a diminution in the values of *r* and therefore an acceleration of our neuronal system with the corresponding spot on the third quadrant (▞). Successive values of *r* as short-long-short are represented by $X_i^{(1)} > 0$ and $X_{i+1}^{(1)} < 0$, while $X_i^{(1)} < 0$ and $X_{i+1}^{(1)} > 0$ will be associated with long-short-long and will populate the second (▜) and fourth (▙) quadrants respectively. The 'second order variability diagram' is generated by correlating four successive values. By proceeding in this way it is possible to have a notion of the system dynamics in a set of 2D diagrams. These three variability diagrams are generated by the tool for each time series in order to compare its patterns.

Because most biological experimental data sets represent no lineal systems with high dimensionality, the variability diagrams are a useful method to understand its underlying dynamics. This method was previously used with very good results on an assessment of human cardiac dynamics by Babloyantz and Maurer (1996) and by Maurer et al. (1997), but it was never applied to other systems before. We had also performed a pharmacodynamic analysis to follow the effects caused by a drug on the neuronal activity. By plotting the variability diagrams at different moments after the drug was provided we were able to study the modifications in the neural dynamics (see Fig. 6).

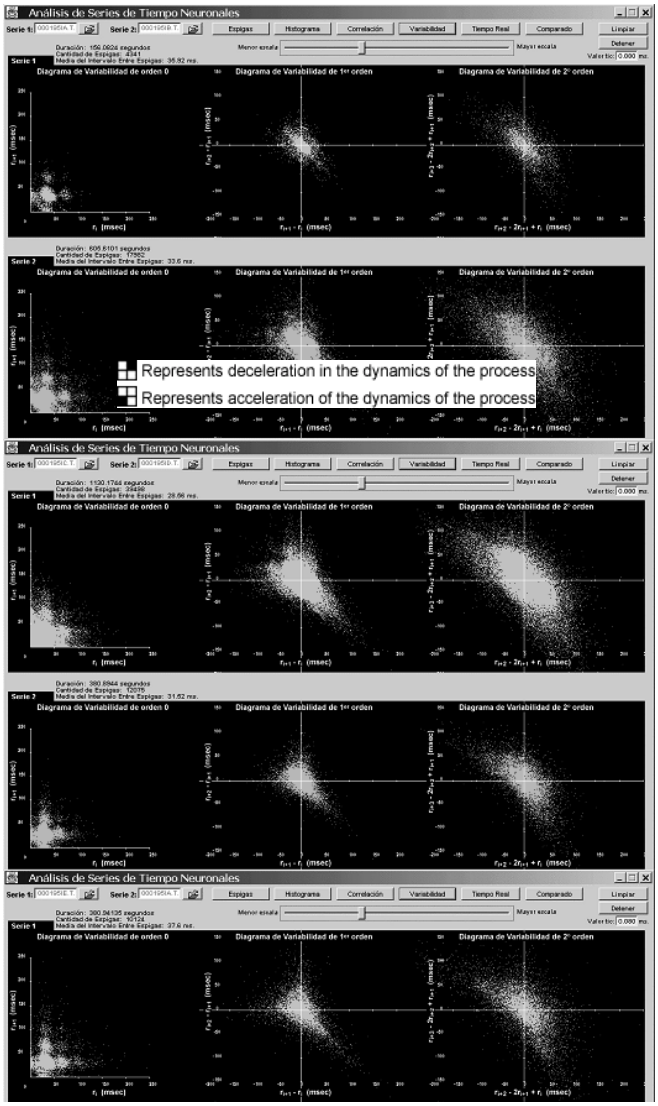


Figure 6: Pharmacodynamics assessment by variability diagrams. The time series belongs to a 45 minutes recording of a single neuron on a laboratory rat to which a drug had been recently provided. The series was divided

into five consecutive time series in order to compare and visualize the neuronal activity of drug effects over time.

2.3.6. Real Time Spike raster-gram comparison

As it was mention in the introduction, due to the particular characteristics of the Parkinson's disease, it is of extreme importance the study of synchrony in certain areas of the brain. Neurophysiologists are not able to completely understand the mechanisms that determine the changes of the temporary firing pattern of the neurons affected by the disease. Nevertheless, observations made in the brain affected area on pairs of neurons revealed the existence of certain underlying synchrony in the mechanics of the system; although the detection of the synchrony could not be easily obtained. As an answer to this problem and in addition to the previous presented methods we implemented a real time graphical method that solves the detection of synchrony.

The real time dynamic graph of the neuron discharges allows us to recreate the original situation at the moment when the recordings were taken. The type of synchrony most commonly found between neurons of the SNpr shows periodic and identical patterns of activity; this means simultaneous discharges on both neurons. Using this computational tool we were able to verify the existence of a different type of synchrony within the neurons in the SNpr, not clearly identified up to date, we denominated 'inhibitory synchrony'. This synchrony behaves exciting or inhibiting a neuron while simultaneously inhibiting or excitating respectively another. A graphical method building a real time spike raster-gram for each neuron, one against the other to have a better comparison of both neuron activities was implemented. Recreating the neuronal activities in this manner we make the presence of synchrony evident and easily detectable. Figure 7 show a snapshot of this parallel dynamic reconstruction of the original recordings.

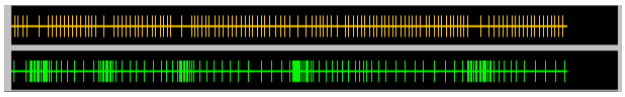


Figure 7: Real time spike raster-gram snapshot showing the comparison of two different neurons where an inhibited synchrony it is make evident.

3 RESULTS

The resulting software is able to detect the existence of determinism (underlying patterns) and synchrony on the neuronal dynamics. It helps the neurophysiologists to make a fast diagnostic in relation to the analysed recordings. This project covers many of the computational necessities of a neurophysiology laboratory and suggests new approaches to the inspection and analysis of the neurological time series data sets.

By developing a computational tool oriented to the analysis of neuronal time series, joining the different traditional methods for studying non linear series with novel methods, presenting the data in a comprehensive way with a user friendly interface we were able to analyze the synchrony aspects of the Parkinson disease providing several new insights not implicated in previous studies.

This unique tool to assess synchrony and other valuable features on time series allows a better and more comprehensive exploration and visualization of the data and acts as a guided tour into the vast and difficult task of time series data analysis. Finally, the software has been generalized to process other signals obtained from different dynamic systems such as the non linear biological systems found in nature.

References

Babloyantz A. and Maurer P. (1996) A graphical representation of local correlations in time series – Assessment of cardiac dynamics. Physics Letters A 221 43-55. Elsevier Science.

Hornykiewicz O. (1966) Dopamine and brain function. Pharmacol Rev. 18:925-964.

Kantz H. and Schreiber T. (1997) Nonlinear Time Series Analysis. Cambridge Nonlinear Science Series 7. Cambridge.

Lozano A. and Lang A. (1998) Parkinson's disease. N Engl J Med. 339(16):1130-43.

Maurer P. et al. (1997) Time structure of chaotic attractors: A graphical view. Physical Review E Volume 56, number 1. The American Physical Society.

Murer G., Dziewczapoiski G., Agid Y., Gershanik O., Raisman-Vozari. (1997) Absence of neurotoxicity of chronic L-DOPA treatment in 6-hydroxydopamine lesioned rats". Neuroreport. 8: 975-79.

Teich M.C. et al. (1997) Estimatring the Fractal Exponent of Point Proceses in Biological Systems.

Tseng K., Kasanetz F., Kargieman L., Pazo J., Murer G., Riquelme L. (2001) Subtalamic nucleus lesions reduce low frecuency oscillatory firing of substantia nigra pars reticulata neurons in a rat model of Parkinson's disease. Brain research. Elsevier Science.

FUZZINESS IN BIOETHICS

Txetxu Ausín

Instituto de Filosofía, CCHS, CSIC
Grupo de Estudios Lógico-Jurídicos(JuriLog)-Unidad de Éticas Aplicadas
txetxu.ausin@cchs.csic.es

Resumen

Planteamos la pertinencia y adecuación de un enfoque *fuzzy* para la bioética. Para ello atendemos a tres dimensiones en esta disciplina filosófica: los hechos y las definiciones; las razones y los argumentos, con especial atención a las analogías y a los argumentos de pendiente resbaladiza; y las normas y valores (deóntica). La conclusión es que un enfoque *fuzzy* para la bioética contempla los principios de proporcionalidad y equidad necesarios para un tratamiento razonable de los conflictos y dilemas bioéticos.

Palabras Clave: bioética, *fuzziness*, hechos, analogías, pendientes resbaladizas, normas, deóntica

1 BIOETHICS

In a broad sense, Bioethics is the study of ethical controversies brought about by the advances in biology and medicine. It could be understood as a kind of "third culture", attending to the classical C.P. Snow's distinction of two cultures, between the humanities and the hard sciences. In that way, Bioethics constitutes a "bridge" between these two sides of research, a sort of "survival science" in terms of V.R. Potter.

Anyway, Bioethics is an interdisciplinary matter, as other applied ethics, since it involves issues that arise in the relations among life sciences, biotechnology, medicine, politics, law, and philosophy.

Bioethics includes polemic issues as euthanasia, abortion, embryonic research, human enhancement, the role of personhood and rational thinking in conferring rights and duties, or animal experimentation.

We find in bioethical debate issues about facts (states of affairs), about definitions (categories), about reasons (arguments), and about norms (values). And it is usual to cope with them from a dichotomyzed point of view, from an "all-or-nothing" approach.

In this article we are going to show the fruitfulness of a fuzzy-logic approach to bioethics when discussing about facts, definitions, and norms in Bioethics.

2 THE ISSUE OF FACTS AND DEFINITIONS IN BIOETHICS

In order to defend a fuzzy-logic approach to Bioethics, I'll stand a strong ontological assumption but very reasonable from a commonsense viewpoint; that is, the gradualism of reality: Any fact, any property, any definition or category, or any state of affairs involved in bioethical debate (as in everyday life and in most sciences) is likely to have fuzzy edges and as regards borderline cases, without precise lines of demarcation [15]. That is the case to face life and death (related to euthanasia, abortion, and organ transplantation), personhood (abortion, embryonic research, hybrids, animal experimentation), health and illness, ordinary treatments versus extraordinary ones, etc. etc.

Remember that a property or a set is fuzzy in the sense that there are degrees of possession thereof or belonging to such a set –so that the relationship of a member belonging to a group varies in different degrees, in the same way as a property may possess it or lack it in various amounts.

This gradualism of reality appears itself in two types of linguistic expressions: Adverbs of intensity and decay: pretty, little, a lot, entirely, somewhat, …; and, comparative constructions: healthier, more or less, so and so, likeness, similar, close, etc.

In the case of life and death, for example, they are undoubtedly gradual (fuzzy) processes, not discrete phenomena that happen in one moment and don't happen at the next instant. The issue is that we learn to recognize the life (to live) and the death (to die) more than to define them in a precise way. So, following Wittgenstein, recognition precedes and goes before definition and has a gradual character since we speak about "more or less life" and "more or less death", including not only the mere

biological sense of life but also its biographical and narrative sense. In some languages, as Spanish, it is usual the expression "she is more dead than alive".

Another example is about the status of personhood, about what or who is a person. We could consider two dichotomic viewpoints: a) A person is every entity or being (material) belonging to the *homo sapiens* kind (species). b) A person is every thinking living being with reason, memory, reasoning, and self-consciousness (as Locke said). The former puts in the same level any material with human DNA (like embryos) to other human and non-human entities capable of psychological properties and skills for intentional action (desires, interests, feelings, etc.). The latter is too restrictive since it gives up of the privileged status of personhood to many beings that not satisfy completely the required conditions (for example, children, disabled people, lost of memory, other primates, etc.).

In fact, as we have pointed out, what prevails in Bioethics about facts and definitions is a dichotomyzed approach dividing sets into two, separately or exclusively, as wholly or fully possessing or not a particular property. Thus, there is a profound disagreement between a continuous and gradual reality, riddled with nuances and transitions, a reality in gray, and a logic (an analysis and description of it) bivalent, between sheer truth and complete falsehood, in "all-or-nothing" terms, black or white.

This leads us to do a critical analysis of some relevant dichotomies in Bioethics [2]. On the one hand, the issue of actions and omissions: Doing and not doing have a very different moral consideration in Bioethics. For example, in the euthanasia debate, usually it is banned active euthanasia (killing) but it is allowed passive euthanasia (letting die). Obviously, it is not the same killing than letting die as if they were equivalent, we would be just as (morally) responsible of the deaths that we cause as of the deaths that we do not prevent; it would be the same not help the starved in Africa as to send them poisoned food. However, the distinction between actions and omissions is more a matter of degree as is the responsibilities they carry out. Omissions, as actions do, cause change in *status quo* and, therefore, have consequences. So, omissions are not radically different to actions and some scholars have analyzed them as a sort of 'negative actions'. Thus, omissions should also be morally evaluated since omissions are not morally neutral. See, for example, the legal punishment of aid refusal; or cases of medical carelessness or negligence; or some parents who don't give food to their children; … In euthanasia debate, someone as Helga Kushe [8] has defended that killing is not always worse, and sometimes better, than letting die (in some situations, doctors assume that death is, from the patient's viewpoint, a benefit rather than a harm –for example, when giving up a breathing tube without deep sedation).

On the other hand, the issue of ordinary means versus extraordinary means (proportionate treatments versus disproportionate ones): A treatment will be proportionate if it offers a reasonable hope of benefit to the patient and disproportionate if not –what is called "therapeutic obstinacy" (cruelty, futility). Related to the refusal of (burdensome) treatment, Catholic Church accepts "the right to die peacefully with human and Christian dignity" and says in its Declaration on Euthanasia (Sacred Congregation for the Doctrine of the Faith, 1980).

But this distinction, which is considered morally significant, is not an abstract distinction on the different treatments; it is, rather, a distinction on the proportion of the benefits that treatment can offer, which will be depend on the patient's situation. The key element is the 'degree' of benefit and welfare which may be provided with such an action or omission.

In short, facts, states of affairs, and categories in Bioethics are fuzzy and, therefore, should be treated in a fuzzy way. As Prof. Sadegh-Zadeh [10] says, "everything in medicine is fuzzy". In the case analyzed before, that on euthanasia, all other relevant features involved in it, apart from actions and omissions and ordinary and extraordinary means, come in degrees too: terminalness, painfulness, unbearableness, and so on.

3 THE ISSUE OF REASONS AND ARGUMENTATION IN BIOETHICS

As we have said, Bioethics involves polemic issues and controversies since its starting point is the pluralism of society about preferences, interests, doctrines, and diverse conceptions of goodness. The other element involved in Bioethics is uncertainty, the issue of how manage risks. The answer is clear: by means of discussion, by the public debate and the exercise of public reason, which imply a strong argumentative exchange.

Argumentation in Bioethics includes two main kinds of argumentative schemes that have a fuzzy and gradual character: Analogies and Slippery Slope Arguments. Let's see in detail.

3.1. ANALOGIES IN BIOETHICS

Analogy has its root in the Greek word 'analogia', understanding as proportion, correspondence, and resemblance. It is the process of transferring information from a particular subject (the analogue or source) to another particular subject (the target). Thus, an analogy establishes an interrelation between two different spheres or domains; it enables us to see aspects of a particular domain in the light of another domain. In short, we can define analogy as similarity in some respects between

things that are otherwise dissimilar; or a comparison based on such similarity.

Analogies have two main functions, interconnected: a) Epistemic: An explanatory function, to shape our perceptions and conceptualizations (comprehension) of phenomena. b) Moral: An argumentative function, to guide us in our handling of phenomena, to argue how thing should be, by comparison with others. This double sense of analogies is patent in Bioethics. Let's present some examples.

In case of abortion, there are some very well known analogies used in the debate [7]. For example, the Violinist Analogy:

> You wake up in the morning and find yourself back to back in bed with an unconscious violinist. He has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist's circulation system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, "Look, we're sorry the Society of Music Lovers did this to you –we would never have permitted it if we had known. But still, they did it, and the violinist now is plugged into you. To unplug you would be to kill him. But never mind, it's only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you".

The rationale is: A woman carries in her womb an unwanted child is like a person who is forced to remain connected to the circulatory system (body) of another unconscious person to keep her alive. This analogy refers to the possibility of abortion in case of pregnancy against own will (for example, in the case of rape).

Another realm in Bioethics where the use of analogies in relevant is about the moral status of non-human animals: The issue is that most, at least nearest to us evolutionary non-humans animals (for examples, apes), have similar, analogous interests, preferences, and sensations (pain, fear) than humans do. However, their moral status is radically different.

> We must fight against the spirit of unconscious cruelty with which we treat the animals. Animals suffer as much as we do. True humanity does not allow us to impose such sufferings on them. It is our duty to make the whole world recognize it. Until we extend our circle of compassion to all living things, humanity will not find peace. (Albert Schweitzer [12])

In analogies we use comparative expressions as more or less, so and so, likeness, similar, near, close,… that have and undoubtedly gradual and fuzzy character. In fact, the degree of similarity between the analogue and the target gives weight to the analogy.

3.2. SLIPPERY SLOPE ARGUMENTS

Another example of fuzziness in bioethical argumentation is the case of slippery slope arguments. This kind of argumentation resembles the old Arabian proverb: If the camel once gets his nose in the tent, his body will soon follow. Slippery slope arguments state that a relatively small first step inevitably leads to a chain of related events culminating in some significant impact, much like an object given a small push over the edge of a slope sliding all the way to the bottom. That is, taking certain step may be fraught with indirect results which could eventually turn our disastrous. But not as an unavoidable outcome, but as a quite possible upshot. Slippery slope arguments could be understood as a kind of presumption: Once that certain step has been taken, there is a presumption that some kind of pressure towards further steps may quite possibly be hard to resist, the thus triggered process yielding a bleak result.

Slippery slope arguments are far from being necessarily fallacious. What alone is fallacious with some slippery slope arguments is wording them as if they were conclusive, non-defeasible or deductively valid reasoning allowing a strong rejection of the proposal and thus closing the discussion [16]. In Bioethics, slippery slope arguments operate as a kind of 'precautionary principle': Caution in advance, a measure taken beforehand against possible danger or failure (risk).

For example, a typical slippery slope argumentation in the euthanasia debate is the following:

> A person apparently hopeless ill may be allowed to take his own life. Then he may be permitted to deputize other to do it for him should he no longer be able to act. The judgment of others then becomes the ruling factor. Already at this point euthanasia is not personal and voluntary, for others are acting "on behalf of" the patient as they see fit. This may will incline them to act on behalf of others patients who have not authorized them to exercise their judgment. It is only a short step, them, from voluntary euthanasia (self-inflicted or authorized), to directed euthanasia administered to a patient who has given no authorization, to involuntary euthanasia conducted as a part of a social policy. … The dangers of euthanasia are too great to all to run the risk of approving it any form. The first slippery step may well lead to a serious and harmful fall. (J. Gay-Williams [5]).

It is usual to identify the sorites (paradox of the heap) type of slippery slope arguments with the vagueness of key concepts or terms. In the discussion on euthanasia, Walton [16] says that slippery slope argument are due to the vagueness of such an expressions as 'letting die', 'actively causing the death', 'complying with patient's willingness', and so on. However, such phrases can be

used in a vague manner, but they can also be used with no vagueness. But, yet all of them denote properties which admit of degrees. As we have pointed out before, all relevant features involved in euthanasia come in degrees: action, omission, terminalness, painfulness, proportionate means, unbearableness, life, and death.

4 THE ISSUE OF NORMS AND VALUES IN BIOETHICS

One of the most important issues in Bioethics is about norms and values involved in as guidance to action. As we have seen before, the core of Bioethics are controversies and hard cases. Moral conflicts and dilemmas constitute the touchtone of Bioethics. Some have tried to minimize, or even deny, the existence of genuine moral dilemmas —in many occasions to avoid acknowledging contradictions in the normative domain, especially among analytic philosophers of law. However, ethical pluralism, the protection of conflicting interests, the diversity of sources of obligations, casuistry itself, or even psychological answers, such as remorse or guilt, seem to make us inclined to accept the existence of normative conflicts. [6]. So, weighing and balancing are the essential elements for rational argumentation in a contingent domain as Bioethics.

In relation to bioethical dilemmas or conflicts, given certain circumstances (with specific particular elements, with different agents involved), the choice in a particular moment for one certain precept, norm, or value over another does not cancel the one being rejected. Even if it weakens it, the now rejected norm can be adopted under different circumstances. In other words, the non-applied norm still holds a degree of normativity that will fade only insofar as it is not being used or in dialectic relation with its use or lack of use in ethical practice. This is especially contemplated in the case of legal norms, in which the jurist's interpretation in favor of a norm and against incompatible one does not eliminate conflict, since it does not necessarily cause its revision or derogation.

This way of normative reasoning would provide reasonable justified 'inclinations' in favor on one of the pans of the scale or one of the points of view at stake. This amounts to an inclination that is consistently non-definitive, non-conclusive, and revisable (*incliner sans necessiter*, in Leibniz' words).

All the previous invite us to a gradualist and fuzzy analysis of normative qualifications, values, and principles. The sense is a sort of 'modest' reason, typical in Bioethics, where moral reasoning should be related to circumstances and situations, against the oversimplification of moral realm and against the 'tyranny of principles' [13]. The goal is 'practical wisdom', rejecting 'moral geometry' in the line of a new casuistry which does not offer 'prescriptions' or 'recipes' but upholds the necessity of analysis, weighing and valuation of circumstances; it is, basically, 'prudent.'

As a result, deontic qualifications are a matter of degree: Licitness, prohibition, and duty should be treated as gradual notions. From this perspective, a particular action can have a more or less degree of licitness and to the extent that it is not completely licit, it will have some degree of illicitness. This leads us to develop a soft deontic logic based on transitive logic, a fuzzy-paraconsistent nonconservative extension of relevant logic E [4].

The core of this type of analysis is the principle of graduation, according to which, when two facts are similar, their deontic treatment must also be similar. The principle of graduation in relation to norms leads to the rejection of the idea that matters involving rights, duties, and prohibitions, that is, moral or juridical matters (bioethical issues), are "all-or-nothing" issues. This entails an anti-absolutist approach to human rights and duties which were unconditional but not absolute. Graduation leaves room for flexibility and adaptability when dealing with particular and contingent circumstances in bioethical debate, and it gives an important role to prudence –understood as reasonableness– in the bioethical domain. From a normative point of view, this gradualist approach brings in an ingredient of malleability and flexibility, which rehabilitates, in some ways, a certain spirit closer to Anglo-Saxon common law.

Moreover, our approach to deontic realm is fact-sensitive. Norms and facts are partly interdependent –despite Hume's and Moore's qualms and strictures. What norms are in operation is not an issue entirely independent of what facts happen in the world. And the other way round. There is certain solidarity between facts and norms. Facts can abolish norms. So, we reject the sharp dichotomy between facts and norms [9], [3].

5 CONCLUSION

There is a profound disagreement between a continuous and gradual reality [deontic and ontological], riddled with nuances and transitions (a reality in gray), and a logic (an analysis and description of it) bivalent, between sheer truth and complete falsehood, in "all-or-nothing" terms, black or white.

As an alternative to the 'principle of bivalence' that permeates the standard approach to reality in general and Bioethics in particular, we maintain the 'principle of gradualism', which says that everything is a matter of degree and therefore a fuzzy-logic approach is an appropriate theoretical method in bioethics. However, gradualness is not the same as vagueness or lack of precision; it is not "indifference" about truth-value, but

determination of a different kind: by degrees, neither exhaustive nor strong, with transitions and edges more or less blurred. Moreover, the fuzzy approach to the world is much more precise than that offered by their correlates sharp and drastic —that yield a simpler and more simplistic view. Otherwise, speech that uses fuzzy terms brings us closer and better to reality with adverbs of intensity and decay and comparative constructions.

But, fuzzy logic does not benefit relativism (moral or epistemological) in any way, but it makes manifest truth's relational character. Thereby the opposition between 'truth' and 'opinion', which has been a matter of concern in Western thought since pre-Socratic times, is rejected as irrelevant and confusing [14]. In a fuzzy calculus, propositions take their truth-value in the interval [0, 1]. Consequently, the notion of 'truth' can be 'modulated', as it happens in ordinary experience and argumentation, so that 'true' and 'false' have lost their static and abstract character.

The notions of weighing (norms and values) as well as those of comparison and similarity (analogies) take place in a gradual manner, by means of transitions instead of leaps or breaks. So, the fuzzy approach to Bioethics entitles us to soften the sharp dichotomies usually stated about bioethical issues, as the well-known ones mentioned above (actions vs. omissions; ordinary means vs. extraordinary ones; proportionate vs. disproportionate).

The main consequence of the fuzzy approach to Bioethics is that it allows us to cope with thousands of dilemmas that arise in our discipline in a way less wrenching, traumatic, and arbitrary than the "all-or-nothing" approach. For example, it allows us rethinking about ethical and legal debates on the limits or boundaries of life (abortion, euthanasia) that presuppose a sharp and clear definition of the beginning and end of life and, thus, a sharp, clear, and absolutely different normative qualifications (and punishments) for very close, complex, and similar facts and situations.

Consequently, similar behaviours and situations can receive a similar normative (ethical and legal) treatment, in the sense of the elementary principles of fairness and proportionality.

References

[1] T. Ausín: *Entre la lógica y el derecho. Paradojas y conflictos normativos*. Plaza y Valdés, México/Madrid, 2005.

[2] T. Ausín y L. Peña: Derecho a la vida y eutanasia: ¿Acortar la vida o acortar la muerte? *Anuario de Filosofía del Derecho* XV(1998), pp. 13-30.

[3] T. Ausín y L. Peña: Arguing from Facts to Duties (and Conversely). In F. van Eemeren et al. (eds.), *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*, Sic Sat, Amsterdam, pp. 45-47, 2003.

[4] T. Ausín & L. Peña: Soft Deontic Logic. In: R. Seising & V. Sanz (eds.), *Soft Computing in Humanities and Social Sciences*, vol.I, Springer, Dordrecht (forthcoming).

[5] J. Gay-Williams: The Wrongfulness of Euthanasia. In: R. Munson (ed.), *Intervention and Reflection: Basic Issues in Medical Ethics*, Wadsworth, Belmont CA, 1979.

[6] C.W. Gowans: *Moral Dilemmas*. Oxford University Press, Oxford, 1987.

[7] J. Jarvis Thompson: A Defence of Abortion. *Philosophy & Public Affairs* 1:1(1971), pp. 47-66.

[8] H. Kushe: Why Killing is not Always Worse –and Sometimes Better– Than Letting Die. *Cambridge Quarterly of Healthcare Ethics* 7:4(1998), pp. 371-374.

[9] H. Putnam: *On the Collapse of Fact/Value Dichotomy and Other Essays*. Harvard University Press, Cambridge MA, 2002.

[10] K. Sadegh-Zadeh: The Fuzzy Revolution: Goodbye to the Aristotelian Weltanschauung. *Artificial Intelligence in Medicine* 21(2001), pp. 1-25.

[11] K. Sadegh-Zadeh: Fuzzy Deontics. In: R. Seising & V. Sanz (eds.), *Soft Computing in Humanities and Social Sciences*, vol.I, Springer, Dordrecht (forthcoming).

[12] A. Schweitzer: *The Philosophy of Civilization*, MacMillan, New York, 7th edition, 1955.

[13] S. Toulmin: The Tyranny of Principles. *The Hastings Center Report* 6(1981), pp. 31-39.

[14] E. Trillas, C. Alsina y J.M. Terricabras: *Introducción a la lógica borrosa*, Ariel, Barcelona, 1995.

[15] M. Vásconez y L. Peña: ¿Qué es una ontología gradual?. *Ágora* 15:2(1996), pp. 29-48.

[16] D. Walton: *Slippery Slope Arguments*. Clarendon Press, Oxford, 1992.

ALGUNOS PROBLEMAS FILOSÓFICOS SOBRE IMPRECISIÓN QUE PERMANECEN EN EL ÁMBITO COMPUTACIONAL

Santiago Fernández Lanza

Escuela Superior de Ingeniería Informática,
Universidade de Vigo (Campus Ourense),
As Lagoas, 32004 - Ourense,
sflanza@uvigo.es

Resumen

El presente trabajo consta de una serie de apuntes y reflexiones sobre cómo algunos problemas filosóficos permanecen aun hoy día en el ámbito de la Recuperación y Extracción de Información, el Procesamiento del Lenguaje Natural, etc. No se trata de realizar ninguna propuesta aplicada al respecto sino una mera puesta en escena de tales problemas a partir de una definición intuitiva del concepto de calidad de un buscador web. Por este motivo, este trabajo consiste en el establecimiento de un listado de preguntas y no en la propuesta de soluciones o respuestas a las mismas.

Palabras Clave: Calidad de un buscador, recuperación de información, extracción de información, comunicación persona-máquina, expansión de consultas, significado, ambigüedad, imprecisión, verdad.

1 INTRODUCCIÓN

Desde el advenimiento de Internet, y más concretamente con el desarrollo de las tecnologías de la Web, los usuarios tenemos a nuestro alcance cada vez más información, y en consecuencia, mayores dificultades de localización de la misma debido a su volumen excesivo. Este es uno de los motivos por el que la utilidad y popularidad de los buscadores en la Web ha crecido más y más desde los orígenes de ésta. La necesidad de construir buenos sistemas de búsqueda que faciliten a los internautas la obtención de información ha tenido como consecuencia el establecimiento de campos de investigación como la *Recuperación de Información (RI)* o la *Extracción de Información (EI)* muy vinculados a otros como el *Procesamiento del Lenguaje Natural* o la *Lingüística Computacional*. En estos ámbitos siempre que se propone una nueva estrategia, procedimiento o metodología aplicable a un buscador, se evalúa si tal propuesta supone una mejora de la calidad del mismo o no. En consecuencia, parece relevante reflexionar sobre la cuestión de la calidad de un sistema de búsqueda.

En el presente trabajo se tratará de indagar si existe alguna definición de calidad aplicada a un buscador que, de forma implícita o explícita, es utilizada en los ámbitos de la *RI* y la *EI*. Esto nos conducirá al manejo de conceptos como los de "precisión", "cobertura", etc. que podrían requerir un tratamiento vago o aproximado. Al tratar estas cuestiones, en última instancia llegaremos hasta problemas clásicos de la filosofía como puede ser el *significado*, la *imprecisión*, la *ambigüedad*, la *verdad*, etc.

2 SOBRE LA CALIDAD DE UN BUSCADOR

¿Cómo se podría definir la calidad de un buscador? Intuitivamente, parece plausible pensar que un buen sistema de búsqueda debería ser aquel que permite realizar consultas de forma cómoda o amigable y que responde a éstas en un tiempo razonablemente corto de forma correcta con la máxima información posible que tenga a su disposición.

Los parámetros de calidad que se coligen de esta definición intuitiva ¿son empleados por los investigadores de *RI* y *EI* cuando evalúan la calidad de un buscador? La respuesta a esta pregunta requiere un análisis previo que permita indicar cuáles son tales parámetros y para ello hace falta estudiar un poco lo dicho en la definición dada:

1. En primer lugar, se indica que un sistema de búsqueda que tenga calidad debe permitir "realizar consultas de forma cómoda o amigable". El primer parámetro de calidad parece apuntar a la interfaz del buscador. En el ámbito computacional en general se conoce este parámetro como *Facilidad de uso* o *amigabilidad*. El formulario de búsqueda es mejor cuanto más simple, un cuadro de edición de texto y un botón para lanzar la búsqueda son suficientes. Añadir nuevos controles debe estar altamente justificado en aras de una clara mejora de la calidad del buscador.

2. Según nuestra definición intuitiva, la respuesta del buscador debe ser realizada "en un tiempo razonablemente corto". Este es un parámetro que se exige, en general, a la mayoría de los sistemas informáticos y se conoce con el nombre de *latencia*. Es el tiempo transcurrido entre el momento en que un usuario realiza una solicitud a un sistema y el momento en el que éste termina de devolver la respuesta. Cuanto más baja sea la latencia, más rápido será el sistema y, en consencuencia, mejor evaluada será su calidad en este punto. Es importante señalar que, en el contexto de la Web, en ocasiones, esta rapidez no va a depender exclusivamente del sistema de búsqueda sino de otros factores como la cantidad de tráfico o la calidad de las subredes existentes entre el equipo del usuario y el servidor de búsqueda.

3. Un buscador con calidad debe responder "de forma correcta". Esta característica apunta a dos temas importantes; el primero de ellos la relaciona con el concepto de *precisión* utilizado en el contexto de la *RI* y la *EI*; el segundo de ellos relaciona el concepto de *corrección* de la respuesta con el concepto de *verdad*. El sistema de búsqueda debe proporcionar exactamente lo que el usuario busca, si la información es muy voluminosa, debe colocar lo relevante al inicio de la respuesta y, por otra parte, si consideramos al buscador como un buscador de calidad, la información proporcionada debería ser fiable y veraz.

4. Además, un buen sistema de búsqueda debe responder "con la máxima información posible que tenga a su disposición". Este parámetro está vinculado a lo que en *RI* y *EI* se conoce como *cobertura*. Una vez localizado lo que se busca, el usuario requiere con frecuencia ver toda la información disponible sobre lo buscado, toda la información relacionada con ello, toda la información afín y, a ser posible, ordenada de mayor a menor proximidad con lo que se busca.

Todas estas características son susceptibles de un análisis desde el punto de vista de la imprecisión. Las dos primeras son comunes en ámbitos más genéricos de la computación, sin embargo, las dos últimas son más específicas del contexto *RI-EI*.

2.1 FACILIDAD DE USO

La facilidad de uso parece un parámetro bastante subjetivo, puesto que depende fundamentalmente de la pericia del usuario, generalmente basada en el hábito de uso de las tecnologías por parte de éste. A un usuario le puede resultar sencillo el manejo de un formulario mientras que a otro le podría resultar altamente tedioso. Además, y al igual que sucedía con el caso de la latencia, la facilidad de uso también ha ido cambiando a lo largo del desarrollo de las nuevas tecnologías de la información. Por otra parte, el tipo de lenguaje que debemos utilizar para realizar la consulta a un buscador automático no es igual que el que emplearíamos con un humano, por lo que requiere cierta fase de aprendizaje, aunque ésta sea simple. La pericia para encontrar las palabras adecuadas en la búsqueda se adquiere, en buena medida, con el uso del sistema de búsqueda después de interiorizar, de manera relativamente implícita, su forma de funcionar. En este aprendizaje suelen ser efectivas las tácticas de ensayo y error. En el apartado 3 se indagará brevemente la diferencia entre la comunicación interpersonal y la comunicación persona-máquina dentro de un contexto pregunta-respuesta, algo que nos conducirá al problema del significado.

2.2 LATENCIA

Podría parecer que la latencia es la que menos problemas de definición y análisis podría dar, ya que es medible de manera más o menos precisa. Sin embargo, depende en exceso del grado de paciencia que un usuario pueda tener. Esto hace que sea difícil responder a la cuestión de cuál es el tiempo máximo de latencia que un buscador debe tener para que no sea penalizado excesivamente en este parámetro. Este tiempo máximo no sólo variará de unos usuarios a otros sino también, para el mismo usuario de un momento a otro dentro de la evolución de las tecnologías relacionadas con Internet. Nosotros mismos, éramos capaces de esperar mucho más tiempo por una respuesta de un buscador a principios de los 90 que hoy en día, ya que en aquellas fechas estábamos acostumbrados a otro tipo de velocidades de red que distan mucho de las actuales.

2.3 PRECISIÓN

Respecto a la precisión, no resulta trivial definir lo que se quiere expresar con este término en *RI* y *EI*, debido a que está relacionado con la mayor o menor consecución del objetivo del usuario cuando éste realiza la búsqueda. Pero, ¿cuál es ese objetivo? ¿existe alguna analogía entre distintos objetivos de distintos usuarios ante una consulta? ¿puede establecerse una clasificación de tipos de objetivos del usuario? En principio, parece que la única información que se tiene *a priori* relativa al objetivo de la consulta de un usuario de un buscador es la cadena de texto que introduce en el formulario de búsqueda y, eventualmente, alguna opción más que el buscador le brinde, como puede ser la selección del formato de la respuesta (imágenes, texto, audio, vídeo, etc.), la selección del idioma de la consulta, etc. *A posteriori* el usuario podrá juzgar en que medida el buscador ha sido útil para el objetivo de su consulta o no, sin embargo, el resultado de este juicio suele permanecer oculto a la mayoría de los buscadores. Un análisis sobre el objetivo de la búsqueda figura en el apartado 4. Este análisis nos conducirá a cuestiones como la *precisión*, la *ambigüedad* o a problemas clásicos como el de la *verdad*.

2.4 COBERTURA

La cobertura va ligada a la precisión y consiste en proporcionar toda la información disponible relacionada con el objetivo de la consulta. De nuevo el problema de la indefinición de dicho objetivo nos lleva a catalogar este término como "poco claro", pero además, ¿qué queremos decir con información relacionada o afín al objetivo de la consulta? ¿existen distintos tipos de relaciones o afinidades? ¿son clasificables estos tipos de relación en el supuesto que existan? No parece absurdo pensar que tal relación es más bien imprecisa y que hay información que está más relacionada con el objetivo que otra, existiendo un continuo entre ellas. Por este motivo, cuando tratamos de ampliar la cobertura de una consulta, es relativamente fácil que se incorpore "ruido" en la respuesta, es decir, que se recupere información que no esté relacionada con el objetivo de la consulta. El tema de la cobertura conduce al problema de la expansión de consultas que se tratará en el apartado 5. Este problema llevará a plantear de nuevo cuestiones relativas a la *ambigüedad* y su representación en contextos computacionales.

3 COMUNICACIÓN INTERPERSONAL Y COMUNICACIÓN PERSONA-MÁQUINA

En el contexto de la búsqueda automática, ¿es el objetivo último establecer una comunicación persona-máquina lo más similar posible a una comunicación entre personas?, en caso de que sea así ¿es posible alcanzar este objetivo? Responder a esta pregunta requiere otra previa, ¿qué significa que la comunicación persona-máquina sea lo más similar posible a la comunicación entre personas? Para dar cuenta de ello aparecen dos opciones a primera vista. En primer lugar, que un observador humano cualquiera, presente en una escenificación de comunicación persona-máquina no sea capaz de distinguir entre ésta y una comunicación interpersonal, es decir, que la máquina pase el Test de Turing ([7]). En segundo lugar, que la comunicación persona-máquina esté constituida por una serie de procesos cuya estructura sea análoga a la de la comunicación entre personas. En este caso, ¿cuáles son estos procesos?. La siguiente lista es una propuesta intuitiva que lleva a nuevas preguntas:

1. Supongamos como punto de partida de una comunicación pregunta-respuesta, una necesidad de información por parte de uno de los interlocutores sobre algo concreto que es el objetivo de la búsqueda. Una necesidad de información relativa a unos contenidos concretos y no un modo de evaluar el conocimiento del encuestado. Pero ¿qué significa "información" en este contexto?

2. Para conseguir la información sobre el objetivo, el emisor la expresa en lenguaje natural. ¿Existe algún lazo de unión entre el objetivo y lo expresado? Estudiar este lazo de unión parece útil para describir el proceso de comunicación, tanto para la comunicación interpersonal como para la comunicación persona-máquina. Pero, ¿es útil este estudio para establecer propuestas computacionales de comunicación persona-máquina?

3. El receptor de la consulta debe interpretar lo que el usuario ha expresado y tratar de identificar el objetivo a partir de lo expresado. ¿Cómo se hace esto? Un humano, para ser eficaz en este contexto, tendría que entender lo expresado y debería indagar si tiene la información que respondería a la consulta. En ocasiones, el receptor puede tener dificultades para interpretar lo expresado por el emisor, y es muy usual que aquel haga preguntas a éste para alcanzar la mayor comprensión de la consulta, estableciéndose así una especie de *feedback*."Entender lo expresado", "interpretar", "comprender" son fenómenos altamente vinculados a la captura del significado. Llegamos así a un problema presente en la Filosofía desde sus orígenes en Grecia y que se consolidó como un tema filosófico central a partir del siglo XX: ¿qué es es significado?.

4. Si el receptor tiene la respuesta, la debe devolver expresándola en lenguaje natural. De nuevo un proceso análogo al del punto 2 se vuelve a dar pero en este caso por parte del receptor de la pregunta.

5. El que realizó la consulta debe comprender la respuesta para que la comunicación haya sido satisfactoria y el ciclo pregunta-respuesta esté completo. Una vez más el proceso descrito en el punto 3 se repite aquí intercambiando a los interlocutores.

En el caso de la comunicación persona-máquina, los puntos 1, 2 y 5 serían llevados a cabo por un humano, mientras que los puntos 3 y 4 deberían ser ejecutados por la máquina si pedimos que la comunicación sea estructuralmente análoga con respecto a la interpersonal. Pero una máquina no parece operar del modo descrito en el esquema anterior. la forma estándar que un usuario tiene para interactuar con un buscador es mediante el uso de un cuadro de texto. La cadena introducida por el usuario podría considerarse como un medio que utiliza éste para conseguir información sobre el objetivo. Dar con la palabra o palabras clave para recuperar el objetivo no siempre es una tarea fácil para quién realiza la consulta. Por eso, este modo de interacción hace que pueda existir cierta distancia entre el objetivo de la consulta y lo que el sistema responde. Esto se debe a que, en esencia, y simplificando mucho, la estrategia más básica, y por otra parte más eficiente, de los sistemas de *RI* y *EI* es la de comparar de forma más o menos sofisticada, las cadenas del texto de la consulta expresada con cadenas de texto presentes en la documentación que posee en su base de datos.

En consecuencia, parece que los puntos 3 y 4 colapsan de algún modo.

La clave de esta diferencia con respecto a la comunicación interpersonal parece estar en las cuestiones respecto al significado indicadas del punto 3. Hasta el momento una máquina no es capaz de capturar el significado tal y como lo hacemos los humanos. Aunque no se tenga clara la respuesta a la cuestión sobre qué es el significado, parece interesante preguntarse ¿existen procedimientos para representar computacionalmente el significado? En el caso de que existan ¿son éstos eficientes para la totalidad del lenguaje o sólo para fragmentos del mismo?

Otra de las diferencias con la comunicación interpersonal es que en el ámbito computacional se suele tratar de evitar las propuestas que implican *feedback*, dado que este tipo de estrategias ralentizan en exceso la comunicación y pueden confundir al usuario.

4 EL OBJETIVO DE LA BÚSQUEDA

El tipo de información que puede ser objeto de una búsqueda de un usuario es muy variopinta. La Web ofrece texto, imágenes, audio, vídeo, servicios de venta de productos, de tramitación burocrática, etc., pero una de las mayores dificultades para un sistema de recuperación suele radicar en la información textual. La información susceptible de constituir el objeto de una consulta en un sistema de *RI* o *EI* se podría clasificar según las siguientes categorías imprecisas sin pretender que la lista sea exhaustiva:

- Según el grado de generalidad o especificidad: Es posible solicitar al sistema información general relacionada con la historia de España del siglo XIX, o información específica como la lista de los reyes godos.
- Según el grado de imprecisión o precisión: Se puede consultar información imprecisa como la predicción meteorológica para mañana o precisa como cuál es el tiempo del récord mundial de los 100 metros lisos actualmente.
- Según el grado de ambigüedad: Los sistemas tratan información relativa a términos como “ratón”, que puede versar sobre, un dispositivo electrónico, un animal o sobre un toro famoso.

El caso es que el objetivo de la consulta de un usuario, puede ir dirigida a recuperar información más o menos general, más o menos precisa o más o menos ambigua. El carácter impreciso de estas características supone, en sí mismo, una dificultad para la recuperación o extracción de tal información. ¿Pueden calcularse estos grados de generalidad, precisión y ambigüedad? En caso afirmativo ¿puede ser este cálculo una herramienta para paliar la citada dificultad?

Existen también otra serie de parámetros borrosos que tienen que ver con las características del usuario respecto a la consulta. De nuevo sin pretensión de exhaustividad, algunos de estos parámetros podrían ser:

- El grado de claridad que el usuario pueda tener del objetivo de la búsqueda. En ocasiones un usuario puede buscar sin tener claro lo que busca. Puede buscar un listado de nombres propios, de tal forma que hiciese válida una respuesta como la lista de los reyes godos, una alineación de un partido de fútbol o la parrilla de salida de un gran premio de Fórmula 1.
- El grado de conocimiento previo que éste tenga del objetivo buscado. No es lo mismo que un usuario busque la lista de los reyes godos, que la lista de unos reyes de los cuales el usuario desconoce el apelativo “godos”. Es diferente buscar una información que el usuario conocía pero se le ha olvidado, que buscar algo nuevo o desconocido para el usuario.
- La pericia del usuario para realizar consultas. No todos los usuarios tienen la misma destreza para consultar, ya sea por la falta de experiencia con el uso de buscadores, por falta de manejo del idioma de la consulta, desconocimiento del idioma de la respuesta, etc.
- La capacidad del usuario para identificar el objetivo en la respuesta del sistema. Un sistema puede proporcionar la respuesta que coincide con el objetivo del usuario pero éste no ser capaz de identificarla por simple desconocimiento o falta de información sobre el tema. La respuesta al tiempo del récord mundial de los 100 metros lisos puede estar en la página personal del atleta, devuelta por el sistema, cuyo nombre puede ser desconocido para el usuario.
- El grado de fiabilidad necesario en la respuesta. La Web no garantiza la veracidad de los datos que contiene. Puede suceder que la información que obtengamos sea total o parcialmente falsa o que esté incompleta. Hay consultas que exigen un mayor grado de fiabilidad en la respuesta que otras. No es lo mismo una consulta sobre los récord mundiales de atletismo que sobre listados de alimentos que pueden comer los celíacos.
- La capacidad del usuario para identificar la fiabilidad de la información de la respuesta. Ante una respuesta, el usuario puede no ser capaz de saber si la fuente es o no fiable. Por ejemplo, es plausible recibir respuestas con información de sitios web que publican noticias de broma, o sitios web serios que publican noticias que posteriormente fueron desmentidas.

Estos dos últimos parámetros conducen al problema de la verdad. Una vez obtenida la respuesta ¿es esta verdadera o

falsa?, una versión menos exigente de la misma pregunta, ¿es esta información veraz o mendaz? ¿existen estrategias para averiguar la veracidad o mendacidad de la información obtenida? Si las hay, ¿son éstas automatizables? ¿se basan exclusivamente en cuestiones de opinión?

5 EXPANSIÓN DE CONSULTAS

Con el fin de obtener la máxima información sobre el objetivo de la búsqueda, los sistemas de *RI* y *EI* deben tender a responder con la máxima cobertura posible. Esto hace que no todas las respuestas tengan el mismo grado de adecuación a la consulta y que se deban establecer procedimientos para ordenar las respuestas de mayor a menor grado de proximidad con respecto al objetivo. Por tanto, parece relevante para la mejora de los sistemas de *RI* y *EI*, diseñar estrategias que contribuyan a aumentar la cobertura de las respuestas. El estudio de estas estrategias, es conocido en el ámbito con el nombre de *expansión de consultas*. Dicho de una forma intuitiva, la mayor parte de estas estrategias suelen "manipular" o "retocar" la consulta del usuario con el fin de aumentar la cobertura. Suponiendo que el sistema contiene documentos con información relativa al objetivo de la consulta, hay varios motivos por los que una consulta o bien no produce ninguna respuesta, o bien produce una respuesta completamente inadecuada al objetivo, o bien produce respuestas adecuadas e inadecuadas al objetivo y además hay respuestas adecuadas que el sistema no devuelve, o bien sólo produce respuestas adecuadas pero hay respuestas adecuadas que el sistema no devuelve. Algunos de estos motivos son los siguientes:

- El usuario ha cometido errores tipográficos a la hora de escribir. Este problema es tratado en el ámbito de la expansión de consultas mediante el uso de correctores ortográficos automáticos, que sustituyen o sugieren formas correctas similares a las de la consulta.

- El usuario ha empleado una palabra correcta pero morfológicamente distinta a la que figura en los documentos que contienen el objetivo. Ampliando la consulta con palabras que difieren morfológicamente de las que figuran inicialmente en ella, puede que se obtengan mayor número de resultados, algunos de los cuales podrán contener el objetivo de la consulta. En este caso, se utilizan lematizadores que añaden a la consulta expresiones morfológicamente parecidas a las empleadas en ella.

- El usuario ha introducido una palabra correcta, pero semánticamente relacionada con aquella o aquellas que aparecen en los documentos que contienen el objetivo. Las palabras introducidas en la consulta pueden ampliarse con aquellas que están semánticamente relacionadas con éstas, con lo que se obtendrá una mayor cobertura. Para dar cuenta de este problema se utilizan diccionarios que añaden a la consulta palabras semánticamente próximas a las introducidas por el usuario.

- El sistema trata todas las expresiones de la consulta como si tuvieran la misma importancia para la consecución del objetivo. Unas palabras pueden tener más peso que otras para conseguir que la consulta tenga éxito. Tratarlas a todas por igual puede provocar respuestas no deseadas. Para dar cuenta de este problema, se analiza la relevancia de las palabras de la consulta y se le da prioridad a las que tienen mayor relevancia.

Independientemente de si se puede evaluar o no que la expansión sea correcta, el aumento de la cobertura suele provocar que el sistema devuelva información no relevante para el objetivo de la consulta, es decir, suele provocar perdida de precisión o aumento de ruido. Esto se debe fundamentalmente a que los diccionarios empleados en la expansión reflejan un fenómeno que se da en el lenguaje natural: la ambigüedad o polisemia. Podría considerarse que una palabra o expresión, tomada fuera del contexto oracional, es ambigua si posee varias acepciones o significados. Esta característica de algunas palabras se refleja, por regla general, en los diccionarios de sinónimos o expresiones semánticamente relacionadas, mediante la asociación de varios conjuntos de sinónimos o palabras relacionadas para una única entrada. Cuando se emplea una expresión polisémica en una consulta, si ésta figura como entrada del diccionario y éste se emplea para realizar expansión, no se sabe *a priori* cuál de las distintas acepciones es la adecuada al uso de la expresión en el contexto de la consulta, por lo que se expande con todas las acepciones disponibles. Generalmente, sólo una de estas acepciones es la adecuada y, en consecuencia, es la única que se debería emplear en la expansión; pero localizar y seleccionar esta acepción, o lo que es lo mismo, eliminar la ambigüedad de las expresiones utilizadas, no es una tarea fácil. Son necesarias, por tanto, propuestas de análisis y tratamiento de la ambigüedad. ¿Son posibles propuestas de tratamiento de la ambigüedad que sean computacionalmente adecuadas? En caso afirmativo, ¿solucionarían algún tipo de problema tales propuestas en este ámbito?

6 EVALUACIÓN DE UN BUSCADOR

Cuando se evalúa un sistema de *RI* o *EI*, ¿se tienen en cuenta los parámetros derivados de la definición propuesta en el apartado 2 y todas sus consecuencias, compromisos teórico-prácticos, problemas, carencias, etc. aquí relatados? Volviendo a los parámetros relatados en 4, para evaluar un buscador, se suele pensar en un usuario ideal, que tiene claro lo que busca, tiene conocimiento previo suficiente sobre lo buscado, conoce el grado de fiabilidad

necesario para la respuesta y lo sabe identificar, tiene pericia para realizar consultas tanto a nivel técnico como idiomático, es capaz de identificar el objetivo en la respuesta proporcionada por el sistema, etc., pero en la realidad no siempre se dan estas condiciones ideales. Por este motivo, no resulta fácil la evaluación de la calidad de tales sistemas, así como tampoco lo es la mejora de los mismos.

Respecto a la evaluación de precisión y cobertura, es usual utilizar *corpus* que, en algunas ocasiones están elaborados a mano por humanos. ¿Es este el mejor procedimiento teniendo en cuenta la frecuencia con la que los humanos cometemos errores? ¿son capaces estos humanos de tener en cuenta todas las posibles respuestas a una consulta teniendo en cuenta la ambigüedad, imprecisión y otras caracteristicas del lenguaje aquí reseñadas? Además, este tipo de corpus ¿están disponibles para cualquier idioma?

7 CONCLUSIONES

Partiendo de la pregunta ¿Cómo se podría definir la calidad de un buscador?, se ha propuesto una definición intuitiva para dar respuesta a la misma. De tal definición se coligen cuatro parámetros de índole imprecisa que conducen a una serie de cuestiones, algunas de las cuales son clásicas en el ámbito de la Filosofía. Las preguntas más fundamentales derivadas de cada parámetro son las siguientes:

- Facilidad de uso: ¿es el objetivo último establecer una comunicación persona-máquina lo más similar posible a una comunicación entre personas?, ¿es posible alcanzar este objetivo?, ¿qué significa que la comunicación persona-máquina sea lo más similar posible a la comunicación entre personas?, ¿cuáles son los procesos que tienen lugar en la comunicación interpersonal?, ¿qué significa información en el contexto de la *RI* y la *EI*?, ¿existe algún lazo de unión entre el objetivo de una consulta y lo expresado por el usuario cuando escribe la consulta en el buscador?, ¿es útil el estudio de este lazo de unión para establecer propuestas computacionales de comunicación persona-máquina?, ¿cómo se indentifica el objetivo de la consulta a partir de lo que expresa el usuario para tal fin?, ¿qué es es significado?, ¿existen procedimientos para representar computacionalmente el significado?, ¿son éstos eficientes para la totalidad del lenguaje o sólo para fragmentos del mismo?
- Latencia: ¿cuál es el tiempo máximo de latencia que un buscador debe tener para que no sea penalizado excesivamente en este parámetro?
- Precisión: ¿cuál es el objetivo de un usuario cuando realiza una consulta en un buscador?, ¿existe alguna analogía entre distintos objetivos de distintos usuarios ante una consulta?, ¿puede establecerse una clasificación de tipos de objetivos del usuario?, ¿pueden calcularse los grados de generalidad, precisión y ambigüedad del objetivo de una consulta? ¿puede ser este cálculo una herramienta para paliar dificultad que suponen estos factores en *RI* y *EI*?, ¿es verdadera o falsa la respuesta de un buscador?, ¿es veraz o mendaz la respuesta de un buscador? ¿existen estrategias para averiguar la veracidad o mendacidad de la información obtenida?, ¿son éstas automatizables?, ¿se basan exclusivamente en cuestiones de opinión?
- Cobertura: ¿qué queremos decir con información relacionada o afín al objetivo de la consulta? ¿existen distintos tipos de relaciones o afinidades?, ¿son clasificables estos tipos de relación en el supuesto que existan?, ¿son posibles propuestas de tratamiento de la ambigüedad que sean computacionalmente adecuadas?, ¿solucionarían algún tipo de problema tales propuestas en *RI* y *EI*?

Estos cuatro parámetros ¿son empleados por los investigadores de *RI* y *EI* cuando evalúan la calidad de un buscador?, ¿se tienen en cuenta además todas sus consecuencias, compromisos teórico-prácticos, problemas, carencias, etc. aquí relatados?

Agradecimientos

Este trabajo ha sido financiado por los proyectos FFI2008-03902 del Ministerio de Ciencia y Tecnología, FFI2009-08828 del Ministerio de Ciencia e Innovación y INOU11A-04 de la Diputación Provincial de Ourense.

Referencias

[1] R. Baeza-Yates, B. Ribeiro-Neto: *Modern Information Retrieval*. Addison-Wesley, ACM Press, New York, 1999.

[2] I. M. Bochenski: *Historia de la Lógica Formal*. Gredos, 1966.

[3] V. Cross: Fuzzy information retrieval, *Journal of Intelligent Information Systems*, Volume 3, Number 1, pp. 29–56, 1994.

[4] E. Herrera-Viedma, G. Pasi, F. Crestani: *Soft computing in Web information retrieval: models and applications*, Springer, 2006.

[5] W. Kneale, Kneale, M: *El desarrollo de la Lógica*. Tecnos, 1972.

[6] G. Pasi: Flexible information retrieval: some research trends, *Mathware and Soft Computing* 9, 2002.

[7] A.M. Turing: Computing machinery and intelligence. *Mind*, 59, pp. 433–460, 1950.

MÉTODO PARA EL DESCUBRIMIENTO DE CONOCIMIENTO ABSTRACTO BASADO EN PROXIMIDAD Y SU INCORPORACIÓN AL SISTEMA BOUSI~PROLOG

Pascual Julián-Iranzo[1], **Clemente Rubio-Manzano**[2]

[1]*Dept. de Tecnologías y Sistemas de Información, Universidad de Castilla-La Mancha*, Pascual.Julian@uclm.es
[2]*Dept. de Sistemas de Información, Universidad del Bío-Bío*, clrubio@ubiobio.cl

Resumen

En el presente trabajo se propone e implementa un método genérico para el descubrimiento de conocimiento abstracto y/o general a partir de las ecuaciones de proximidad que modelan la imprecisión dentro de un marco de Programación Lógica Borrosa basado en unificación débil. Este método hace uso de las λ-clases de similaridad y los λ-bloques de proximidad que definen una relación de similaridad y proximidad, respectivamente. Cuando el universo de discurso lo constituyen conceptos lingüísticos las λ-clases y los λ-bloques pueden ser vistos como agrupaciones jerarquizadas por niveles de abstracción. Entonces, cada agrupación puede etiquetarse, con la ayuda del usuario, con un descriptor lingüístico más general que identificará al grupo simulando un proceso de generalización basado en proximidad. Gracias a este método el sistema puede aprender conceptos que inicialmente no conocía y responder a preguntas que en un principio no era capaz de responder. La novedad de este trabajo reside en que es la primera vez que un método de estas características se introduce en un marco de Programación Lógica Borrosa. Para comprobar la viabilidad del método se ha desarrollado una herramienta y se ha integrado en el seno del sistema Bousi~Prolog. Una motivación adicional es continuar mejorando el lenguaje Bousi~Prolog para el desarrollo de aplicaciones de Computación Flexible en general y, la Computación con Palabras en particular, ya que el proceso de abstracción/generalización forma parte del razonamiento ordinario y por tanto un sistema que tenga como objetivo la computación con palabras debería gestionar este proceso cognitivo.

Palabras Clave: Relaciones de Proximidad y Similaridad, Programación Lógica Borrosa, Descubrimiento de Conocimiento Abstracto.

1 INTRODUCCIÓN

La Programación Lógica es un paradigma de programación que permite la construcción de software especificando "qué" queremos computar y no el "cómo", aspecto este último que se delega de forma automática al sistema de control. Además, como la Programación Lógica está basada en la Lógica, está bien dotada para la representación del conocimiento y la obtención de nuevo conocimiento a partir del representado. La Programación Lógica, en principio, no permite representar ni manejar de forma natural la Vaguedad debido a su forma rígida de contestar a una pregunta. Esta característica puede considerarse una limitación a la hora de modelar ciertos problemas. Esto se pone de manifiesto en el siguiente ejemplo.

Ejemplo 1 *Supongamos un fragmento de una base de datos deductiva que almacena información sobre personas y sus preferencias. Ante la cuestión de si a alguien le gusta el 'Whiskey', una persona podría responder: 'No me gusta', 'Me gusta un poco', 'Me gusta mucho', 'Me gusta muchísimo'. Una sencilla regla podría ser que cuando a alguien le gusta el 'Whiskey' es posible que lo compre. Esta situación puede ser modelada por el siguiente programa* Prolog.

```
%% HECHOS
gusta(juan, whiskey, un_poco).
gusta(maria, whiskey, mucho).
gusta(pedro, whiskey, muchisimo).
gusta(pablo, whiskey, no_me_gusta).

%% REGLAS
comprar(X,P):-gusta(X, P, mucho).
```

Si preguntamos al sistema sobre las personas que podrían comprar Whiskey*, lanzando para ello el objetivo* ?-gusta(X,whiskey)*, solamente se nos daría como respuesta:* X=maria. *Sin embargo, es razonable pensar que* juan *y* pedro *también serían candidatos para comprar* Whiskey.

Del ejemplo se deduce que la Programación Lógica no cuenta con la capacidad de manejar conceptos que pueden estar relacionados por su significado. En el ejemplo

anterior 'un poco', 'mucho', 'muchísimo' y 'no me gusta' hacen referencia al grado en que le gusta el Whiskey a un individuo. Por lo tanto, sería interesante contar con mecanismos que permitan relacionar, dentro de un dominio, conocimiento próximo o semejante, cuantificar dicha relación, e introducir estos nuevos conceptos en un marco basado en Programación Lógica. Este hecho, motivo la aparición de la Programación Lógica basada en Similaridad [4, 11] y la Programación Lógica basada en Proximidad [7]. Una realización práctica de ambos marcos es el lenguaje Bousi~Prolog [1] (BPL, para abreviar) [8] , donde el algoritmo de unificación de la Programación Lógica ha sido sustituido por un algoritmo de unificación basado en relaciones de proximidad/similaridad. Por tanto, su mecanismo operacional es una variante del principio de resolución SLD presentado en [12]. Bousi~Prolog permite la comunicación con los tesauros WordNet-Similarity [10] y FOG [13]. Además, se ha investigado la aplicación de BPL para: Búsqueda Flexible de Respuestas, Emparejamiento Avanzado de Patrones [6], Recuperación de Información donde la información textual se selecciona o analiza usando una ontología), Catalogación de textos [5], Base de Datos Flexibles, Sistemas Basados en el Conocimiento o Razonamiento Aproximado [8]. El siguiente ejemplo ayuda a entender la sintaxis y semántica del lenguaje.

Ejemplo 2 *Supongamos que queremos modelar el problema descrito en el Ejemplo 1 en un marco de* Programación Lógica Borrosa *basado en unificación débil. Un posible programa* BPL *que implemente esta situación sería el siguiente:*

```
%% ECUACIONES DE PROXIMIDAD
mucho~muchisimo=1.0.    un_poco~muchisimo=0.4.
un_poco~mucho=0.6.      no_me_gusta~un_poco=0.5.

%% HECHOS
gusta(juan, whiskey, un_poco).
gusta(pedro, whiskey, muchisimo).
gusta(maria, whiskey, mucho).
gusta(pablo, whiskey, no_me_gusta).

%% REGLAS
comprar(X,P):-gusta(X, P, mucho).
```

Si preguntamos al sistema sobre las personas que podrían comprar Whiskey*, lanzando para ello el objetivo* ?-gusta(X,whiskey)*, se nos daría como respuesta:* X=maria *con* 1.0*;* X=juan *con* 0.6 *y* X=pedro *con* 1.0*. Gracias a que el sistema aprovecha el conocimiento encerrado en las ecuaciones de proximidad, que modelan la imprecisión de este problema.*

El hecho de emplear un mecanismo que relaciona conceptos por su significado hace que podamos manejar fenómenos lingüísticos que se basan en esas relaciones, por ejemplo: la sinonimia, la antonimia o la generalidad. Aquí nos centramos en este último. Es importante mencionar que la Programación Lógica, a diferencia de lo que ocurría con la vaguedad, sí permite tratar la generalidad de forma natural. De hecho, la Programación Lógica permite la definición de redes semánticas de conceptos donde está presente la relación "es-un". Sin embargo, en un marco de Programación Lógica basada en Similaridad y/o Proximidad al emplear ecuaciones de proximidad para relacionar conceptos puede inducirse nuevo conocimiento abstracto y/general basado en proximidad entre los mismos. Nosotros estamos interesados en automatizar este proceso de generalización y/o abstracción.

El objetivo de este trabajo es, por tanto, proponer e implementar un método para el descubrimiento de conocimiento abstracto dentro de un marco de Programación Lógica Borrosa basado unificación débil e integrarlo en el seno del sistema BPL. Dotar a BPL con esta capacidad mejora su mecanismo de inferencia al soportar el aprendizaje inductivo de nuevos conceptos y, por tanto, deducir conocimiento que inicialmente el sistema no era capaz de deducir. Este es pues un paso más en el objetivo de afianzar a BPL como un auténtico lenguaje para la Computación Flexible en general y, la Computación con Palabras en particular, ya que el proceso de abstracción/generalización forma parte del razonamiento ordinario y por tanto un sistema que tenga como objetivo la computación con palabras debería gestionar este proceso cognitivo.

La organización del trabajo es como sigue. La Sección 2 se encarga de introducir el tópico de la generalización y la abstracción de conocimiento general, además de revisar el concepto de inducción orientada a atributos. En la Sección 3 se definen las relaciones de proximidad y se estudia su aplicación a la abstracción de cocimiento abstracto cuando se definen sobre un conjunto de conceptos lingüísticos. La Sección 4 describe una aproximación genérica para el descubrimiento de conocimiento general en un marco de Programación Lógica Borrosa basado en unificación débil. En la Sección 5 se describe una herramienta que implementa la funcionalidad del método presentado. Por último, la Sección 6 se dedica a las conclusiones y el trabajo futuro.

2 GENERALIZACIÓN Y ABSTRACCIÓN DE CONOCIMIENTO

La generalización es una propiedad fundamental de la lógica y el razonamiento humano. Dados dos conceptos, $\mathscr{C}_1$ y $\mathscr{C}_2$, el concepto $\mathscr{C}_1$ es una generalización del concepto $\mathscr{C}_2$ si y sólo si: (a) cada instancia del concepto $\mathscr{C}_2$ es también una instancia del concepto C_1; (b) existen instancias del concepto C_1 que no son instancias del concepto $\mathscr{C}_2$. De forma equivalente, $\mathscr{C}_1$ es una generalización de $\mathscr{C}_2$ si $\mathscr{C}_2$ es una especialización de $\mathscr{C}_1$. Por ejemplo, "animal" es una generalización de "ave" porque toda "ave" es un "animal",

[1] BPL esta disponible públicamente y puede ser encontrado en la URL: http://dectau.uclm.es/bousi

y hay animales que no son aves (un perro, por ejemplo). La generalización/especialización lingüística esta bien definida mediante el par hiperónimo/hipónimo.

El tratamiento de la generalización en los sistemas software se ha realizado mediante el uso de tesauros que proporcionan el par hiperónimo/hipónimo o por medio de ontologías donde se representa la relación de jerarquía entre conceptos. Otra solución ha sido la abstracción de conocimiento general a partir del conocimiento específico [9]. Para llevar a cabo esta tarea se han empleado técnicas de minería de datos. Estas técnicas simulan el proceso de generalización a partir de los datos representados en el sistema. La salida de este proceso es una jerarquía por niveles donde el nivel cero representa los datos más generales y los niveles posteriores datos más específicos. Por ejemplo, dados los conceptos "Francia", "Reino Unido", "Venezuela", "España", "Brasil", "Portugal" se puede realizar un proceso de generalización o abstracción donde el concepto "Europa" identifique al conjunto "Francia", "Reino Unido", "España" y "Portugal", y de este grupo podríamos hacer un subgrupo adicional "Países en Crisis" que estaría formado por "España" y "Portugal". Esta técnica se denomina "inducción orientada a atributos" [3] y permite obtener una jerarquía a partir de los datos básicos disponibles en el sistema. Una extensión borrosa de está técnica se aplicó en [1] para el descubrimiento de conocimiento abstracto sobre bases de datos borrosas. En este trabajo se emplearon las clases de similaridad como agrupaciones. Nosotros pretendemos generalizar tal mecanismo y llevarlo a un marco de Programación Lógica Borrosa que trabaje con relaciones de proximidad, permitiendo, de esta manera, el descubrimiento de conceptos lingüísticos a partir de las λ-clases de similaridad y los λ-bloques de proximidad que definen una relación de proximidad. En el caso particular de las relaciones de similaridad, se emplearían λ-clase de similaridad.

El tratamiento de la generalización en un marco de Programación Lógica Borrosa basado en unificación débil pasa por representarla de forma explicita mediante las ecuaciones de proximidad. Si no fuera así, pueden surgir algunos problemas:

Ejemplo 3 *Supongamos una base de datos que almacena información sobre jugadores de cartas y su país de procedencia. Además definimos un conjunto de ecuaciones de proximidad que representan la cercanía de un país a otro:*

```
%%HECHOS
jugador(john,usa).        jugador(hugo,venezuela).
jugador(mary,canada).     jugador(rico,colombia).
jugador(alfred,mexico).   jugador(lloyd,australia).
jugador(alomu,nueva_zelanda).

%%ECUACIONES DE PROXIMIDAD
canada~usa=0.8.           canada~mexico=0.5.
canada~colombia=0.1.      canada~venezuela=0.1.
usa~mexico=0.8.           usa~colombia=0.3.
usa~venezuela=0.2.        mexico~colombia=0.4.
mexico~venezuela=0.2.     colombia~venezuela=0.8.
nueva_zelanda~australia=0.8.
```

Si preguntamos al sistema qué jugadores son de "américa_del_sur"*, lanzando para ello el objetivo* ?- jugador(X,america_del_sur) *el sistema respondería* "No" *ya que en ningún momento se modela la generalización que existe en la inclusión de un país a un continente.*

Una solución a este problema es obtener las ecuaciones de proximidad que representan la generalización de un tesauro que permita obtener el par hiperónimo/hipónimo o, que la generalización se represente correctamente en la ontología definiendo cada una de las ecuaciones de proximidad necesarias, por ejemplo, que "américa_del_sur" sea próxima a "colombia" y a "venezuela". Sin embargo, aún contando con estas soluciones, no tratamos correctamente los procesos de generalización que no están fijados por un par hiperónimo/hipónimo (por ejemplo, "países en crisis" es una generalización que puede ser cierta ahora y no serlo en un futuro). Además, podría darse la necesidad de emplear abstracciones más subjetivas y libres de cualquier formalización, como por ejemplo, agrupar a {naranja, manzana, plátano} como "fruta favorita de mi hijo" o al conjunto { ron, whiskey, vodka } como "bebidas alcohólicas que no me gustan".

Por tanto, sería adecuado contar con un proceso semiautomático de generalización que permita la clasificación de elementos en grupos en base a las decisiones del usuario para identificar cada uno de los elementos mediante un descriptor que sea, al mismo tiempo, un concepto más general. Este método puede verse como un proceso de aprendizaje porque el usuario enseña al sistema nuevos conceptos a partir del conocimiento representado, además podría ayudar al usuario al descubrimiento de nuevas relaciones y, por tanto, al descubrimiento de nuevo conocimiento. Por otro lado, este método simula un proceso de abstracción que: (1) reduce la dependencia de las ecuaciones de proximidad; (2) mejora, mediante resúmenes de datos, la compresión de la ontología y; (3) proporciona mayor libertad al usuario para definir su propia generalización.

3 RELACIONES DE PROXIMIDAD SOBRE CONCEPTOS LINGÜÍSTICOS

Una relación borrosa binaria, $\mathscr{R}$, sobre un conjunto ordinario U, es un subconjunto borroso sobre $U \times U$ (esto es, una aplicación $U \times U \rightarrow [0,1]$). Una relación de similaridad cumple las propiedades reflexiva, simétrica y t-transitiva, mientras una relación de proximidad no tiene porque cumplir la t-transitividad. Existe una estrecha relación entre las relaciones de similaridad y las de equivalencia: el λ-corte de una relación borrosa R sobre U, $R_\lambda = \{\langle x,y\rangle \mid \mathscr{R}(x,y) \geq \lambda\}$, es una relación de equivalencia sobre U. Recordemos que una relación de equivalencia

sobre un conjunto define una partición en este conjunto, dado por las clases de equivalencia y que esto caracteriza aquellos recubrimientos (converings) de un conjunto que son familias de clases de relaciones de equivalencia.

Las relaciones de proximidad son más complicadas a este respecto. Para una proximidad sobre el conjunto U, la noción de clase de equivalencia se divide en dos conceptos diferentes: los bloques y las clases de proximidad. Dada una relación de proximidad $\mathscr{R}$ sobre un conjunto U, un bloque de nivel λ (λ-bloque), $\mathscr{B}^{\lambda}$, es un subconjunto de U tal que la restricción de $\mathscr{R}_{\lambda}$ para este conjunto es una relación total y maximal con esta propiedad. Por otro lado, la clase de nivel λ (λ-clase) de un elemento x en U es un subconjunto $\mathscr{K}_{\lambda}(x) = \{y \in U | R(x,y) \geq \lambda\}$ de aquellos elementos de U que son próximos a x (en un determinado nivel λ). Los bloques y las clases de una relación de proximidad sobre U forman recubrimientos, pero no necesariamente particiones (si uno de estos recubrimientos es una partición de U, entonces R es una relación de similaridad y los recubrimientos por bloques y clases coinciden). Estos nos dan una imagen detallada de la estructura de una relación de proximidad. En particular, los recubrimientos por bloques de $\mathscr{R}$ determinan completamente $\mathscr{R}$ porque se puede observar que, para $x, y \in U$, $R(x,y)$ es igual al mayor valor $\lambda \in [0,1]$ para el cual x e y están en el mismo bloque de nivel λ.

Ejemplo 4 *Dada $\mathscr{R}$ una relación de proximidad con $R(p,r) = 0.8$, $R(a,b) = 0.7$ y $R(a,c) = 0.5$. La Figura 1 describe nivel a nivel la estructura por bloques de esta relación de proximidad, la cual está caracterizada por un conjunto $\Lambda = \{0.5, 0.7, 0.8, 1.0\}$ de niveles de aproximación. En particular para el valor $\lambda = 0.5$, existen tres bloques de nivel 0.5: $\mathscr{B}_1 = \{p,r\}$, $\mathscr{B}_2 = \{a,b\}$, $\mathscr{B}_3 = \{a,c\}$. Por otro lado, las clases de nivel 0.5 para cada uno de los elementos son: $\mathscr{K}_{0.5}(p) = \{p,r\}$, $\mathscr{K}_{0.5}(r) = \{r,p\}$, $\mathscr{K}_{0.5}(a) = \{a,b,c\}$, $\mathscr{K}_{0.5}(b) = \{b,a\}$, $\mathscr{K}_{0.5}(c) = \{c,a\}$, formando el recubrimiento de nivel 0.5.*

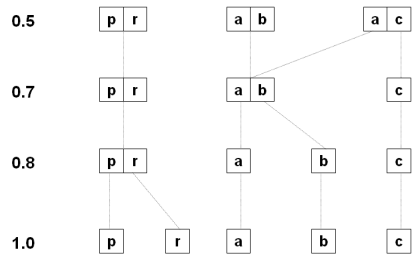


Figura 1: Bloques por niveles de aproximación

Cuando el universo de discurso está formado por conceptos lingüísticos, una relación de proximidad o similaridad establece una relación entre ellos en función de un determinado criterio de clasificación o abstracción. Un criterio de abstracción o clasificación es una característica o propiedad por la cual se realiza una agrupación de elementos (conceptos) en función de que cumplan en menor o mayor grado dicho criterio. Por ejemplo, un criterio de clasificación sería "pertenecer a Europa o América" (localización), "países de habla inglesa" (lenguaje), "tener el pelo claro" (intensidad del color de pelo), etc. Entonces, decimos que dos conceptos $\mathscr{C}_1$ y $\mathscr{C}_2$ tienen el mismo significado si bajo un determinado criterio de abstracción dichos conceptos son próximos (semánticamente). Por ejemplo, "España" y "Reino Unido" son próximos bajo el criterio "localización", i.e., "pertenecer a Europa", "tener el pelo rubio" y "tener el pelo canoso" son próximos bajo el criterio "tener el pelo claro", "lentejas" y "fruta" son próximos bajo el criterio "comida sana", etc. De esta manera, las λ-clases o λ-bloques representan agrupaciones de conceptos que son similares bajo un mismo criterio. Cuando disponemos de un conjunto de conceptos lingüísticos $\{\mathscr{C}_1, \ldots, \mathscr{C}_n\}$ que pueden identificarse por un concepto que los hace similares, ese concepto se denomina descriptor lingüístico más general de $\{\mathscr{C}_1, \ldots, \mathscr{C}_n\}$. Por ejemplo, "América" sería el descriptor lingüístico del conjunto {Usa, Canada, México, Colombia, Venezuela} ya que "América" identifica al criterio "pertenecer al continente americano". "Claro" sería un descriptor del conjunto {Rubio, Canoso} ya que "Claro" identifica al criterio "tener el pelo Claro". Por tanto, estas agrupaciones de λ-clases o λ-bloques se pueden identificar por un descriptor lingüístico más general, estableciendo así un proceso de generalización basado en proximidad.

4 UNA APROXIMACIÓN GENÉRICA

En esta sección proponemos una aproximación genérica que nos permita realizar, de forma automática, el proceso descrito en la sección anterior. Para ello, es necesario disponer de un algoritmo que obtenga los λ-bloques de proximidad o λ-clases de similaridad. Una vez los λ-bloques o las λ-clases de similaridad se han generado se etiquetan mediante un descriptor único que los identificará ante el resto. Posteriormente se generan nuevas ecuaciones de proximidad que ponen en relación ese nuevo descriptor con cada elemento del conjunto descrito por el mismo. Entonces, este nuevo concepto puede participar en el proceso de inferencia de un lenguaje de programación lógica borrosa basado en unificación débil permitiendo al usuario lanzar preguntas sobre conceptos que inicialmente el sistema no conocía y que han sido generados de forma semiautomática. Informalmente el método se define, paso a paso, como sigue: (1) Se parte de un conjunto de ecuaciones de proximidad (definidas por el usuario u obtenidas de un tesauro; (2) Se genera una relación de proximidad o similaridad; (3) Se ejecuta el algoritmo que permite obtener las λ-clases de similaridad o los λ-bloques de proximidad; (4) Se etiquetan aquellas agrupaciones que son susceptibles de generalización (este proceso necesitará de la intervención del usuario); (5) Se generan las nuevas ecuaciones de proximidad en base al etiquetado realizado; (6) Se cargan las ecuacio-

nes de proximidad en un sistema de programación lógica borrosa basado en unificación débil.

La formalización de este método se describe a continuación. En primer lugar, se define un algoritmo que computa automáticamente una jerarquía de conceptos a partir de un conjunto de ecuaciones de proximidad.

Algoritmo 1

Entrada: E, ecuaciones de proximidad y $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$, niveles de aproximación.
Salida: Una jerarquía de bloques $\mathscr{J}$.
Inicialización: $\mathscr{J} := \emptyset$; $\mathscr{R}$:= generarRelación(E).
Para todo nivel $\lambda_j \in \Lambda$:

1. BP^{λ_j}:=generarBloquesProximidad($\mathscr{R}_{\lambda_j}$)
2. insertar BP^{λ_j} en $\mathscr{J}$

Devolver $\mathscr{J}$

El siguiente paso, una vez que disponemos de la jerarquía de bloques de proximidad $\mathscr{J}$, consiste en etiquetar algún $\mathscr{B}_k^{\lambda_j} \in BP^{\lambda_j} \subseteq \mathscr{J}$ con un descriptor d_k y por cada elemento $e \in \mathscr{B}_k^{\lambda_j}$ generar la ecuación de proximidad "$d_k \sim e = 1.0$". Finalmente, observemos que cuando la relación de proximidad $\mathscr{R}$ es una similaridad, la noción de bloque de proximidad, empleada en la descripción del Algoritmo 1, se reduce a la noción de clase de similaridad, que permite una generación más eficiente.

Ejemplo 5 *Dada la relación de proximidad del ejemplo 3 y suponiendo el nivel* $\lambda = 0.5$. *Los* λ*-bloques que se generan en ese nivel son:* $\mathscr{B}_1 = \{canada, usa, mxico\}$, $\mathscr{B}_2 = \{colombia, venezuela\}$ *y* $\mathscr{B}_3 = \{nueva_zelanda, australia\}$. *Entonces, el usuario podría asignar el descriptor "*américa_del_sur*" al bloque* $\mathscr{B}_2$, *"*américa_del_norte*" al bloque* $\mathscr{B}_1$ *y "*oceanía*" al bloque* $\mathscr{B}_3$. *Entonces se generarían las ecuaciones que se muestran (sólo indicamos un subconjunto significativo de las mismas):* "`america_del_sur~colombia=1.0.`" *y* "`america_del_sur~venezuela=1.0.`".

Esta técnica puede verse como un proceso de generalización donde el usuario se encarga de seleccionar el mejor descriptor para un λ-bloque de proximidad o λ-clase de similaridad. Cuando el usuario selecciona el descriptor se generarán, automáticamente, un conjunto de ecuaciones de proximidad que podrán participar en el mecanismo de inferencia de un lenguaje de programación lógica borrosa basado en unificación débil, como por ejemplo, Bousi~Prolog.

5 HERRAMIENTA PARA EL DESCUBRIMIENTO DE CONOCIMIENTO ABSTRACTO

Para la realización práctica del método presentado en la sección anterior, se ha construido una herramienta (ver Figura 2) que permite automatizar el proceso de generación de λ-clases de similaridad y λ-bloques de proximidad, facilita la asignación de un identificador lingüístico a cada uno de las agrupaciones seleccionadas y genera las ecuaciones de proximidad correspondientes. Esta herramienta cuenta con ventanas gráficas que facilitan la interacción con el usuario: Visualizador de Jerarquía (1) que muestra la jerarquía generada y que permite seleccionar un bloque o clase pulsando sobre él. Cuando se selecciona la información que contiene aparecerá en la Ventana de Selección de Descriptor (2) que además cuenta con un campo de texto para introducir el descriptor. Elegido este, se pulsa el botón "Generar Ecuaciones". Entonces, las ecuaciones se generan y se cargan en el Visualizador de Ecuaciones (3). Una vez hecho esto, las ecuaciones de proximidad pueden cargarse en el sistema pulsando el botón "Cargar Ecuaciones".

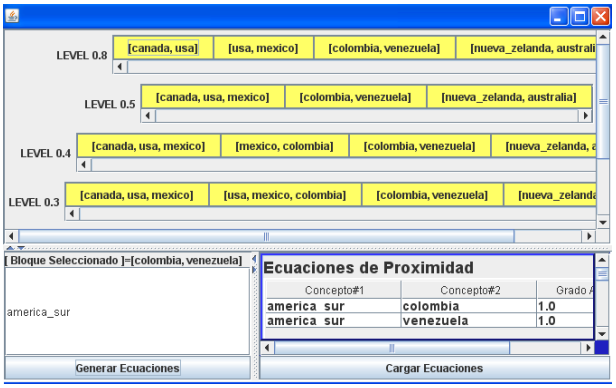


Figura 2: Herramienta para el descubrimiento de conocimiento abstracto

La herramienta se ha implementado en Java 7. Para ello se han construido dos capas. Una capa de dominio encargada de implementar toda la funcionalidad del método descrito y que se compone de cuatro clases: la clase Bloque que es principalmente un lista de conjuntos (bloques o clases), la clase Nivel compuesta de una lista de objetos de la clase Bloque, además de un campo "grado" que indica el nivel, la clase Jerarquía compuesta de una lista de objetos de la clase Nivel y la clase Generador que se encarga de implementar una extensión borrosa del Algoritmo del Clique [2]. La segunda capa, de presentación, tiene como función implementar la componente gráfica de la herramienta. Esta herramienta se ha integrado en el sistema Bousi~Prolog lo que permite hacer uso de las ecuaciones generadas por la herramienta como se muestra en el siguiente ejemplo.

Ejemplo 6 *Volviendo al programa del Ejemplo 3 y suponiendo que han sido cargadas las ecuaciones del Ejemplo 5, entonces al preguntar al sistema sobre que jugadores son de* "america_del_sur" *el sistema respondería* "X=hugo" *con* 1.0 *y* "X=rico" *con* 1.0.

6 CONCLUSIONES Y TRABAJO FUTURO

Hemos propuesto e implementado un método que permite al usuario definir su propia generalización de conceptos, dentro de un marco de programación lógica borrosa basado en proximidad y/o similaridad, y lo hemos integrado en el sistema Bousi∼Prolog. Gracias a este método conseguimos manejar de forma semiautomática la relación de generalidad inducida en una representación del conocimiento construida a partir de las relaciones semánticas establecidas por las ecuaciones de proximidad.

Algunas ventajas importantes de nuestro marco son: (a) los descriptores aprendidos se fijan una vez y se usan muchas, añadiendo nuevo conocimiento; (b) las clases y los bloques pueden ayudar a la depuración de la ontología y son una forma de compresión de información; (c) proporciona libertad al usuario a la hora de definir la generalización. Sin embargo, podemos apuntar posibles desventajas: (a) el usuario tiene que intervenir en la elección del descriptor lingüístico; (b) el tiempo de generación de clases y bloques puede ser alto si contamos con una gran número de ecuaciones de proximidad.

Respecto al trabajo futuro queremos flexibilizar la generación automática de ecuaciones de proximidad permitiendo, por ejemplo, la posibilidad de establecer relaciones entre los descriptores lingüísticos. Por otro lado queremos usar la programación multinúcleo para mejorar el algoritmo encargado de procesar los bloques y las clases. En definitiva, seguir mejorando el sistema Bousi∼Prolog para que se convierta en una herramienta eficaz para el desarrollo de aplicaciones que realicen Computación Flexible.

Referencias

[1] R.A. Angryk and F.E. Petry Discovery of Abstract Knowledge from Non-Atomic Attribute Values in Fuzzy Relational Databases. In: *Modern Information Processing: From Theory to Applications* (2005).

[2] C. Bron and J.Kerbosh Algorithm 457: Finding All Cliques of an Undirected Graph. In: *Communications of ACM Vol. 16*, Issue 9. ACM Press (1973).

[3] Y. Cai, N. Cercone and J. Han Attribute-Oriented Induction in Relational Databases. In: *Proc. IJCAI-89*, pp. 26–36 (1989).

[4] F. Formato, G. Gerla, and M.I. Sessa. Similarity-based Unification. In: *Fundam. Inform.*, 41(4):393–414 (2000).

[5] F.P. Romero, P. Julián, A. Soto, M. Ferreira, J. Gallardo. A Fuzzy Declarative Approach for Classifying Unlabeled Short Texts Using Thesauri. In. *Proc. WILF 2011*, pp. 180-187 (2011).

[6] P. Julián, C. Rubio and J. Gallardo. Bousi∼Prolog: A Prolog Extension Language for Flexible Query Answering. In: *ENTCS, vol 248*, pp. 131–147. Elsevier (2009).

[7] P. Julián and C. Rubio. A Declarative Semantics for Bousi∼Prolog. In: *Proc. of 11th ACM SIGPLAN Symposium on PPDP'09*, pp. 149–160 (2009).

[8] P. Julián and C.Rubio. Bousi∼Prolog - A Fuzzy Logic Programming Language for Modeling Vague Knowledge and Approximate Reasoning. In: *Proc. of IJCCI (ICFC-ICNC)*, pp. 93–98 (2010).

[9] D. H. Lee and M. H. Kim. Database Summarization Using Fuzzy ISA Hierarchies. IEEE Transactions on Systems, Man and Cybernetics–part B, 27(1), pp.68-78 (1997).

[10] T. Pedersen, S. Patwardhan and Jason Michelizzi. WordNet::Similarity - Measuring the Relatedness of Concepts. In: *Proc. AAAI-04*, pp. 1024-1025 (2004).

[11] M.I. Sessa. Approximate Reasoning by Similarity-based SLD Resolution. *Theorical Computer Science*, 275:389–426 (2002).

[12] J.A. Robinson. A Machine-oriented Logic Based on the Resolution Principle. *Journal of the ACM*, 12(1):23–41, 1965.

[13] E. Fdz-Viñas, J. Serrano-Guerrero, J. A. Olivas, J. de la Mata, A. Soto. FOG: Arquitectura Flexible para la Generación Automática de Ontologías. In: *Proc. ESTYLF'10*, pp. 525–530 (2010).

RAZONAMIENTO SILOGÍSTICO APROXIMADO CON CUANTIFICADORES GENERALIZADOS

M. Pereira-Fariña[1], **Juan C. Vidal-Aguiar**[1], **P. Montoto**[2], **F. Díaz-Hermida**[1], **A. Bugarín**[1]

[1]*Centro de Investigación en Tecnoloxías da Información (CITIUS), Universidade de Santiago de Compostela*, {martin.pereira,juan.vidal,felix.diaz,alberto.bugarin.diz}@usc.es

[2]*Escola Técnica Superior de Enxeñaría, Universidade de Santiago de Compostela*, pmgberm@gmail.com

Resumen

En este trabajo se plantea una extensión de propuestas anteriores en el razonamiento silogístico aproximado. Se manejan tipos de cuantificadores procedentes de la Teoría lingüística de Cuantificadores Generalizados, como los comparativos ("*hay el doble de profesores que estudiantes*") o de excepción ("*todas las muestras menos dos han dado negativo en sustancias dopantes*"). Otra ventaja de este modelo es que el número de premisas no está limitado, lo que permite obtener conclusiones en casos de razonamiento con un número amplio de proposiciones. Para manejar estas características, el problema de razonamiento se transforma en un problema equivalente de optimización matemática donde las premisas del silogismo constituyen las restricciones del espacio de búsqueda.

Palabras Clave: Silogismo borroso, cuantificadores generalizados, razonamiento aproximado, optimización matemática.

1 INTRODUCCIÓN

El silogismo o razonamiento silogístico es una forma de razonamiento deductivo en el que tanto las premisas como la conclusión son oraciones cuantificadas. Un ejemplo bien conocido de este tipo de inferencias se muestra en la Tabla 1, donde *PR*1 denota la primera premisa, *PR*2 la segunda y *C* la conclusión.

La forma usual de las oraciones que aparecen en el silogismo es "*Q As* son *Bs*", donde *Q* denota un cuantificador, *A* el término-sujeto o restricción y *B* el término-predicado o alcance.

La primera aproximación sistemática a este tipo de razonamiento, desarrollada por Aristóteles [1] y conocida como

Tabla 1: Silogismo clásico.

*PR*1	Todos los seres humanos son mortales
*PR*2	Todos los griegos son seres humanos
C	Todos los griegos son mortales

"Silogística", trata argumentaciones formadas por dos premisas con exactamente un término en común (término medio) y una conclusión (cuyos términos no son el término medio). En el ejemplo de la Tabla 1 *seres humanos* es el término medio, *griegos* y *mortales* son los términos de la conclusión y *todos* el cuantificador de las premisas y la conclusión. La silogística aristotélica se definió para los cuatro cuantificadores lógicos clásicos (*todos*, *ninguno*, *alguno* y *alguno... no / no todos*), por lo que su capacidad expresiva para sentencias del lenguaje natural es muy limitada.

Las aproximaciones posteriores de la bibliografía proponen ampliar la capacidad expresiva del silogismo explorando fundamentalmente dos vías alternativas: *i*) añadiendo nuevos cuantificadores a los cuatro clásicos, tanto nítidos [2] como borrosos [3, 4, 5, 6, 7, 8] (como *la mayor parte, pocos,...*), limitados a ciertos patrones de inferencia; *ii*) contemplando más de dos premisas, pero considerando únicamente los cuatro cuantificadores clásicos [9]. En cuanto al tratamiento de cuantificadores borrosos en el silogismo, del análisis de las propuestas más relevantes [3, 4, 5, 6] se concluye [10] que sólo se consideran los cuantificadores de la dicotomía típica de la lógica borrosa absoluto/relativo,[1] sin tener en cuenta otro tipo de cuantificadores, como los comparativos (*hay el doble de,...*), de excepción (*todos menos tres,...*) u otros; por otro lado, los patrones de inferencia propuestos tampoco permiten incluir dichos tipos de cuantificadores. En [11], también se analizan las mencionadas propuestas de silogismo borroso, pero enfocándose en su operativa y no en los tipos de cuantificadores considerados.

En este contexto, el objetivo principal de este trabajo es

[1]Ejemplos de cuantificadores absolutos son *alrededor de 5, casi ninguno,...* ; ejemplos de relativos son *pocos, casi todos,...*

plantear una formulación general del silogismo que pueda manejar diferentes tipos de cuantificadores (no sólo del tipo absoluto/relativo) así como un número mayor de premisas.

Para alcanzar este objetivo, tomamos como punto de partida el análisis de la *Teoría de los Cuantificadores Generalizados* [12] (TGQ[2]) respecto del fenómeno de la cuantificación: los cuantificadores son predicados de orden superior que denotan relaciones entre conjuntos. De este modo, una oración cuantificada queda definida como una relación entre conjuntos (las propiedades expresadas en la oración) en la cual el cuantificador denota su cardinalidad, que puede ser expresada según la dicotomía absoluto/proporcional (*cinco/la mayoría*) así como con la de nítido/borroso (*cinco/alrededor de cinco*). En este trabajo, abordamos la generalización comentada mediante el uso de cuantificadores definidos como intervalos $[a,b]$. Este tratamiento, además de ser compatible con los modelos de la TGQ, permite también tratar cuantificadores borrosos, bien directamente a través de su aproximación intervalar $\{Ker, Sup\}$, como utilizando definiciones basadas en α-cortes.

Se proporciona también un método de resolución para la conclusión del silogismo, según el marco de cuantificación adoptado, ya que el razonamiento puede ser expresado mediante un conjunto de inecuaciones. Esto nos permite definir el proceso de razonamiento como un problema de optimización matemática y así aplicar sus métodos de resolución.

2 FORMULACIÓN GENERAL DEL SILOGISMO

Sea E el conjunto de elementos del universo de discurso y sea $P = \{P_s,\, s = 1,\ldots,S\}$ el conjunto de las propiedades de interés definidas sobre E. Por ejemplo, P_1 puede ser la propiedad "ser estudiante" y P_2 "ser gallego". Respecto de la estructura de las oraciones que constituyen el silogismo, mantenemos la forma estándar:

$$Q\ L_1 \text{ son } L_2 \tag{1}$$

donde Q denota un cuantificador lingüístico y las propiedades P_s son denotadas por L_1 (término-sujeto o restricción) y L_2 (término-predicado o alcance). En este caso, también supondremos que tanto L_1 como L_2 pueden ser combinaciones booleanas de las propiedades (ej., L_1 =*ser gallego y no ser estudiante*), lo que nos permitirá abstraernos del tratamiento casuístico habitual en la bibliografía [10], donde en muchas argumentaciones aparecen este tipo de combinaciones tanto en la restricción como en el alcance del cuantificador. Por otra parte, dado el carácter general de nuestra aproximación, una argumentación silogística está constituida por N premisas relacionadas entre si por los términos de las sentencias bajo el siguiente esquema:

$$\begin{array}{rl} PR1: & Q_1\ L_{1,1} \text{ son } L_{1,2} \\ PR2: & Q_2\ L_{2,1} \text{ son } L_{2,2} \\ \ldots & \\ PRN: & Q_N\ L_{N,1} \text{ son } L_{N,2} \\ \hline C: & Q_C\ L_{C,1} \text{ son } L_{C,2} \end{array} \tag{2}$$

donde Q_n, $n = 1,\ldots,N$ son el conjunto de cuantificadores lingüísticos de las N premisas; los $L_{n,j}$, $n = 1,\ldots,N$, $j = 1,2$ representan combinaciones booleanas arbitrarias de las propiedades consideradas en el silogismo y Q_C es el cuantificador de la conclusión, calculado en términos de los cuantificadores de las premisas.

Entenderemos que en las premisas de la argumentación silogística las restricciones $L_{n,1}$, $n = 1,\ldots,N$ son no vacías. Esto es muy razonable desde el punto de vista del uso del lenguaje, ya que es esperable que los hablantes cuantifiquen acerca de conjuntos que contengan algún elemento. Esta restricción evita ciertos problemas de indefinición de las proporciones asociadas a cuantificadores proporcionales que tienden a trivializar la conclusión (esto es, la argumentación silogística no permite realizar ningún tipo de conclusión sobre el cuantificador de la conclusión). En todo caso, la consideración de que el alcance de las expresiones cuantificadas pueda ser vacío y su impacto en la argumentación silogística se analizará en más detalle en futuras ampliaciones del presente trabajo.

Respecto de la semántica de los cuantificadores, asumimos inicialmente que esta se define a partir de una relación intervalar entre las cardinalidades de las propiedades a las que se aplica el cuantificador. Por ejemplo, consideramos la oración "*Entre 3 y 6 estudiantes son gallegos*"; donde *estudiantes* es denotado por Y_1 y *gallegos* por Y_2. Así pues, sea $Q : P^2(E) \to \{0,1\}$ el cuantificador binario absoluto "*entre tres y seis*." Entonces, una definición intervalar de Q es:

$$Q(Y_1,Y_2) = \begin{cases} 0 : |Y_1 \cap Y_2| \notin [3,6] \\ 1 : |Y_1 \cap Y_2| \in [3,6] \end{cases} \tag{3}$$

En la sección 3.2 plantearemos otras definiciones intervalares asociables a cuantificadores lingüísticos de la TGQ.

3 TRANSFORMACIÓN DEL RAZONAMIENTO SILOGÍSTICO EN UN PROBLEMA DE OPTIMIZACIÓN

En la aproximación que proponemos en este trabajo, convertimos el problema del razonamiento (determinar el cuantificador de la conclusión en función de los cuantificadores de las premisas) en un problema de optimización. Esta idea ya se apunta en [5], pero específicamente para

[2]De sus siglas en inglés, *Theory of Generalized Quantifiers*.

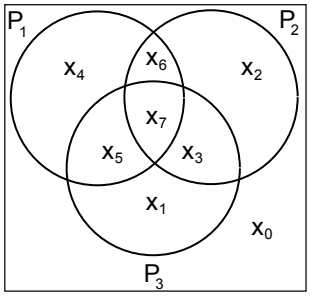


Figura 1: División del universo en conjuntos disjuntos para S=3 propiedades.

algunos patrones de inferencia silogística y con cuantificadores proporcionales. Nosotros proponemos su extensión para argumentaciones silogísticas con N premisas y para cualquier tipo de cuantificador compatible con la TGQ. Inicialmente consideramos Q definidos mediante un intervalo $Q=[a,b]$, lo que permite tratar de forma directa cuantificadores *crisp* y de forma indirecta cuantificadores borrosos.

Distinguimos tres etapas en la transformación del problema: *i)* la división del universo de referencia en regiones disjuntas; *ii)* la definición de los cuantificadores como conjuntos de inecuaciones; *iii)* el método de optimización que puede ser utilizado en cada caso.

3.1 DIVISIÓN DEL UNIVERSO EN REGIONES DISJUNTAS

El primer paso es transformar las S propiedades P_s, $s=1,\ldots,S$ del universo E en un nuevo conjunto de propiedades disjuntas $PD=\{x_0,\ldots,x_K\}$, con $K=2^S-1$, definidas como:

$$\begin{aligned} x_0 &= \overline{P_1}\cap\overline{P_2}\cap\ldots\overline{P_{S-1}}\cap\overline{P_S} \\ x_1 &= \overline{P_1}\cap\overline{P_2}\cap\ldots\overline{P_{S-1}}\cap P_S \\ &\ldots \\ x_K &= P_1\cap P_2\cap\ldots P_{S-1}\cap P_S \end{aligned} \tag{4}$$

Los x_k, $k=0,\ldots,K$ son de la forma $x_k=T_1\cap\cdots\cap T_S$, donde $T_s=P_s$ ó $T_s=\overline{P_s}$, $s=1,\ldots,S$, con lo que quedan definidas las diferentes regiones disjuntas en las que las intersecciones de las propiedades originales y sus complementarios dividen el universo de referencia E. Un ejemplo se muestra en la Figura 1, para S=3 propiedades, donde los x_k, $k=0,...,7$ denotan cada uno de los conjuntos disjuntos que se generan a partir de las propiedades P_1,P_2 y P_3.

Así, tenemos en ese ejemplo, que: P_1=$x_4\cup x_5\cup x_6\cup x_7$, $\overline{P_2}\cap P_3=x_1\cup x_5$, o que $P_1\cap P_2\cap P_3=x_7$

3.2 DEFINICIÓN DE LOS CUANTIFICADORES COMO CONJUNTOS DE INECUACIONES

En la Tabla 2 recogemos las definiciones de algunos tipos de cuantificadores generalizados según la TGQ. Por simplicidad, consideramos el caso de dos propiedades Y_1,Y_2 (cuantificadores binarios).

Tabla 2: Cuantificadores binarios generalizados. $Q=[a,b]$

Cuantificadores lógicos	
algún (Y_1,Y_2)	$0: Y_1\cap Y_2=\emptyset$ $1: Y_1\cap Y_2\neq\emptyset$
todos (Y_1,Y_2)	$0: Y_1\not\subseteq Y_2$ $1: Y_1\subseteq Y_2$
ningún (Y_1,Y_2)	$0: Y_1\cap Y_2\neq\emptyset$ $1: Y_1\cap Y_2=\emptyset$
no todos (Y_1,Y_2)	$0: Y_1\subseteq Y_2$ $1: Y_1\not\subseteq Y_2$
Cuantificadores binarios absolutos	
$Q_{BA}(Y_1,Y_2)$	$0: \lvert Y_1\cap Y_2\rvert\notin[a,b]$ $1: \lvert Y_1\cap Y_2\rvert\in[a,b]$
Cuantificadores binarios de excepción	
$Q_{BE}(Y_1,Y_2)$	$0: \lvert Y_1\cap\overline{Y_2}\rvert\notin[a,b]$ $1: \lvert Y_1\cap\overline{Y_2}\rvert\in[a,b]$
Cuantificadores binarios proporcionales	
$Q_{BP}(Y_1,Y_2)=$	$1: \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert}\geq a\wedge\frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert}\leq b$ $0: \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert}<a\vee\frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert}>b$ $\lvert Y_1\rvert=0$
Cuantificadores binarios comparativos	
$Q_{BC-ABS}(Y_1,Y_2)=$	$0: \lvert Y_1-Y_2\rvert\notin[a,b]$ $1: \lvert Y_1-Y_2\rvert\in[a,b]$
$Q_{BC-PROP}(Y_1,Y_2)=$	$0: \frac{\lvert Y_1\rvert}{\lvert Y_2\rvert}\notin[a,b]$ $1: \lvert\frac{Y_1}{\lvert Y_2\rvert}\rvert\in[a,b]$
Cuantificadores de similaridad	
$Q_S(Y_1,Y_2)=$	$1: \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\cup Y_2\rvert}\geq a, Y_1\cup Y_2\neq\emptyset$ $0: \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\cup Y_2\rvert}<a, Y_1\cup Y_2\neq\emptyset$ $1: Y_1\cup Y_2=\emptyset$

A partir de estas definiciones, se generan los diferentes conjuntos de inecuaciones que definen los cuantificadores. Supongamos las siguientes combinaciones boolenas de S=2 conjuntos disjuntos:

$$\begin{aligned} x_0 &= \overline{Y_1}\cap\overline{Y_2} \\ x_1 &= \overline{Y_1}\cap Y_2 \\ x_2 &= Y_1\cap\overline{Y_2} \\ x_3 &= Y_1\cap Y_2 \end{aligned} \tag{5}$$

Abusando de la notación, y por simplicidad, los valores $x_0,\ldots,x_K$, con $K=2^S-1$ denotarán también las cardinalidades de las propiedades disjuntas del mismo nombre $|x_0|,\ldots,|x_K|$. A modo de ejemplo, desarrollaremos la transformación de los cuantificadores binarios proporcionales (que permiten modelar expresiones como *Muchos A son B*) y de los cuantificadores de similaridad (que permiten modelar expresiones como *A se parece a B*), los cuales, ge-

neran patrones silogísticos aproximados no considerados hasta ahora en la bibliografía.

Para un cuantificador binario proporcional, definido como un intervalo $[a,b]$, si la relación entre la restricción y el alcance está en $[a,b]$, el cuantificador tiene un valor 1 mientras que si no lo está, su valor es 0. Así pues, según las cardinalidades definidas en la expresión (5) y la definición correspondiente de la Tabla 2, un cuantificador binario proporcional $Q_{BP}(Y_1,Y_2)$ se reescribe del siguiente modo:

$$\frac{x_3}{x_1+x_2+x_3} \geq a, \qquad \frac{x_3}{x_1+x_2+x_3} \leq b,$$

de donde obtenemos

$$-ax_1-ax_2+(1-a)x_3 \geq 0, \quad -bx_1-bx_2+(1-b)x_3 \leq 0 \quad (6)$$

En el caso de los cuantificadores de similaridad, se fija el umbral de similaridad en un valor a (expresado en tanto por 1), por lo que obtenemos la siguiente inecuación:

$$\frac{x_3}{x_1+x_2+x_3} \geq a$$

de donde obtenemos:

$$-ax_1-ax_2+(1-a)x_3 \geq 0$$

En la Tabla 3 se muestran las inecuaciones que definen a todos los cuantificadores de la TGQ descritos en la Tabla 2.

3.3 DEFINICIÓN Y MÉTODO DE RESOLUCIÓN DEL PROBLEMA DE OPTIMIZACIÓN EQUIVALENTE

3.3.1 Caso binario proporcional

Una vez obtenido el conjunto de inecuaciones que definen los cuantificadores generalizados, podemos plantear la resolución de un silogismo que involucre a dichos cuantificadores como la resolución de un problema de optimización matemática. Por simplicidad, consideraremos inicialmente el caso en que los cuantificadores presentes en las premisas y la conclusión son todos de tipo proporcional. Las inecuaciones a resolver son, en este caso, las indicadas en la expresión (6)[3]. Por completitud se deben añadir tres restricciones adicionales, relacionadas con asegurar: que no existen conjuntos con un número negativo de elementos,

$$x_k \geq 0, \forall k = 0,...,2^S-1 \quad (7)$$

[3]Debemos considerar que, al tratar con la formulación original 2, los valores x_1, x_2, x_3 aquí indicados denotan en realidad a las sumas de las cardinalidades asociadas a las respectivas particiones disjuntas $L_{n,j}$ en cada proposición

Tabla 3: Conjuntos de inecuaciones que definen los cuantificadores binarios generalizados.

Cuantificadores lógicos	
$algun(Y_1,Y_2)$	$x_0 > 0$
$todos(Y_1,Y_2)$	$x_2 = 0$
$ningun(Y_1,Y_2)$	$x_3 = 0$
$no\ todos(Y_1,Y_2)$	$x_2 > 0$
Cuantificadores binarios absolutos	
$Q_{BA}(Y_1,Y_2)$	$x_3 \geq a$ $x_3 \leq b$
Cuantificadores binarios de excepción	
$Q_{BE}(Y_1,Y_2)$	$x_2 \geq a$ $x_2 \leq b$
Cuantificadores binarios proporcionales	
$Q_{BP}(Y_1,Y_2)$	$-ax_1-ax_2+(1-a)x_3 \geq 0$ $-bx_1-bx_2+(1-b)x_3 \leq 0$
Cuantificadores binarios comparativos	
$Q_{BC-ABS}(Y_1,Y_2)$	$-x_1+x_3 \geq a$ $-x_1+x_3 \leq b$
$Q_{BC-PROP}(Y_1,Y_2)$	$-ax_1+(1-a)x_2+x_3 \geq 0$ $-bx_1+(1-b)x_2+x_3 \geq 0$
Cuantificadores de similaridad	
$Q_S(Y_1,Y_2)$	$-ax_1-ax_2+(1-a)x_3 \geq 0$

que los $L_{n,j}$ son no vacíos, por lo que los sumandos asociados a sus particiones disjuntas $x_1^{n,j}\cdots x_r^{n,j}$ deben verificar, para evitar indefinición en los resultados, que:

$$x_1^{n,j}+\cdots+x_r^{n,j} > 0, \forall n = 1,...,N \quad (8)$$

y que la suma de las cardinalidades es igual al tamaño del referencial:

$$\sum_{k=0}^{K} x_k = |E| \quad (9)$$

Aunque el valor de $|E|$ es arbitrario, en este caso proporcional resulta resulta más conveniente normalizar el problema al intervalo $[0,1]$, suponiendo que los x_k representan las proporciones en tanto por 1 y que ninguno de los conjuntos de $x_r^{n,j}$ es vacío, tomando por tanto en este caso $|E| = 1$.

La sentencia a optimizar es la conclusión del argumento silogístico, que tiene la misma forma que el resto de las premisas, esto es, "$Q_C\ L_{C,1}$ son $L_{C,2}$", donde Q_C es igualmente un cuantificador proporcional binario. Usando una notación similar a la de la expresión (5), resolver el silogismo es equivalente a

$$minimizar\ \frac{x_3}{x_1+x_2+x_3}\ y\ maximizar\ \frac{x_3}{x_1+x_2+x_3}$$
$$sujeto\ a\ las\ restricciones\ (6),(7),(8),(9)$$

Este problema de optimización se resuelve directamente utilizando técnicas de *linear fractional-programming* como las descritas en [13].

3.3.2 Caso binario absoluto con *N* premisas

Como segundo ejemplo, consideramos el caso de un silogismo como el descrito en la expresión (2), para el caso de cuantificadores absolutos. En este caso podemos prescindir de la restricción de trabajar en el universo de proporciones, y por tanto los x_k se definen ya en términos de cantidades absolutas. Si denotamos por $Q_n : [a_n, b_n]$ el cuantificador absoluto asociado a la premisa *PRn*, resolver un silogismo como el descrito en la expresión es equivalente a:

$$minimizar\ x_{C,k_1} + \cdots + x_{C,k_C} \quad (10)$$

$$maximizar\ x_{C,k_1} + \cdots + x_{C,k_C} \quad (11)$$

sujeto a las restricciones

$$x_{1,k_1} + \cdots + x_{1,k_1} \geq a_1; \qquad x_{1,k_1} + \cdots + x_{1,k_1} \leq b_1 \quad (12)$$

$$\cdots$$

$$x_{N,k_1} + \cdots + x_{1,k_N} \geq a_N; \qquad x_{N,k_N} + \cdots + x_{N,k_N} \leq b_N \quad (13)$$

$$x_{i,j} \geq 0 \qquad \forall i, j \quad (14)$$

donde los x_{n,k_1} ,..., x_{n,k_n} son las propiedades disjuntas correspondientes a los respectivos conjuntos $|L_{n,1} \cap L_{n,2}|$ de la premisa *PRn*, n=1,...,N.

Este problema de optimización puede resolverse directamente por el método SIMPLEX [13].

3.3.3 Nuevos tipos de silogismos: TGQ y borroso

Mediante un planteamiento similar es posible considerar el resto de cuantificadores descritos en la TGQ, entre ellos diversos casos no tratados hasta el momento en la bibliografía, como los cuantificadores comparativos (como "*hay alrededor del doble de As que de Bs*"), de excepción ("*todos menos 3 As son Bs*") o de similaridad ("*A es similar a A' en un factor 0.95*"). Las expresiones a optimizar son, como en los ejemplos anteriores, las correspondientes a la conclusión, que se describirán en términos de la suma de cardinalidades de las propiedades disjuntas que correspondan. Las restricciones se toman, para cada una de las premisas, de las correspondientes de la Tabla 3, añadiéndole también como restricciones las expresiones (7), (8) y (9).

Adicionalmente, el caso borroso puede tratarse mediante el procedimiento indicado utilizando la aproximación de Dubois y Prade [4], que describe los números borrosos trapezoidales de núcleo *Ker* y Soporte *Sup* como un par de intervalos $\{Ker, Sup\}$. El planteamiento de este problema consiste en considerar para cada una de las sentencias (tanto las premisas como la conclusión), dos problemas asociados, uno considerando el núcleo y otro el soporte del cuantificador correspondiente. Así, para una premisa *PRn*, tenemos que el cuantificador borroso Q_n se define como $Q_n := \{Ker_{Q_n}, Sup_{Q_n}\}$. La resolución del silogismo completo consiste en este caso, en plantear la resolución de dos problemas intervalares crisp similares, uno de ellos tomando $Q_n := Ker_{Q_n}$ y otro tomando $Q_n := Sup_{Q_n}$, de modo que se obtiene para cada uno de ellos una conclusión silogística diferente. Ambas conclusiones se integran en una única solución correspondiente al cuantificador borroso Q_C de la conclusión del problema original de la siguiente manera: Ker_{Q_C} vendrá dado por la resolución del silogismo con $Q_n := Ker_{Q_n}$, n=1,...,N y Sup_{Q_C} vendrá dado por la resolución del silogismo con $Q_n := Sup_{Q_n}$, n=1,...,N. En esta aproximación puede darse el caso de que se obtenga solución para el problema de soportes y no para el de núcleos. La solución sería en esta situación un número borroso no normalizado, cuyo cálculo puede aproximarse de forma más precisa utilizando la operación de α-cortes. Para ello, basta resolver el silogismo considerando diferentes α-cortes de los cuantificadores, de modo que para $\alpha = 0$ tenemos el problema $Q_n := Sup_{Q_n}$ y para $\alpha = 1$ tenemos el problema $Q_n := Ker_{Q_n}$. El mayor de los α-cortes (α) para el que exista solución en el problema de optimización definirá el intervalo de máxima pertenencia de Q_C, precisamente con valor de pertenencia α.

Tabla 4: Silogismo con cuantificadores intervalares.

*PR*1 :	Entre el 85 % y el 95 % de los estudiantes son gallegos
*PR*2 :	Entre el 25 % y el 35 % de los gallegos son estudiantes
*PR*3 :	Entre el 60 % y el 80 % de los gallegos son altos
*PR*4 :	Entre el 90 % y el 100 % de los estudiantes son altos
C :	Q_C estudiantes gallegos son altos

4 EJEMPLO DE SILOGISMO Y SU RESOLUCIÓN

Para ilustrar nuestra propuesta, a continuación desarrollamos el análisis de una argumentación silogística para un ejemplo con cuantificadores proporcionales.

En el silogismo de la Tabla 4, sean las propiedades $estudiantes = P_1$, $gallegos = P_2$ y $altos = P_3$, los cuantificadores de las premisas $Q_1 = [0.7, 0.9]$, $Q_2 = [0.4, 0.6]$, $Q_3 = [0.8, 0.85]$ y $Q_4 = [0.7, 0.9]$ y el cuantificador de la conclusión $Q_C = [a, b]$.

Como hemos indicado, el primer paso consiste en transformar las premisas en el correspondiente conjunto de inecuaciones. En este caso, todos los cuantificadores son binarios

proporcionales:

$$PR1: \frac{x_6+x_7}{x_4+x_5+x_6+x_7} \geq 0.85; \quad \frac{x_6+x_7}{x_4+x_5+x_6+x_7} \leq 0.95$$
$$PR2: \frac{x_6+x_7}{x_2+x_3+x_6+x_7} \geq 0.25; \quad \frac{x_6+x_7}{x_2+x_3+x_6+x_7} \leq 0.35$$
$$PR3: \frac{x_7+x_3}{x_2+x_3+x_6+x_7} \geq 0.60; \quad \frac{x_7+x_3}{x_2+x_3+x_6+x_7} \leq 0.80$$
$$PR4: \frac{x_5+x_7}{x_4+x_5+x_6+x_7} \geq 0.90; \quad \frac{x_5+x_7}{x_4+x_5+x_6+x_7} \leq 1$$
$$C: \frac{x_7}{x_6+x_7} \geq a \quad ; \quad \frac{x_7}{x_6+x_7} \leq b$$

Aplicando las transformaciones correspodientes a las anteriores inecuaciones, obtenemos:

$$\begin{aligned}
PR1: & x_6+x_7-0.85x_4-0.85x_5-0.85x_6-0.85x_7 \geq 0;\\
& x_6+x_7-0.95x_4-0.95x_5-0.95x_6-0.95x_7 \leq 0\\
PR2: & x_6+x_7-0.25x_2-0.25x_3-0.25x_6-0.25x_7 \geq 0;\\
& x_6+x_7-0.35x_2-0.35x_3-0.35x_6-0.35x_7 \leq 0\\
PR3: & x_7+x_3-0.6x_2-0.6x_3-0.6x_6-0.6x_7 \geq 0;\\
& x_7+x_3-0.8x_2-0.8x_3-0.8x_6-0.8x_7 \leq 0\\
PR4: & x_5+x_7-0.9x_4-0.9x_5-0.9x_6-0.9x_7 \geq 0;\\
& x_5+x_7-1x_4-1x_5-1x_6-1x_7 \leq 0.9
\end{aligned}$$

Estas inecuaciones tienen la forma adecuada para poder optimizar las inecuaciones de C mediante el método *linear fractional-programming*, obteniendo como resultado $Q = [0.88, 1]$, que puede ser interpretada como "*El* 88 *% o más de los estudiantes gallegos son altos*".

5 CONCLUSIONES

En el trabajo se ha planteado un esquema de razonamiento silogístico aproximado, sin limitación en el número de premisas y compatible con la TGQ. Para varios casos de interés con cuantificadores binarios se ha demostrado como es posible transformar el problema de razonamiento en un problema equivalente de optimización, cuya resolución proporciona las conclusiones del silogismo.

Como trabajo futuro, se plantea completar este esquema con la incorporación de nuevos casos nítidos (en particular los cuantificadores ternarios y cuaternarios), y ampliarlo para considerar en toda su extensión el problema de cuantificadores borrosos siguiendo la aproximación de α-cortes aquí apuntada, la consideración conjunta en un mismo esquema silogístico de distintos tipos de cuantificadores y la implementación de todo el proceso de razonamiento silogístico en una librería que facilite su aplicación experimental.

Agradecimientos

Este artículo fue realizado con el apoyo Ministerio de Ciencia e Innovación (Proyectos TIN2008-00040 y TIN2011-29827-C02-02) y del Ministerio de Educación (Becas del Programa FPU).

Referencias

[1] Aristóteles. *Organon*. Gredos, Madrid, 1982-1986.

[2] P. L. Peterson. *Intermediate Quantities. Logic, linguistics, and Aristotelian semantics*. Ashgate, Alsershot. England, 2000.

[3] L. A. Zadeh. Syllogistic reasoning in fuzzy logic and its applications to usuality and reasoning with dispositions. *IEEE Transactions On Systems, Man and Cybernetics*, 15(6):754–765, 1985.

[4] D. Dubois and H. Prade. On fuzzy syllogisms. *Computational Intelligence*, 4(2):171–179, 1988.

[5] D. Dubois, H. Prade, and J-M. Toucas. *Intelligent Systems. State of the Art and Future Directions*, chapter Inference with imprecise numerical quantifiers, pages 52–72. Ellis Horwood, Great Britain, 1990.

[6] D. Dubois, L. Godo, R. López de Mántaras, and H. Prade. Qualitative Reasoning with Imprecise Probabilities. *Journal of Intelligent Information Systems*, 2:319–363, 1993.

[7] V. Novák. A formal theory of intermediate quantifiers. *Fuzzy Sets and Systems*, 159:1229–1246, 2008.

[8] D. G. Schwartz. Dynamic reasoning with qualified syllogisms. *Artifical Intelligence*, 93:103–167, 1997.

[9] F. Sommers. *The Logic of Natural Language*. Clarendon Press, Oxford, 1982.

[10] M. Pereira-Fariña, F. Díaz-Hermida, and A. Bugarín. An analysis of reasoning with quantifiers within the Aristotelian syllogistic framework. In *Proc. 2010 International Fuzzy Conference on Fuzzy Systems (FUZZ-IEEE 2010)*, pages 635–642. IEEE, 2010.

[11] Y. Liu and E. E. Kerre. An overview of fuzzy quantifiers (ii). reasoning and applications. *Fuzzy Sets and Systems*, 95:135–146, 1998.

[12] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.

[13] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, The Edinburgh Building, Cambridge, UK, 2004.

ON THE FUZZINESS OF SCIENTIFIC DISCIPLINES – USING THE EXAMPLE OF COMPUTER SCIENCE

Rudolf Seising

European Centre for Soft Computing
Edificio de Investigación
Calle Gonzalo Gutiérrez Quirós S/N
33600 Mieres, Asturias,
rudolf.seising@softcomputing.es

Abstract

Using the example of the scientific discipline of Computer science (CS) this paper deals with the view of scientific disciplines as fuzzy concepts. Historically, Computer science emerged from electrical engineering in the 1960s. From the content of some unpublished documents and also some rather non-famous papers by Lotfi A. Zadeh it is argued that the emergences of Computer science and his theory of Fuzzy Sets have been historically interlinked. Zadeh's task as Chair of the Electrical Engineering Department in Berkeley in the 1960s, his activities in Education of Engineering and his creation of the theory of Fuzzy sets generated his view on the scientific discipline of Computer science as a fuzzy set. This view could establish a new approach to history and philosophy of science.

Keywords: Philosophy of Science, Fuzziness, Science, Scientific Disciplines, Electrical Engineering, Computer Science

1 INTRODUCTION

The world is like Neapolitan ice cream – as the British writer Aldous Huxley (1894-1963) once put it – and its physical, biological, sociological and moral spheres correspond to the layers of chocolate, strawberry and vanilla. Ludwig von Bertalanffy (1901-1972), the founder of *General System Theory*, quoted this witticism when he considered a "unitary conception of the world that may be based, not upon the possibly futile and certainly farfetched hope finally to reduce all levels of reality to the level of physics, but rather on the isomorphy of laws in different fields". ([3], p. 8.) Bertalanffy was not pleading for a reductionism but for a "perspectivism": "We cannot reduce strawberry to chocolate – the most we can say is that possibly in the last resort, all is vanilla, all mind or spirit." ([3], p. 8.) In the face of the different scientific disciplines he wrote: "We can not reduce the biological, behavioral, and social levels to the lowest level, that of the constructs and laws of physics. We can, however, find constructs and possibly laws within the individual levels. ([3], p. 8.)

The unified principle was more the organizational structure of the individual areas which join together to form a whole, and this principle aimed to find it (fig. 1).

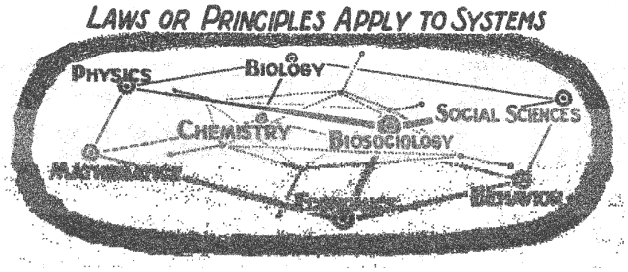


Figure 1: Illustration of Bertalanffy's *General System Theory* in [3].

In figure 1 we see only the "classical" scientific disciplines, e.g. on the left s we have the natural sciences physics, mathematics, chemistry and biology but we will consider in this paper some aspects of electrical engineering (EE) and of the "new" field of computer science (CS) that accrued from EE after the computer has been conceived and built. However, when discussing science and its disciplines in Spain, perhaps we should compare it with what is called "la paella (fig. 2):

02 La paella auténtica

VIERNES 30 SEPTIEMBRE DE 2011 Qué!

LOS PLATOS MÁS TÍPICOS LUCHAN POR MANTENER SU ESENCIA

No me toques la paella

Los mejores chefs homenajean en Valencia a este plato universal. Aceptan las novedades, pero sin prescindir de la receta tradicional. La misma batalla libran el cocido madrileño, la fabada, la tortilla...

"La paella valenciana es la de pollo y conejo y no lleva ni cebolla ni pimiento"
RAFAEL BLASCO/Centro de Turismo Valencia

"Hay evolución, pero el chorizo en la paella es una aberración"
SENTO ALEJANDRE/Restaurador

"En la huerta le ponemos pato y la hacemos con leña"
TONI MONTOLIU/Restaurante Barraca

"Como mínimo tarda hora y media o dos horas en hacerse. Si no, no es paella de verdad"

¿Cómo es para ti la paella ideal? Danos tu opinión en www.que.es

LOS 'GURÚS' DE LA TRADICIÓN NOS CUENTAN CÓMO CUIDAR OTROS PLATOS TÍPICOS

Figure 2: La paella. Excerpt from *Que es?*, 30.9.2011.

As in the article of the newspaper *Que es*? at September 30, 2011 was argued, there is – perhaps – no strict definition on how to create the paella: "La paella auténtica no existe, es un mito". – The authentic paella does not exist, it is a myth! To say it in terms of scientific language: the concept of the paella is fuzzy.

In the paper at hand I will show that we are dealing with fuzzy concepts not only in our culinary life and – of course in other parts of our daily life, too – but also in science. My example in this paper is the scientific discipline of computer science (CS) that emerged from the scientific discipline of electrical engineering (EE) in the 1960s. In chapter 2 I will sketch some aspects of that historical episode from the viewpoint of the electrical engineer Lotfi A. Zadeh who was on the one hand a decisive protagonist of this development at the University of California, Berkeley and who founded on the other hand the theory of fuzzy sets (FS) during that period. From non-published or not well known documents that I found in Zadeh's private archives in his office and at his home I will argue that the two developments in science, the birth of CS and FS, resp. have been interlinked, viz.:

- FS emerged from scientific studies in Circuit theory, Network theory, System theory, and Information theory as parts of EE, as well as from mathematical theories and logic, after computers entered the field of technology.
- We can treat the appearance of CS from EE and other scientific theories and methodologies as a fuzzy relationship because the subjects of CS – some of them are also subjects of EE or mathematics or logic etc. and some of them were new created – became part of the new scientific discipline CC only to some degree and CS can be considered as a fuzzy set!

Chapter 3 deals with this fuzziness of CS in the 1960s. The outlook gives a short remark to consider the fuzziness of scientific disciplines in general.

2 EE AND THE COMPUTER ERA

Computers have been the most famous technical product of Second World War research − next to the atomic bomb. Computers became popular as "electronic brains" or "thinking machines". Even the "era of computers" started already with the analogue *MIT Differential Analyzer* of Vannevar Bush (1890-1974), the technological development of digital computers was an eminent push with ENIAC (*Electronic Numerical Integrator and Computer*) and EDVAC (*Electronic Discrete Variable Computer*), both designed by John Presper Eckert (1919-1995) and John William Mauchly (1907-1980).

2.1. COMPUTERS: A NEW FIELD IN EE

Already in 1950 the British mathematician Alan Mathison Turing (1912-1954) asked "Can machines think?"in his famous *Mind*-article "Computing Machinery and Intelligence" [15]. Instead of an answer he proposed the well-known imitation game, now called the "Turing test", that became of philosophical interest to decide whether a computer or a program could think like a human being or not. In those days Lotfi Aliasker Zadeh (born 1921) was a young electrical engineer with deep interest in these new computing machines. He wrote the paper "Thinking Machines. A New Field in Electrical Engineering", which appeared in *The Columbia Engineering Quarterly* in the same year [19]. Here, Zadeh put up for discussion the questions "How will 'electronic brains' or 'thinking machines' affect our way of living?" and "What is the role played by electrical engineers in the design of these devices?" ([19], p. 12.) He was looking for "the principles and organization of machines which behave like a human brain. Such machines were then variously referred to as "thinking machines", "electronic brains", "thinking robots", and similar names.

The fundamental principles of thinking machines, Zadeh stressed, were developed by mathematicians, but the physical realization, the construction of the thinking machines, was the task of electrical engineers, who design and build the memory chips, processors, "computors", decision makers, etc. Until now, if electrical engineers had come into contact at all with such heretofore far-removed subjects as Boolean algebra, polyvalent logic, etc., it was through so far off when it would be just as important for post-graduate electrical engineers to take classes in mathematical logic as classes on complex variables: "Time marches on." ([19], p. 31)

2.2. FROM CIRCUITS TO SYSTEMS

Also in the *Columbia Engineering Quarterly* Zadeh published in 1954 the article "System Theory" [21] where, he characterized systems as a "black boxes" with inputs and outputs, and in the case that these inputs and outputs are describable as time dependent functions then the dynamic behaviour of the system can be studied mathematically,

and the input-output-relationship of the system says that the output is just a function of the input.

In the early 1950s, engineers were, in general, inadequately trained to think in abstract terms, but nevertheless, Zadeh believed that it was only a matter of time before system theory would attain acceptance. It turns out that he was right: Eight years later, he wrote the article "From Circuit Theory to System Theory" [24] for the anniversary edition of the *Proceedings of the IRE* appeared in May 1962 to mark the 50th year of the *Institute of Radio Engineers* (IRE). He could describe problems and applications of system theory and its relations to network theory, control theory, and information theory: "It has been brought about, largely within the past two decades, by the great progress in our understanding of the behavior of both inanimate and animate systems-progress which resulted on the one hand from a vast expansion in the scientific and technological activities directed toward the development of highly complex systems for such purposes as automatic control, pattern recognition, data-processing, communication, and machine computation, and, on the other hand, by attempts at quantitative analyses of the extremely complex animate and man-machine systems which are encountered in biology, neurophysiology, econometrics, operations research and other fields" ([24], p. 856f.).

After 1959, when Zadeh became a professor of electrical engineering at the University of California, Berkeley, he published papers on system theory and two well known books with colleagues at his EE department [26], [31].

2.3. ELECTRICAL FILTERING

Already in his time in New York, when he joined the faculty of Columbia University as an instructor and since 1950 he was appointed to an assistant professor, he was enhanced by *Cybernetics* [17] of Norbert Wiener (1894-1964) and he was interested in the theory of ideal and optimal filtering. Together with his supervisor John Ralph Ragazzini (1912-1988) he published in 1950 "An Extension of Wiener's Theory of Prediction" [18] and this article was an important milestone in the development of network synthesis.

The mathematical techniques of the theory of prediction and filtering have been commonly employed in mathematical physics, particularly in quantum mechanics, e.g. multi-dimensional Euclidean spaces and Hilbert space representation. In his many papers on prediction and filtering, linear and nonlinear systems, time-variying networks, etc. in the first 1950s, Zadeh used the mathematical calculus of functionals and operators. The analogy between projection in a function space and filtration with an ideal filter led Zadeh in the early 1950s to a functional symbolism of filters [20]. In [20] Zadeh introduced the concept of optimal filters contrary to ideal filters. Ideal filters are defined as filters which achieve a perfect separation of signal and noise but in reality there are no ideal filters. He knew from experience that characteristics of electrical filters don't show an exact step at the limiting frequency but smooth functions. Zadeh regarded optimal filters that give the "best approximation" of a signal and he noticed that "best approximations" depend on reasonable criterions. In that time he formulated these criterions in statistical terms, but later he had turned away from statistical methods during the course of his work on optimum filters in the mid-50s, and he recognized that a more promising approach was that of finding an optimum filter relative to a distance to be minimized in the function space of the signals.

In that time, Zadeh was working constructively to bridge the gap between theory and practice but he was forced to recognize, however, that these attempts would not be successful: "As a mathematically oriented system theorist, I had been conditioned to believe that the analytical tools based on set theory and two-valued logic are all that is needed to build a framework for a precise, rigorous and effective body of techniques for the analysis of almost any kind of man-made or natural system. Then, in 1961-1963, in the course of writing a book on system theory (with C. A. Desoer), I began to feel that complex systems cannot be dealt with effectively by the use of conventional approaches largely because the description languages based on classical mathematics are not sufficiently expressive to serve as a means of characterization of input-output relations in an environment of imprecision, uncertainty and incompleteness of information." [9]

There were two ways to overcome this situation. In order to describe the actual systems appropriately, he could try to increase the mathematical precision even further, but Zadeh failed with this course of action. The other way presented itself to Zadeh in the year 1964, when he discovered how he could describe real systems as they appeared to people. "I'm always sort of gravitated toward something that would be closer to the real world" [10]. What was closer to the real world was the Theory of Fuzzy Sets (FS)!

2.4. MATHEMATICS AND LOGIC

Zadeh's resignation with usual mathematics would soon lead to capitulation. He was nearing a crossroads. Herbert E. Robbins (1915-2001) (fig. 3) was the chairman of Columbia University's department of mathematical statistics at the time. He was good friends with Lotfi Zadeh as well as with Deane Montgomery (1909-2002) (fig. 3), a member of the *Institute for Advanced Study* (*IAS*) in Princeton. Robbins and Montgomery campaigned for the approval of the *IAS* guest residency for which Zadeh had applied, even though it was rare for requests by scientists who were neither mathematicians nor theoretical physicists or historians to receive a positive response [11]. Zadeh initially took a half-year sabbatical from *Columbia University* in 1956. He wanted to learn more about logic, an interest he had cultivated since 1950, when he pre-

dicted that logic, and particularly multi-valued logic, would become increasingly more important to the problems of electrical engineering in the future. [19]

Figure 3: Robbins, Montgomery, Kleene.

Princeton inspired Zadeh very quickly; it was the "Mecca for mathematicians" he said [13]. He attended lectures by Steven Kleene (1909-1994) (fig. 4), who had also continued developing the multi-valued logic devised by the Polish school of logic. Kleene became Zadeh's friend and mentor at Princeton: "Steven Kleene was my teacher in logic. Yes, I learned logic from Steven Kleene!" [9].

This residency had revealed to Zadeh some completely new perspectives of scientific life and work. New ways of thinking had come from the mathematics philosophers in Princeton thanks to the mathematicians there he had learned new mathematical methods from statistics, from game and decision theory. He also experienced new views of system theory and the newly established automata theory. He had apparently become familiar with *Automata Studies*, published during this period by Shannon and John McCarthy (born 1927). All of the knowledge, impetus and impressions he found at the IAS would have a lasting effect on Zadeh's future endeavors!

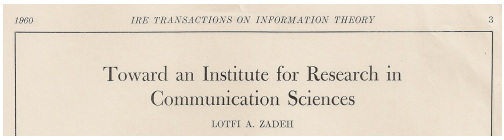
Zadeh found multi-valued logic to be a natural generalization of the conventional logic of just two values into *n* values, similar to the leap from two-dimensionality to *n*-dimensionality in mathematics. [11] He was now also toying with the idea of introducing multi-valued logic into automata theory and implementing it in electric circuits, and once he had returned to Columbia University in New York he assigned two dissertations that dealt with the subjects of multi-valued logic in the design of transistor circuits and with multi-valued coding:

- Oscar Lowenschuss wrote the dissertation *Multi-Valued Logic and Sequential Machines or Non-Binary Switching Theory* the following year: [4]. Parts of this paper had been published previously: [5] See also the later publication [6].
- Werner Ulrich managed to finish his dissertation *Nonbinary Error Correction Codes* in 1957: [16]

"That's why I wanted to know about logics!" Zadeh recalls [9].

2.5. INFORMATION AND COMMUNICATION

The value to that scientific community which Zadeh recognized in an institute like *IAS* in Princeton can also be seen in his dedication to establishing a similar institute for his own scientific community. Since the late 1940s he was enhanced by the "Mathematical Theory of Communication" (later called for short *Information theory*) [13] of Claude E. Shannon (1916-2001) and when he once again composed an editorial for the *IRE Transactions on Information Theory* in 1960, it was entitled "Toward an Institute for Research in Communication Science" [23] (fig. 4). He had in the mean time become a professor at Berkeley and he apparently sometimes longed for the freedom he had enjoyed as a guest scientist in Princeton. He called for an *Institute for Communication Science* to be founded where scientists could spend a year or two concentrating exclusively on their research without being distracted by teaching and administrative duties, contract negotiations and doctoral advising. This was the only way to guarantee a free choice of research topics and scientifically communicative exchange with no outside pressure!


1960 IRE TRANSACTIONS ON INFORMATION THEORY 3

Toward an Institute for Research in Communication Sciences

LOTFI A. ZADEH

Figure 4: Zadeh's editorial in 1963, [23], excerpt.

A number of well-known institutions both in the United States and abroad have these characteristics, but they are embodied perhaps in their purist form in the *IAS*, which since its inception in 1930 had played a very significant role in the development of mathematics in their country [23]. The *IAS*, Germany's *Max-Planck-Institutes* or the *Institute for Automatics and Telemechanics* in the Soviet Union did not necessarily have to serve as models for the institute he hoped to establish. Rather it would be designed to meet the specific needs and interests of workers in the fields of information theory, communication theory, system theory, control theory, automata, biological systems, computation, machine translation of languages and related fields. It would be concerned with both theoretical and experimental research in this area [23].

After Zadeh had also joined the editorial board of the *IRE Transactions on Automatic Control* in 1962, he wrote "A Critical View of Our Research in Automatic Control" for the April edition, an article in which he repeated his call for the founding of such an institute. Here he suggested establishing an *Institute for Control Science and Engineering*, then he added: "(or, more broadly, an *Institute for Research in Information Sciences*) which would serve as a focal point on the national level for research in control theory and its applications as well as in such related fields as system theory, information and communication theories, circuit theory, machine computation, automata theory, bionics, etc." ([25], p. 74).

The many new impressions, encounters, discussions and insights in Princeton had steered Zadeh's future scientific work in new directions.

2.6. FUZZY SETS

As we have seen before, Zadeh had extensively criticized the relationship between mathematics and his own scientific-technical EE discipline. The tool offered by mathematics was not appropriate to the problems that needed to be handled in the engineering sciences. Information and communications technology had led to the construction and design of systems that were so complex that it took much more effort to measure and analyze these systems than had been the case just a few years before. Much more exact methods were now required to identify, classify or characterize such systems or to evaluate and compare them in terms of their performance or adaptivity.

In order to provide a mathematically exact expression of experimental research with real systems, it was necessary to employ meticulous case differentiations, differentiated terminology and definitions that were adapted to the actual circumstances, things for which the language normally used in mathematics could not account. The circumstances observed in reality could no longer simply be described using the available mathematical means.

In the summer of 1964 Zadeh was thinking about pattern recognition problems and grades of membership of an object to be an element of a class as he returned to mind almost 50 years later: "While I was serving as Chair, I continued to do a lot of thinking about basic issues in systems analysis, especially the issue of unsharpness of class boundaries. In July 1964, I was attending a conference in New York and was staying at the home of my parents. They were away. I had a dinner engagement but it had to be canceled. I was alone in the apartment. My thoughts turned to the unsharpness of class boundaries. It was at that point that the simple concept of a fuzzy set occurred to me. It did not take me long to put my thoughts together and write a paper on the subject. This was the genesis of fuzzy set theory." ([33], p. 7).[1]

Zadeh submitted his first article "Fuzzy Sets" to the editors of *Information and Control* in November 1964 and it appeared in this journal in the following June [27]. He introduced new mathematical entities as classes or sets that ''are not classes or sets in the usual sense of these terms, since they do not dichotomize all objects into those that belong to the class and those that do not.'' He introduced ''the concept of a fuzzy set, that is a class in which there may be a continuous infinity of grades of membership, with the grade of membership of an object x in a fuzzy set A represented by a number $f_A(x)$ in the interval [0,1].'' [27]

[1] A detailed presentation of the history of the theory of FS is given in the author's book [12].

The question was how to generalize various concepts, union of sets, intersection of sets, and so forth. Zadeh defined equality, containment, complementation, intersection and union (fig. 5) relating to fuzzy sets A, B in any universe of discourse X as follows (for all $x \in X$):

- $A = B$ if and only if $\mu_A(x) = \mu_B(x)$,
- $A \subseteq B$ if and only if $\mu_A(x) \leq \mu_B(x)$,
- $\neg A$ is the complement of A,

  if and only if $\mu_{\neg A}(x) = 1 - \mu_A(x)$,
- $A \cup B$ if and only if $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$,
- $A \cap B$ if and only if $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$.

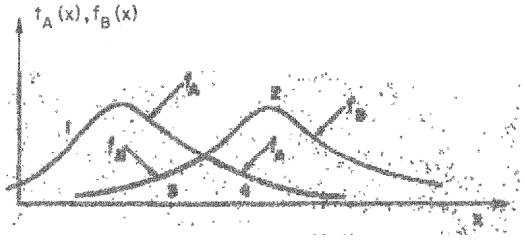
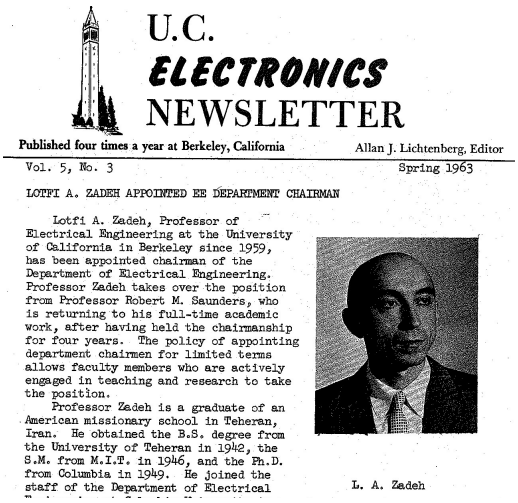


Figure 5: Zadeh's Illustration of fuzzy sets in R[1]: "The membership function of the union is comprised of curve segments 1 and 2; that of the intersection is comprised of segments 3 and 4 (heavy lines)." ([27], p. 342).

3 FROM EE AND CS TO EECS

In one of my interviews Zadeh recalled: "System Theory came grown up but then computers and computers then took over. In other words: the center of attention shifted. ... So, before that, there were some universities that started departments of system sciences, departments of system engineering, something like that, but then they all went down. They all went down because computer science took over." [10]


U.C.
ELECTRONICS
NEWSLETTER

Published four times a year at Berkeley, California — Allan J. Lichtenberg, Editor

Vol. 5, No. 3 — Spring 1963

LOTFI A. ZADEH APPOINTED EE DEPARTMENT CHAIRMAN

Lotfi A. Zadeh, Professor of Electrical Engineering at the University of California in Berkeley since 1959, has been appointed chairman of the Department of Electrical Engineering. Professor Zadeh takes over the position from Professor Robert M. Saunders, who is returning to his full-time academic work, after having held the chairmanship for four years. The policy of appointing department chairmen for limited terms allows faculty members who are actively engaged in teaching and research to take the position.

Professor Zadeh is a graduate of an American missionary school in Teheran, Iran. He obtained the B.S. degree from the University of Teheran in 1942, the S.M. from M.I.T. in 1946, and the Ph.D. from Columbia in 1949. He joined the staff of the Department of Electrical

L. A. Zadeh

Figure 6: Zadeh became chair in 1963, (excerpt).

When Zadeh had become chairman of the department of electrical engineering (EE) at Berkeley in 1963 (fig. 6), he experienced these shifts very intensively, for it was during his five-year tenure in this position that his department was renamed the *Department of Electrical Engineering and Computer Science* (EECS) [10]. In November 1992 Zadeh was asked to give an after dinner talk on the history of CS at Berkeley. In the manuscript for this talk he wrote: "I joined the EE Department in 1959. At that time, the EE Department was best known for its work in electromagnetism and microwaves. A decision was made to build up the area of circuits, systems and control. Don Pederson, Ernie Kuh and Charles Desoer came from Bell Labs. Eli Jury, Art Bergen and I came from Columbia University. […] There wasn't much activity in the computer field at that time, but there was was significant. […] There was a Computer Centre in Cory Hall that was run by the EE Department. The principal and only figures in computer science and engineering in EE at that time were Paul Morton and Harry Huskey. They can be rightly regarded as the progenitors of Computer science and engineering at Berkeley." […] When I was appointed as Chairman in 1963, I was not a computer person and I am not a computer user to this day, I regret to say. But I was always a very strong believer in the importance of computers and digital technology. My first action as Chairman was to send a memo to the faculty in which I suggested that we assign the highest priority to the development of computer science in EE. But what it was obvious today was not so obvious then. The reaction to my memo was mixed and some influential faculty members objected strongly to my proposal." [32]

After some difficulties, Zadeh was finally successful with his request to change the name of the department to EECS. In his article "Electrical Engineering at the Crossroads" that he published in 1965 [28] (fig. 7) he described the problem that way: "The slowness with which many electrical engineering departments have reacted to the rapidly growing demand for computer scientists and engineers, and their unwillingness to make substantive changes in their curricula to meet the need for specialized training in computer sciences and related fields, is generating strong pressures on some campuses to establish separate computer science departments." ([28], p. 30) In the same article he presented the new EE curriculum at Berkeley that "reflects the fact that, today, electrical engineering is no longer an aggregation of a small number of subject areas sharing a large common body of concepts and techniques – as it was in the thirties, forties, ant to a lesser extent, in the fifties. Rather, it is an assemblage of a wide range of subjects, falling into three major areas which have a relatively small common core. […] If this premise is accepted, then the only logical conclusion is that the student must be provided with a choice of several basic programs, which could permit him to focus his studies in one of the major areas falling within or nearest to his main field of interest." ([28], p. 31)

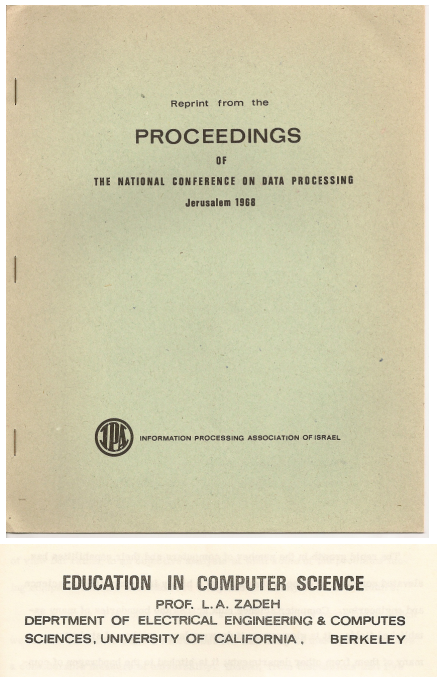
Electrical Engineering at the Crossroads

L. A. ZADEH, FELLOW, IEEE

Figure 7: Excerpt of Zadeh's article in1965, [28].

At Berkeley they established the three programs *A* (Electronics, Fields, and Plasmas), *B* (Systems, Information, and Control) and *C* (Computer Sciences) and a "General EE program" *D* ([28], p. 31f)

Three years later Zadeh gave a talk on "Education in Computer Science" at Israel's *4th National Conference on Data Processing* that took place in the Hebrew University Jerusalem (fig. 8). Initially, he claimed that computer science "as a collection of concepts and techniques which serve to systematize the employment of the means with which modern technology provides us for purposes of stage, representation and processing of information." ([29], p. E157) In the printed version of this talk he emphasized that "Computer science cuts across the boundaries of many established fields. It is glamorous; it draws a large number of students – many of them from other departments; it is hitched to the bandwagon of computers and the information revolution." ([29], p. E158)


Reprint from the

PROCEEDINGS

OF

THE NATIONAL CONFERENCE ON DATA PROCESSING

Jerusalem 1968

INFORMATION PROCESSING ASSOCIATION OF ISRAEL

EDUCATION IN COMPUTER SCIENCE

PROF. L. A. ZADEH

DEPRTMENT OF ELECTRICAL ENGINEERING & COMPUTES SCIENCES, UNIVERSITY OF CALIFORNIA, BERKELEY

Figure 8: Reprint of Zadeh's paper in1968, [29].

In this paper he discussed some details of "Berkeley's solution" of the CS-problem that bred finally a Department of CS in the College of Letters and Science and a program in CS in the College of Engineering within the Department of EECS. By way of explanation he added:

"Essentially, the main premise of Berkeley's "solution" is that computer science is not a homogeneous and unified field – at least not at this time – and that, in paraphrased words of Professor A. Oettinger of Harvard, [2] "… it has some components which are the purest of mathematics and some that are the dirtiest of engineering." This split personality of computer science makes it very difficult to create a single academic unit within the university structure where mathematically oriented automata theorists, formal language experts, numerical analysts and logicians could establish a comfortable modus vivendi with non-mathematical oriented hardware designers, systems programmers and computer architects. […] In essence, the Berkeley "solution" provides a partial answer to the dilemma by dividing computer science not into two non-overlapping parts but into two overlapping parts which differ from one another mainly in degrees of emphasis places on various subject areas." ([29], p. E164f)

In his article "Computer Science as a Discipline", that appeared in the same month in the *Journal of Engineering Education* [30] (fig. 9), he grasped CS also that way. He wrote what was "pointed out in the authoritative reports of the *ACM Curriculum Committee on Computer Science*[3] computer science is not simply concerned with the design of computing devices, nor is it just the art of numerical calculation. Essentially, computer science is concerned with information in much the same way that physics is concerned with energy; it is devoted to the representation, storage, manipulation and presentation of information in an environment permitting automatic information systems." ([30], p. 913)

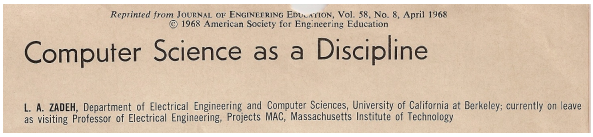


Figure 9: Excerpt of Zadeh's article [30], 1968.

Concerning his view that CS "cuts across the boundaries of many established fields" and also that the parts of CS differ from one another "in degrees of emphasis", Zadeh then presented a new idea in this field of CS education: because of the broadness and vagueness of this statement "to convey even a rough idea of the boundaries of computer science and its relation to mathematics, electrical engineering and other neighboring fields", he introduced his idea of fuzzy sets, to employ a new approach: "Specifically, let us regard computer science as a name for a fuzzy set of subjects and attempt to concretize its meaning by associating with various subjects their respective degrees of containment (ranging from 0 to 1) in the fuzzy set of computer science. For example, a subject such as "programming languages" which plays a central role in computer science, will have a degree if containment equal to unity. On the other hand, a peripheral subject such as "mathematical logic" will have a degree of containment of, say, 0.6." ([30], p. 913.)

Under the heading "Containment Table for Computer Science" (fig. 9) he arranged the most relevant "subjects in question and their degrees of containment in computer science".

TABLE 1
Containment Table for Computer Science

Subject	Degree of Containment in Computer Science
Programming languages	1
Computer design and organization	1
Data structures	1
Models of computation	1
Operating systems	1
Programming systems	1
Formal languages and grammars	0.9
Computational linguistics	0.8
Automata theory	0.8
Finite-state systems	0.8
Theory of algorithms	0.9
Discrete mathematics	0.8
Mathematical logic	0.6
Combinatorics and graph theory	0.8
Dynamic programming	0.7
Mathematical programming	0.7
Numerical methods	0.8
Switching theory	0.8
Analog and hybrid computers	0.7
Computer graphics	0.7
Digital devices and circuits	0.7
Artificial intelligence and heuristic programming	0.9
Information retrieval	0.7
Information theory and coding	0.6
Pattern recognition and learning systems	0.6

Figure 10: Illustration in Zadeh's article [30], 1968.

Zadeh explained: "Clearly, such numerical values of degrees of containment represent merely this writer's subjective assessment, expressed in quantitative terms, of the current consensus regarding the degrees of inclusion of various subjects in computer science." ([30], p. 913.) He also emphasized "that a high degree of containment of a particular subject in computer science does not imply that it cannot have a similar high or even higher grade of containment in some other field. For example "automata theory" has the degree of containment of 0.8 in computer science; it also has the same, or nearly the same, degree of containment in system theory. Also, the subjects listed in the table may have substantial overlaps with one another. This is true, for example, of "Automata theory" and "finite state systems"." ([30], p. 914.)

4 OUTLOOK

This work will be continued by generalization of Zadeh's "fuzzy conception" of Computer science to other scientific disciplines. Finally this should breed to a new concept in philosophy of science.

[2] [7]

[3] [1], [2].

Acknowledgements

I would like to thank Lotfi A. Zadeh for his generous support to my work on the history of fuzzy sets and systems and its applications for many years. Especially, I would like to thank him for his offer to let me browse in his private archives in his office in Soda Hall, UC Berkeley and at his home to search and find many important and interesting materials for my historical research. Work leading to this paper was partially supported by the Foundation for the Advancement of Soft Computing Mieres, Asturias (Spain).

References

[1] W.F. Atchison et al.: An Undergraduate Program in Computer Science—Preliminary Recommendations, *Communications of the ACM*, vol. 8, pp. 543-552, September 1965.

[2] W.F. Atchison et al.: Recommendation for Academic Programs in Computer Science, *Communications of the ACM*, vol. 11, March 1968.

[3] L. von Bertalanffy,: General System Theory. In: L. von Bertalanffy, A. Rapoport (Eds.): *General Systems* (= Yearbook of the Society for the Advancement of General Systems Theory. Vol. I.). Tunbridge Wells: Abacus Press, 1965, pp. 1-10, 1965.

[4] O. Lowenschuss: *Multi-Valued Logic in Sequential Machines*. Ph.D. Thesis. School of Engineering, Columbia University, New York, 1958.

[5] O. Lowenschuss: A Comment on Pattern Redundancy, *IRE Transactions on Information Theory*, p. 127, December 1958.

[6] O. Lowenschuss: Restoring Organs in Redundant Automata, *Information and Control*, vol. 2, pp. 113-136, 1959.

[7] A.G. Oettinger: The Harvard-Software-Complementarity, *Communications of the ACM*, vol. 10, October 1967, pp. 604-606.

[8] R. Seising: Interview with L.A. Zadeh on September 8, 1999 in Zittau, at the margin of the 7th Zittau Fuzzy Colloquium at the University Zittau/Görlitz.

[9] R. Seising: Interview with L.A. Zadeh on July, 26, 2000, University of California, Berkeley, Soda Hall.

[10] R. Seising: Interview with L.A. Zadeh on June 15, 2001, University of California, Berkeley, Soda Hall.

[11] R. Seising: Interview with L.A. Zadeh on June 16, 2001, University of California, Berkeley, Soda Hall.

[12] R. Seising: *The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and Its Initial Applications – Developments up to the 1970s*, Studies in Fuzziness and Soft Computing Vol. 216, Berlin, New York, [et al.]: Springer 2007.

[13] C.E. Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal*, vol. 27, pp. 379-423 & 623-656, 1948.

[14] C.E. Shannon, J. McCarthy (Eds.): Automata Studies. Princeton, New Jersey: Princeton University Press, 1956.

[15] A.M. Turing, Computing Machinery and Intelligence. In: *Mind*, vol. LIX, 1950, pp. 433-460.

[16] W. Ulrich: Nonbinary Error Correction Codes, *Bell Systems Technical Journal*, pp. 1341-1142, November 1957.

[17] N. Wiener: *Cybernetics or Control and Communications in the Animal and the Machine*, Cambridge, Massachusetts: MIT Press, 1948.

[18] L.A. Zadeh, J.R. Ragazzini: An Extension of Wiener's Theory of Prediction, *Journal of Applied Physics*, vol 21, 1950, pp. 645-655.

[19] L.A. Zadeh, Thinking machines - a new field in electrical engineering, *Columbia Engineering Quarterly*, January 1950, pp. 12-13, 30-31.

[20] L.A. Zadeh: Theory of Filtering, *Journal of the Society for Industrial and Applied Mathematics*, vol. 1, 1953, pp. 35-51.

[21] L.A. Zadeh, System Theory, *Columbia Engineering Quarterly*, Nov. 1954, pp. 16-19, and 34.

[22] L.A. Zadeh: On the Identification Problem, *IRE Transactions on Circuit Theory*, pp. 277-281, December 1956.

[23] L.A Zadeh: Toward an Institute for Research in Communication Sciences. *IRE Transactions on Information Theory*, Editorial, p. 3, March 1960.

[24] L.A. Zadeh, From Circuit Theory to System Theory, *Proceedings of the IRE*, vol. 50, 1962, pp. 856-865.

[25] L.A. Zadeh: A Critical View of Our Research in Automatic Control, *IRE Transactions on Automatic Control*, AC-7, No. 3, pp. 74-75, April 1962.

[26] L.A. Zadeh, Ch. A. Desoer: *Linear System Theory: The State Space Approach*. New York, San Francisco, Toronto, London: McGraw-Hill Book Comp., 1963.

[27] L.A. Zadeh: Fuzzy Sets, *Information and Control*, vol. 8, 1965, pp 338-353.

[28] L.A. Zadeh: Electrical Engineering at the Crossroads, *IEEE Transactions on Education*, Vol. E-8, Nos 2-3, June-September 1965, pp. 30-33.

[29] L.A. Zadeh: Education in Computer Science, Proceedings of the National Conference on Data Processing, Jerusalem, Information Processing Association of Israel, 1968, pp. E157-E167.

[30] L.A. Zadeh: Computer Science as a Discipline, *Journal of Engineering Education*, Vol. 58, No. 8, April 1968, pp. 913-916.

[31] L. A. Zadeh, E. Polak, *System Theory*, Bombay, New Delhi: McGraw-Hill 1969.

[32] L.A. Zadeh: *History of Computer Science at Berkeley*, Manuscript from November 1, 1992, Personal Archive of L.A. Zadeh, page 2.

[33] L.A. Zadeh: My Life and Work − a Retrospective View. *Applied and Computational Mathematics*, vol. 10, Nr. 1, 2011, pp. 4-9.

¿POR QUÉ USAMOS TANTO LENGUAJE IMPRECISO? A FAVOR DE UNA SEMÁNTICA COMPARATIVA PARA LOS PREDICADOS VAGOS

Alejandro Sobrino

Facultad de Filosofía, Plaza de Mazarelos, s. n. Santiago de Compostela, alejandro.sobrino@usc.es

Resumen

En este trabajo se comentan algunas de las razones más comúnmente aducidas para justificar el uso abundante de lenguaje vago en nuestras conversaciones ordinarias. La utilidad es una de las razones más invocadas y la teoría de juegos ofrece un marco para estudiar y cuantificar esa noción. Ese enfoque favorece, a mi juicio, una semántica comparativa como la forma más plausible de evitar algunos problemas que ocasionan el uso de los términos vagos.

Palabras clave: vaguedad, imprecisión, evolución, utilidad, comparativos.

1 INTRODUCCIÓN

Tradicionalmente se han asociado a la vaguedad lingüística connotaciones negativas y a la precisión notas positivas. Así, es frecuente acompañar un gesto de desaprobación con la expresión 'terriblemente vago' y un rictus de satisfacción al decir 'absolutamente preciso'. La percepción negativa de la vaguedad proviene de una visión esencialista del significado, de la que fueron primeros adalides Platón o Aristóteles. Este último, decía que las cosas tienen que ser llamadas por sus nombres precisos y no usando términos genéricos o vagos. Concepciones del lenguaje más flexibles, como la inaugurada por Wittgenstein en las *Investigaciones Filosóficas*, rehabilitaron el papel de la vaguedad en la comunicación ordinaria, al subrayar que el significado de la mayoría de las palabras no queda tanto definido por características precisas e inmutables, como por el parecido de familia con palabras de significado similar. El lenguaje preciso pertenece fundamentalmente al dominio de la matemática, y el lenguaje vago, al de la comunicación ordinaria, donde, lejos de bloquear o impedir el diálogo, se convierte a menudo en una estrategia dúctil y útil en el intercambio de mensajes.

Es un truismo decir que usamos mucho más lenguaje impreciso que preciso en la conversación ordinaria. Siendo esto algo obvio, no lo es preguntarse por qué. La respuesta a esta cuestión puede situarse, entre otros, en un marco biológico o cultural. Quizás usamos más lenguaje impreciso que preciso porque nuestra configuración neuronal o cognitiva se adapta mejor a este tipo de palabras. O quizás usamos en mayor abundancia lenguaje impreciso porque con él obtenemos beneficios en el proceso comunicativo que de otra manera no podríamos lograr. Quizás ambas tesis sean el anverso y el reverso de una misma moneda: puede que un cerebro impreciso beneficie la comunicación y el éxito de las palabras imprecisas en la comunicación favorezca la estructura neuronal que las genera, en un flujo de retroalimentación que contribuye, como finalidad última, a mejorar las condiciones de pervivencia de la especie. En el rastreo de este tema se han dado argumentos a favor de ambas hipótesis. A continuación desgranaremos algunos de ellos.

En [2], el neurólogo W. Calvin apunta a las posibles bases neurológicas de la imprecisión: *"la naturaleza parece tener cierta predilección por los límites difusos, al menos en lo que al nivel de organización se refiere….La precisión no es sino una propiedad emergente de un gran número de neuronas imprecisas que trabajan de forma redundante en una misma tarea.* (p. 34). Lamentablemente, Calvin no es mucho más explícito al respecto (al menos no en ese texto), si bien trata de referir y ejemplificar lo que dice aludiendo a la memoria: es verdad que la memoria humana posee equivalentes funcionales a la memoria intermedia del teclado, a la memoria inmediata RAM o a la memoria a largo plazo de los discos duros, pero no se puede comparar con la memoria de un ordenador. En el ordenador los datos se almacenan de forma precisa, ocupando ranuras vacías que no se solapan. En cambio, la memoria humana se caracteriza por el almacenamiento distribuido o superpuesto en el que los nuevos datos tienen que ser compatibles con las resonancias redundantes que generan los datos antiguos; ruido que, antes que ser inútil, le proporciona el contexto necesario para que sean comprendidos adecuadamente. Para Bickerton, [2], esta reflexión se traduce al ámbito lingüístico en términos de transmodalidad. Las palabras son transmoda-

les porque no significan siempre de la misma manera. Cuando oímos '¡fuego!' podemos reírnos o escapar, dependiendo de si escuchamos esa palabra cuando se está representando una obra de contenido erótico o cuando ya ha acabado. Entender su significado de modo apropiado depende del contexto, de un conjunto no explícito y a menudo poco definido de notas que, no obstante, nos inducen normalmente de modo bastante preciso a un comportamiento acertado. Si bien la transmodalidad distancia al lenguaje humano de la llamada de los animales, en el siguiente párrafo veremos como las llamadas tampoco son ajenas a la imprecisión.

En el ámbito de la etología "*hay innumerables historias entresacadas de la conducta animal que ilustran el hecho de que puede que los conceptos no consigan ser más precisos de lo estrictamente necesario (de hecho, son en ocasiones tan rudimentarios que es posible cometer errores de bulto, como ocurre, por ejemplo, cuando algunos pájaros atacan a los pollos de su propia progenie una vez que éstos, habiéndose extraviado más allá del anillo de guano que delimita su margen de movilidad, intentan regresar al nido*" [2, p. 41]-. Así pues, en el mundo animal, donde no hay cultura y sí dispositivos automáticos de comportamiento, parece que la imprecisión tiene un papel en el ámbito de la conducta.

En *Conjeturas y refutaciones*, Popper apuntó algo similar, pero referido al comportamiento lingüístico humano: *"Aunque la claridad es apreciable en si misma, la exactitud o la precisión no lo es: no hay razón para ser más preciso que lo que demanda nuestro problema." (p. 28)*. Esta tesis se verá apoyada por algunos argumentos sobre el papel de la vaguedad en el lenguaje, como veremos más adelante. Para el sesgo que tomará este trabajo, nosotros la parafraseamos de la siguiente manera: tratar de obtener la precisión siempre y en toda circunstancia tiene un coste que no es racional asumir.

En el lenguaje humano las palabras se usan frecuentemente como monedas de cambio. Aprender una palabra no es aprenderla para uno, sino para que otros la entiendan, en un proceso dialógico en el que se transita de significados a sonidos (hablante) y de sonidos a significados (oyente), posibilidad que surge de compartir, aunque de manera aproximada e incompleta, un mismo código lingüístico. En [8], Tomasello ha enfatizado la bi-direccionalidad en el aprendizaje lingüístico. Para aprender una palabra, el niño ha de entenderla como algo socialmente compartido: ha de suponer que el oyente la comprende y puede producirla y también que el oyente sabe que él la puede comprender y producir. Este proceso de 'imitación con inversión de roles' constituye la base de la adquisición lingüística. Si es importante emitir un mensaje, es crucial que su contenido sea compartido y tenga la intención de provocar algún tipo de reacción en quién lo escucha.

Con lo dicho anteriormente queremos enfatizar que la comunicación es intencional. A menudo con el lenguaje pretendemos provocar algún tipo de reacción en el oyente usando oraciones informativas, desiderativas, imperativa,… Remedando a Austin, si con las palabras se hacen cosas, no es disparatado preguntarse si se puede evaluar el coste de esa manufactura en términos de las acciones que una preferencia favorece o inhibe en el oyente. Pero antes de abordar este tema, indaguemos acerca de la importancia que tiene el lenguaje vago en la constitución y desarrollo lingüístico del individuo, así como en qué características cognitivas pueden estar asociadas a un uso tan abundante del lenguaje vago.

2 LA IMPORTANCIA Y ABUNDANCIA DE LA VAGUEDAD EN LA COGNICIÓN Y EL LENGUAJE HUMANO

En la adquisición del léxico, la vaguedad tiene un papel prevalente. Las palabras vagas están entre las que primero aprenden los niños; con sólo 24 meses manejan un amplio repertorio de las mismas. En los primeros años del aprendizaje lingüístico, la generalidad tiene un papel primordial en la extensión de la referencia de las palabras: los niños usan 'sombrero' para denominar 'bufandas', 'diademas', 'fulares'…; esto es, para referirse de manera imprecisa a toda aquella prenda o adorno que se coloca por encima del cuello. Usando el parecido de forma, tamaño, textura, sonido o función, el niño amplia la referencia con un término genérico hasta que adquiere nuevo vocabulario más específico que le permite discriminar lo que ahora agrupa inespecíficamente de modo tan indeterminado. En [6], Rosch y Mervis señalaron que los niños almacenan muchas palabras (particularmente nombres y verbos), no como elementos semánticos individuales, sino como prototipos o ejemplares representativos de una clase con respecto a los cuales clasifican otros términos por similitud. De esta pequeña descripción se puede colegir que la vaguedad es un elemento central en la comunicación humana al formar parte, muy tempranamente, de los procesos básicos de adquisición del léxico.

En el ámbito cognitivo, se ha relacionado a la vaguedad lingüística con las limitaciones de nuestra memoria, en particular, y de nuestra capacidad perceptiva, en general. Salvo casos infrecuentes, los humanos tenemos una memoria frágil y limitada, y somos incapaces de recordar muchas cosas. Así, casi todos nos podríamos ver reflejados en el siguiente relato: escuchamos en un noticiario el número exacto de muertos que ha producido el reciente tsunami de Japón. Supongamos que hemos oído que 10.321 personas han perecido en ese desastre natural. Quizás por un momento podamos recordar la cifra exacta, pero seguramente al cabo de unos minutos, si no hay ninguna motivación añadida, ese número sería distorsionado o simplemente errado si alguien nos demandara, de nuevo, el dato exacto. Quizás recordaríamos una aproximación en términos de miles (10.000), o simplemente seríamos capaces de decir que fueron *muchos*. De modo similar, y en lo que respecta a la percepción, coincidiría-

mos en señalar que alguien aficionado a recuperar objetos perdidos en una playa podría marcar, recién descubierto, el lugar exacto del hallazgo, pero unos minutos más tarde seguramente sólo podría apuntar vagamente a una vasta área del arenal. La memoria y la percepción humana son bastante limitadas y parecen tener una tendencia irrefrenable a la economía, al resumen general y aproximado de lo que se dice o percibe, seguramente para dejar hueco al recuerdo de otras cosas y evitar el colapso. Se concluye, por tanto, que la vaguedad es un elemento frecuente e importante de nuestra cognición, de la memoria y de la percepción, así como del aprendizaje lingüístico.

Introduzcamos a continuación, algunos argumentos que explican, quizás, la abundancia de la vaguedad en el lenguaje ordinario. El primero es que la vaguedad permite evaluar situaciones en las que no disponemos de métricas bien definidas o, si las hay, no son participadas de igual manera –con igual pericia- por los agentes de la comunicación. Expresiones como 'fiebre alta' hacen algo más que generalizar la información de una medida precisa: añaden una valoración a lo que se dice en forma de llamada de atención o precaución, en el sentido de que si uno escucha que un parámetro biológico está 'alto', sabe que tiene que tomar precauciones. Normalmente, casi todo el mundo conoce que 39 ºC es una temperatura alta, pero otros datos médicos son mucho más desconocidos. En ese caso, no siempre la cifra que da el experto es útil para el lego; en numerosas ocasiones el uso de un término vago le puede aclarar mucho mejor al inexperto que uno preciso cuál es la situación y qué atención que debe poner a la misma.

Otro argumento es que la vaguedad, lejos de bloquear la comunicación, permite un ahorro al expresar con suficiencia aquello que hay que saber para tomar una decisión acertada. Evita la complejidad que supone el uso de una métrica precisa (disponer de aparatos, calibrarlos, preparar las condiciones de medida, hacer la medida, etc.) para tomar una acción que quizás no requiera esa finura. P. ej. en el contexto de decidir la cantidad de medicación que se le prescribe a un paciente, en muchas ocasiones es redundante conocer su peso exacto. Basta con discriminar si es 'adulto' o 'niño'; cualquier otra consideración es irrelevante. Este ejemplo ilustra bien el *dictum* de Popper de no ser más preciso de lo que la situación demanda. Usar un término genérico e inespecífico como 'adulto' permite una reducción en la prueba (la constatación del referente) y un ahorro cognitivo (una verbalización más fácil) sin que ello avoque a una decisión errada o menos acertada que si se hubiera procedido con una medición precisa (74,38 kg., p. ej.).

Estas razones nos acerca a la segunda parte de este trabajo, donde expondremos algunos escenarios dialógicos en los que dos o más personas se comunican para hacer una petición y atender a la misma. Coincidamos, de momento, en que puedan comunicarse usando lenguaje preciso o impreciso. Algunos ejemplos mostrarán los inconvenientes o beneficios de usar uno u otro. El análisis de las peticiones y de las acciones encaminadas a resolverlas se hará en el marco de la teoría de juegos. Se considerarán juegos de coordinación y juegos de confrontación. Los primeros parecen avalar el uso de lenguaje preciso; los segundos, de lenguaje vago. Alguna solución a estos juegos mostrará el papel benefactor que ejerce una semántica comparativa como salida razonable a los mismos.

3 LOS BENEFICIOS DE USAR LA VAGUEDAD EN LA COMUNICACIÓN

En [3] Lipman plantea el uso de palabras vagas en un juego colaborativo donde los intereses del hablante y del oyente confluyen. Lo ilustra en el escenario de un aeropuerto, dos jugadores (jugador_1 y jugador_2) y, al menos, tres actores, uno de los cuales es invitado del jugador_1. El jugador_2 tienen que ir al aeropuerto a recoger al invitado del jugador_1. El jugador_1 sabe su altura exacta y el jugador_2 se presta a ir al aeropuerto equipado con una cinta de medir. Recordemos que hay otras dos personas en el aeropuerto que, eventualmente, pueden ser confundidas con el invitado. Bajo estos supuestos, usar términos vagos no sería una buena estrategia. El jugador_2 no obtendría nunca mejor función de pago basando su decisión en la apreciación gruesa de que le parece que el individuo que tiene ante sí 'es alto' si puede sacar el metro y verificar, de forma fina, lo que mide. Saber la altura exacta le permite identificar al invitado con total certidumbre, dado que la probabilidad de que dos personas midan lo mismo es nula.

Podría utilizar, en vez de una medida precisa, un predicado comparativo, siempre que la comparación se hiciese respecto a un parámetro exacto, como p. ej. 'más alto que la media'. En el supuesto de que el jugador dispusiese de ese dato, la comunicación sería también óptima, ya que esa definición excluiría casos fronterizos o dudosos, aunque las dudas podrían resurgir si hubiese varios individuos en esa situación. Incluso una instrucción como 'el más alto' podría, caso de que hubiera una diferencia perceptiva significativa entre los tres, llevar al jugador_2 a seleccionar de modo adecuado a la persona en consideración.

En conclusión, Lipman no encuentra ningún motivo para usar un término vago en vez de uno preciso en un juego de coordinación. En ese caso, la mejor estrategia del hablante es siempre la precisión. La vaguedad induce a equívocos perniciosos por los casos fronterizos que presenta, aunque una semántica comparativa –el uso de un superlativo en este caso- puede ser igualmente eficiente que una medida precisa.

El uso de palabras vagas, es cierto, a veces no conviene. Si un piloto no recibe instrucciones precisas de la torre de control sobre la situación de la pista un día de niebla cerrada, quizás pueda tener serios problemas para aterrizar. Pero hay otras ocasiones en que la precisión, en vez de bloquear, favorece a un diálogo. En [5], R. Parikh mostró,

en un ejemplo ilustrativo, que la vaguedad puede suponer una ventaja para la comunicación cuando la palabra que usan dos personas para referirse a la misma entidad se utiliza de modo impreciso. Cuando eso ocurre, el solape es parcial y esa vaguedad puede ayudar a resolver el problema, a satisfacer la petición. El siguiente ejemplo así lo ilustra: Ada y Renato trabajan en un instituto donde Ada enseña Informática y Renato Filosofía. Un día Ada está en el despacho mientras Renato trabaja en casa. Renato tiene que explicar al día siguiente la filosofía de Wittgenstein y necesita el libro de ese autor titulado *Cuaderno Marrón*, que olvidó en su despacho. Supongamos que el *Cuaderno Marrón* tiene la portada de ese color tanto para Ada como para Renato y que Renato no tiene otro libro marrón en su estantería. La pide a Ada que se lo lleve y Ada, que no sabe nada de Filosofía, le pregunta qué aspecto tiene el libro que le pide. Él le dice simplemente: 'es de color marrón'. Parikh supone que tanto Ada como Renato tienen un concepto preciso de 'marrón'. Si ambos coinciden en lo que nombran con la palabra 'marrón', Ada seguramente encontrará el libro que le pide Renato y se lo llevará a su casa, pero si no coinciden, puede que el libro que Renato diría que es marrón, Ada no lo llamara así y que algunos libros que Renato descartaría como marrones sí lo serían en cambio para Ada. Si tal disparidad se produjese, y cada uno siguiese pensando que sólo hay un tipo de marrón, para ayudar a Renato, Ada se vería obligada a inspeccionar todos los libros de su estantería si no lo encuentra entre 'sus' marrones. Si, en cambio, Ada y Renato entendieran uno del otro que pueden haber usado 'marrón' como un término fuzzy, entonces Ada buscaría el libro entre aquellos que para ella son claramente marrones y, si no lo cree haber encontrado, entre aquellos que, aún no siendo absolutamente marrones, tienen un color similar, solapándose así con algunos de los que Renato diría también que son marrones y ampliando, en consecuencia, la posibilidad de satisfacer con éxito su petición. Parikh demostró, usando un ejemplo numérico concreto que, cuanto más amplio sea el solape entre las extensiones de $\text{marrón}_{\text{Ada}}$ y $\text{marrón}_{\text{Renato}}$ proporcionalmente a su diferencia simétrica, más se reduciría el tiempo de búsqueda de Ada para satisfacer a Renato.

Situar a la vaguedad en términos de los beneficios que produce en la comunicación permite evaluar en su justo término a un tipo de vaguedad problemática: la vaguedad de orden superior, un virus que suele atacar de forma importante a la semántica de los grados de verdad. Según ella, el significado de una sentencia declarativa se identifica con las condiciones o circunstancia que la hace, no absolutamente verdadera, como ocurriría con la lógica clásica, sino verdadera en un grado. El problema es que, una vez fijados, los grados de verdad son tan precisos como la propia verdad. Estoy de acuerdo con Sainsbury [7] en que lo que caracteriza a la vaguedad no es introducir más de una marca entre la verdad y la falsedad, sino la ausencia de marcas: las palabras vagas son *boundaryless*. La introducción de divisiones aboca a preguntar por la vaguedad de orden superior.

La vaguedad de orden superior se muestra de modo paradigmático en la paradoja del Sorites. Una descripción habitual de esta paradoja es la que sigue: Un acúmulo de 10^{10} granos de café es un montón. Todo acúmulo de granos no deja de serlo por sustraerle un grano. Iterando esa premisa inductiva tantas veces como granos haya de partida llegamos a la conclusión falsa y absurda de que un acúmulo de 0 granos de café es un montón. La paradoja del Sorites ilustra un problema racional, ya que de premisas sensatas, incluso verdaderas si se toman aisladamente, se concluye algo evidentemente falso. Ello es debido, al menos, a dos factores: 1) se intenta razonar de forma precisa con predicados que no lo son, 2) los predicados vagos suelen tener varias dimensiones y el sorites escoge, para su representación, sólo una. Así, el predicado 'montón' no sólo se define por el número de granos, sino por la forma que estos adoptan. No es lo mismo 10^{10} granos de café alineados uno tras otro que apiñados unos encima de otros. Si el montón depende del número de granos y de su forma no hay inconveniente en suponer que una aplicación singular del paso inductivo produzca que este se derrumbe y no sea más un montón. Aún así, se puede argumentar que hay desmoronamientos que dejan a la configuración final más próxima a montón que a no montón; el acúmulo resultante es más un montón que un no montón; más un montón que otra cosa. Pero esto es lo esperable y quizás todo lo que podemos decir. Puede que poseamos una imagen prototípica de 'montón' con la cual comparemos otros acúmulos y eso nos permita decir si uno es más 'montón' que otro. De nuevo, los comparativos parecen lo más sensato para decidir este tipo de situaciones.

La vaguedad de orden superior y su inevitabilidad la ilustró Parikh en el teorema homónimo que dice así [4]: en un espacio métrico α-conectado, siempre es posible encontrar un punto indistinguible a otro dado. Como hemos visto en su ejemplo anterior, si 'marrón' es un término vago, debemos estar dispuestos a conceder que hay objetos marrones, claramente no marrones y aproximadamente marrones. Haciéndolo así, interponemos entre un concepto y su negación un espacio intermedio reservado a lo dudoso. ¿Pero por qué no diferenciar los más dudosos de aquellos que lo son menos, por estar más próximos a una de las dos categorías claras? Tendríamos así una vaguedad dentro de la vaguedad; esto es, una vaguedad de orden superior. Y este argumento se puede reproducir todas las veces que queramos. La pregunta clave es: ¿dónde detener el proceso? Esta es una cuestión a la que aluden los clásicos para devolver el veneno que le hemos inoculado en su lógica en términos de arrojar dudas sobre la validez del Principio de Bivalencia: ahora en vez de una marca entre la verdad y la falsedad, hay dos. Pero esas dos marcas delimitan de forma precisa tres áreas. Si los conceptos son borrosos, esas marcas admitirán espacios intermedios que, en una ascensión semántica, conducen a infinitas marcas y, paradójicamente, a discriminar con una precisión extraordinaria algo que es impreciso. Tal acontece cuando en una semántica de gra-

dos se dice que algo verifica tal propiedad en grado 0,54373. Atendiendo a los criterios de economía antes comentados, parece difícil apreciar cómo la vaguedad de orden superior puede causar una reducción lingüística o cognitiva mayor que la que logra la vaguedad de primer orden. Esto debe ser un argumento a favor de una semántica menos fina, que opte por sustituir conceptos ordinales por comparativos o simplemente cuantitativos, reflejando mejor la naturaleza de la vaguedad y permitiendo su aprovechamiento óptimo en la comunicación. Retomando a Popper, si es razonable no ser más preciso de lo que la situación pide, también parece igualmente sensato no ser más vago que lo que el escenario demanda. Y a menudo demanda que nos aproximemos todo lo que podamos a la precisión, no a la vaguedad.

Van Deemter, refiere en [9] un ejemplo de cómo una semántica comparativa ayuda a resolver este problema y frena cualquier debilidad de ahondar en una discriminación cada vez más fina de la vaguedad. Retoma el ejemplo de Lipman antes visto y pregunta qué ocurriría si en vez de un predicado vago, que pretendidamente malograría la comunicación, se usa el sintagma nominal vago 'la persona alta' para que el jugador_2 localice al invitado. La respuesta es que se consigue una descripción suficientemente precisa usando una expresión vaga. Siempre que se pueda discriminar la altura de las tres personas para dirimir perceptualmente y sin equívoco quién es la más alta, el uso de ese sintagma nominal sería suficiente. La comparación se hace al instante y, si la diferencia es apreciable, con más confianza que una medida, que siempre admite márgenes de precisión. La clave está en que en un sintagma nominal como 'el hombre alto' no hay vaguedad global, sino sólo vaguedad local. La expresión no permite casos fronterizos: está claro quién es el hombre alto, aunque pueda haber dudas sobre cómo de alto es o cuanto más alto es ese individuo que otros. Pero si esta información no es relevante para resolver el problema, no tenemos por qué tenerla en cuenta.

Otro ejemplo suyo, similar al de Parikh antes comentado, muestra la pertinencia de la semántica comparativa para frenar el ascenso que supone la vaguedad de orden superior. El ejemplo se sitúa en la ciudad prohibida de Beijing. Hace años, en el palacio del Emperador desapareció un diamante. La seguridad de palacio le contó al Emperador que uno de sus miembros creyó haber visto algo sospechoso en un individuo, pero que cuando trató de apresarlo, el sospechoso hirió mortalmente al guardia y lo único que éste fue capaz de decir es que se trataba de un eunuco y que 'era alto'. El dato 'eunuco' es relevante, porque delimita el espacio de búsqueda (había 1000 eunucos), pero ¿cómo de informativo fue para el Emperador el dato 'alto'? Supongamos que el concepto que tiene el emperador de 'alto' es preciso y coincide con el del malogrado guardia, pues ambos entienden por 'alto', 'más alto que el promedio'. En ese caso, el Emperador daría órdenes a su guardia para que inspeccionase a todos los eunucos que son más altos que la media (500). Pero supongamos ahora que el Emperador duda que él y el informador tengan el mismo concepto de 'alto' y cree que este último usa, en realidad, una noción más laxa que la suya. Ya que la lógica clásica no distingue tipos de 'no alto', la guardia buscaría al ladrón primero, de forma arbitraria, inspeccionando a 500 eunucos, pero como en Emperador no tendría la certeza absoluta de que lo que buscaba estuviera en esa población, daría orden de ampliarla en 0.5*500=250, inspeccionando, en total, 750 eunucos. Si las dudas permaneciesen, podría evitar una salida en falso considerando más y más eunucos como altos hasta inspeccionarlos a todos, con lo cual el enunciado del testigo no le habría sido de utilidad alguna. Es más, en esa situación, la información habría resultado perjudicial para el Emperador, ya que, recordemos, sin ningún dato hubiera buscado en 0.5*1000=500 eunucos. ¿Cómo podría haber conciliado el Emperador la información del malogrado guardia y sus expectativas de encontrar al ladrón? Para van Deemter, una semántica de la comparación podría acudir al rescate de la siguiente forma: supongamos que el Emperador ordena a los eunucos según su altura y asume que cuanto más alta sea una persona, más probabilidad tiene de coincidir con la que el testigo ha denominado así. Si el ladrón no está entre los más altos que la media, esta estrategia puede ayudarle a resolver el puzle sin un coste tan importante como acudir a una vaguedad de orden superior.

La teoría de juegos permite calcular la utilidad de una acción. Si como decía Searle, proferir una palabra es un acto de habla, quizás se pueda calcular la utilidad de los actos lingüísticos vagos. E. Aragonés y Z. Neeman hicieron, en el marco de la teoría de juegos, una propuesta al respecto [1]. Se trata del siguiente juego de confrontación: imaginemos a dos políticos sin escrúpulos que se postulan para su elección ante los votantes, asumiendo que lo que es bueno para el político suele ser malo para el votante y viceversa. Sin grandes convicciones, pueden elegir representar a tres ideologías diferentes (izquierda, derecha y centro) dependiendo de lo que le sea de más utilidad (en este contexto, lo que más votos le dé). Adicionalmente, los políticos pueden mostrar dos niveles de compromiso, alto y bajo, $c_{alto}>c_{bajo}$. ¿Qué combinación de ideología y compromiso debería escoger cada político para obtener el mayor número de votos? La respuesta a esta cuestión dependerá en buena medida de los votantes. Supongamos que hay tres bloques de votantes: $votante_{izquierda}$, $votante_{derecha}$ y $votante_{centro}$. Un votante de izquierda prefiere a un político de izquierda y, si es posible, con un nivel de compromiso alto. Pero si tuviera que elegir entre dos políticos de derechas, preferiría que triunfase uno con un nivel de compromiso bajo. Y lo mismo pensaría un votante de derechas si tuviera que elegir a un político de izquierdas. En ese escenario, los políticos escogerán una ideología y un compromiso basados en las estimaciones del número de votos que puedan conseguir en cada bloque de votantes, tratando de maximizar su función de pago, formulada en términos de la probabilidad de ganar las elecciones. La utilidad para el político i, $U_i(I_1, c_1; I_2, c_2)= P_i(I_1, c_1; I_2, c_2)*k- c_i)$, I_i=ideología y c_i=compromiso, será

el producto de la probabilidad de que gane las elecciones por una constante que penaliza el compromiso. Usando este presupuesto, Aragonés y Neeman demuestran que un compromiso bajo puede ofrecer una utilidad global mayor que un compromiso alto. En efecto, supongamos que el político_1 adopta c_{bajo}. En función de cuántos votantes haya de izquierda, P_1(izqda, c_{bajo}; I_2, c_2) será menor que P_1(izqda, c_{alto}; I_2, c_2), aunque, a cambio, puede que un compromiso bajo le de votos de centro y de derecha. Dependiendo de la diferencia entre P_1(izqda, c_{bajo}; I_2, c_2) y P_1(izqda, c_{alto}; I_2, c_2), el político puede lograr una utilidad global mayor escogiendo un nivel de compromiso bajo y no alto. Puede entenderse 'compromiso bajo' como algo vago, p ej., detraer una parte significativa de la renta de unos cuantos ricos para dársela a la mayoría de los más pobres.

4 CONCLUSIÓN. A FAVOR DE UNA SEMÁNTICA COMPARATIVA DE LOS PREDICADOS VAGOS.

En el punto anterior hemos visto diferentes ejemplos en los que se resalta el papel de los comparativos para una utilización óptima de la vaguedad en la comunicación ordinaria. También hemos visto que usando predicados comparativos se mitigan los problemas derivados de una vaguedad de orden superior.

En el ejemplo de Lipman se muestra como una semántica comparativa –en concreto, el uso del superlativo 'el más alto'- puede ser menos costoso, pero tan eficiente, como la utilización de una medida precisa (1,775 cms.).

El Sorites muestra que, en el caso de los montones, cuando no tenemos un criterio claro para decidir si algo sigue siendo o no un montón, quizás sí lo tengamos para decir que nos parece más un montón que un no montón; más un montón que otra cosa.

También se ha indicado que el uso de un sintagma nominal vago como 'la persona alta', siempre que se pueda discriminar perceptualmente a la de mayor altura entre aquellas involucradas en el juego, puede ser suficiente para no errar en la identificación del individuo en consideración.

Finalmente, el ejemplo de los eunucos muestra como se puede usar el predicado vago 'alto' de manera fuzzy sin vernos condenados a una vaguedad de orden superior. Se asume que cuanto más alta sea una persona, más probabilidad tiene de coincidir con la que el testigo ha denominado así. Si el ladrón no está entre los más altos que la media, esa estrategia puede ayudar a encontrarle sin caer en la parálisis a la que conduce una vaguedad de orden superior.

Una de las propiedades de la vaguedad es que es contaminante. Así, decir que tenemos a tres personas invitadas es una aserción precisa, pero si le añadimos el calificativo 'altas', entonces pasa a ser vaga, debido al contagio del predicado vago. En el Sorites, aunque la aplicación de la regla inductiva es precisa, la premisa de partida no lo es, por contener a un predicado vago. Esta vaguedad se propaga en el razonamiento de una manera inevitable. Las construcciones comparativas o superlativas pueden detener esa infección: si hay discriminación suficiente, quizás no se pueda determinar cuando un hombre deja de ser calvo, pero sí si es más calvo que otro o si es el más calvo de todos.

La vaguedad se muestra útil en el discurso ordinario y de los beneficios que supone su uso se puede hacer una justificación racional. Esa justificación muestra que no debemos ser más precisos ni más vagos de lo que la situación demanda. Los predicados comparativos, a medio camino entre los cuantitativos y los ordinales, parecen recoger bien esta filosofía, mostrando cómo se pueden resolver situaciones que involucran el uso de palabras vagas con el mínimo coste lingüístico y cognitivo sin perder eficiencia o utilidad.

Agradecimientos

Proyecto TIN2008-06890-C02-01, MICIN.

Referencias

[1] Aragonés, E. & Neeman, Z., 'Strategic ambiguity in electoral competition', *Journal of Theoretical Politics*, 200, 12, 183-204.

[2] Calvin, W. Bickerton, D., Lingua ex Machina. La conciliación de las teorías de Darwin y Chomsky sobre el cerebro humano. Barcelona, Gedisa, 2001.

[3] Lipman, B. L., Why is language vague? Working paper, dec., 2006, Dpt. Of Economics, Boston University.

[4] Parikh, R., The problema of vague predicates, in *Logic, Language and Method*, R. S. Cohen & M. W. Wartofsky (eds.) Reidel, 183, 241-261

[5] Parikh, R., Vague Predicates and Language Games. *Theoria* 1996, 11 (3), 97-107.

[6] Rosch, E. H. & Mervis, C. "Categorization of Natural Objects," *Annual Review of Psychology*, 1981, 32, 89-113.

[7] Sainsbury, M. Is there higher-order vagueness? *The Philosophical Quarterly*, 1991, 41 (163), 167-182.

[8] Tomasello, M. Los orígenes culturales de la cognición humana. Amorrortu editores, 2007.

[9] van Deemter, K., 'Utility and language generation: the case of vagueness'. *Journal of Philosophical Logic*, 2009, 38, 6, 607-632.

MODELO HEURÍSTICO PARA LA GESTIÓN DE NECESIDADES ESPECIALES DE INFORMACIÓN

José M. Morales-del-Castillo [1], Eduardo Peis [1], José Andrés Delgado López [1]

[1] *Universidad de Granada. Dpto. Biblioteconomía y Documentación. Colegio Máximo de Cartuja (18071) Granada*, {josemdc, epeis, jadelgadol}@ugr.es

Resumen

Presentamos el estado de la investigación en curso para la creación de un sistema de recomendaciones que pretende servir como herramienta para satisfacer necesidades especiales de información de la comunidad científica. El sistema proporciona recomendaciones sobre recursos de temas de investigación tangencialmente relacionados con los de interés del usuario, para permitirle descubrir áreas afines y poder, de esta forma, crear grupos de investigación multidisciplinares. Para ello, se aplican conjuntamente tecnologías de Web Semántica, técnicas de modelado lingüístico difuso y análisis de coocurrencias.

Palabras Clave: Sistemas de recomendación, Tecnologías Web Semántica, Redes de coocurrencia, Modelado Lingüístico Difuso.

1 INTRODUCCIÓN

La actividad científica es un objetivo primordial para la mayoría de los países avanzados, debido a las implicaciones socio-económicas e incluso políticas de su desarrollo. Por poner un ejemplo, sólo la Unión Europea en el séptimo programa marco 2007-2013 en investigación e innovación (FP7) dispone de un presupuesto de alrededor de cincuenta mil millones de euros [1].

Tanto la empresa privada como (especialmente) las universidades son el motor de la innovación, y a día de hoy existen numerosos ejemplos de grupos de investigación universitarios que crean compañías *spin-off* y empresas que aprovechan la innovación generada en las universidades para desarrollar nuevos productos y servicios.

Por lo tanto, para los gobiernos (que están invirtiendo capital público en las universidades y que aspiran a liderar los rankings de investigación) debe ser una prioridad la creación de una infraestructura sólida y propicia para que los grupos de investigación y universidades puedan generar innovación.

Dentro del marco de la teoría de la gestión del conocimiento, la innovación está íntimamente relacionada con el concepto de capacidad de absorción (*absorptive capacity*)[2], que se define como la habilidad de los individuos u organizaciones para reconocer el valor de la información que proviene de fuentes externas, con objeto de procesarla y asimilarla para resolver problemas concretos.

El valor de la capacidad de absorción no sólo reside en la adquisición de habilidades para resolver problemas, sino que es un elemento crucial para la creación de innovación. Para que esta capacidad de absorción se haga evidente, es necesario que el agente (entendido como la entidad que genera innovación) disponga de un conocimiento previo sobre la información que pretende asimilar. Además, cuanto más variado sea éste, mayor será la capacidad de aprendizaje del agente, ya que esta diversidad facilita una fijación más rápida y sólida del nuevo conocimiento. En otras palabras, un grupo de investigación que acoja miembros provenientes de diversas disciplinas, tiene más probabilidades de generar innovación desde el momento en que cada uno de estos individuos puede contribuir desde diferentes puntos de vista para encontrar soluciones a un mismo problema. Por lo tanto, la multidisciplinariedad puede ser considerada como una ventaja competitiva en términos de creación de innovación.

En el mundo académico, el trabajo en equipo representa, de hecho, uno de los principales modos de llevar a cabo tareas de investigación [3][4]. No obstante, encontrar colaboradores potenciales puede ser una tarea ardua, especialmente cuando los individuos o grupos objetivo, trabajan en dominios tangencialmente relacionados con el nuestro.

Tradicionalmente, los sistemas de recomendación (SR) han demostrado ser herramientas eficaces cuando se intenta descubrir personas con intereses similares a los nuestros, productos que se ajustan a nuestros gustos, o artículos científicos que profundizan en nuestro dominio de especialización. Pero cuando se trata de necesidades especiales de información, como es el caso concreto de descubrir áreas afines donde encontrar nuevas oportuni-

dades de colaboración, esta aproximación no resulta adecuada.

En este trabajo, presentamos los elementos más relevantes de una investigación en fase de desarrollo cuyo objetivo es crear un sistema de recomendaciones que permite sugerir, a una comunidad concreta de investigadores, no sólo recursos de su área sino también de áreas afines, de acuerdo al análisis de los intereses que han explicitado en sus perfiles de usuario. Para ello proponemos la aplicación de manera conjunta de tecnologías Web Semántica y redes de coocurrencia.

El resto de la comunicación se estructura de la siguiente manera: en el apartado 2 hacemos un breve repaso de trabajos relacionados con nuestra propuesta. En el apartado 3, presentamos las principales tecnologías implicadas en el desarrollo del modelo. En el apartado 4, hacemos una breve descripción de la arquitectura del modelo, para seguidamente explicar, en el apartado 5, el algoritmo que utilizamos para generar las recomendaciones y la estrategia a seguir para la implementación del modelo. Por último, en el apartado 6 presentamos algunas conclusiones y trabajos futuros.

2 TRABAJOS RELACIONADOS

Los SR se pueden definir como aplicaciones que exploran fuentes de información para sugerir a un individuo determinado distintos tipos de productos o recursos, de manera que éste pueda disponer de elementos de valoración que le ayuden en la toma de decisiones [5]. Se clasifican principalmente en dos categorías según la técnica usada para generar la recomendación [6]: por un lado encontramos los SR colaborativos, los cuáles utilizan información relativa a los intereses o usos de individuos con características parecidas, para generar recomendaciones para el usuario activo; y, por el otro, estarían los SR basados en contenido, en los que la recomendación se genera contrastando las preferencias del usuario (expresadas por éste de forma explícita o implícita, mediante la monitorización de su comportamiento en el sistema) con las descripciones de los recursos o productos que se van a recomendar. La mayoría de estos sistemas presentan pobres resultados cuando se dispone de poca información sobre los usuarios o cuando éstos tienen gustos muy heterogéneos. La tendencia actual, consiste en desarrollar SR híbridos, que aglutinan características de ambos SR para mejorar el resultado de las recomendaciones.

Las áreas de aplicación de los SR son muy diversas y los podemos encontrar, por ejemplo, en servicios de comercio electrónico [7], o plataformas de *e-learning* [8]. El SR que presentamos en este trabajo se centra en la generación de recomendaciones en el ámbito académico y en la literatura podemos encontrar varios ejemplos de modelos que favorecen la cooperación entre investigadores [9], o que presentan a empresas e investigadores oportunidades de colaboración al recomendar proyectos de investigación que pueden ajustarse a los intereses de ambos [10]. Para su desarrollo se basan principalmente en la aplicación de técnicas de modelado lingüístico difuso multigranular [11] para representar y gestionar la información de una manera flexible por medio de etiquetas lingüísticas.

Nuestra propuesta, por su parte, apuesta por un enfoque basado en el uso de tecnologías semánticas para desarrollar esquemas de conceptos u ontologías, que sirvan como base de conocimiento para generar las recomendaciones, tal y como se hace en otros trabajos de índole similar [12] [13].

3 TECNOLOGÍAS IMPLICADAS EN EL DESARROLLO DEL MODELO

Para la construcción del SR nos basamos principalmente en tres pilares: por un lado en las tecnologías de Web Semántica, por otro en el análisis de coocurrencias y, por último, en las técnicas de modelado lingüístico difuso. Vamos a comentar brevemente cada una de estas tecnologías.

La Web Semántica [14] es una extensión del modelo de Web actual que utiliza una serie de vocabularios estándar que permiten la descripción enriquecida de recursos para hacerlos semánticamente accesibles. Para ello, se apoya en dos pilares básicos: i) el etiquetado semántico de los recursos, y ii) el desarrollo de agentes software capaces de interpretar su contenido [15]. De esta manera, es posible pasar de la búsqueda con palabras clave a la recuperación mediante la definición de conceptos. La piedra angular del modelo es el vocabulario RDF (*Resource Description Framework*) [16], ya que proporciona el modelo de datos necesario para representar, intercambiar y enlazar metadatos estructurados en forma de tripletas (*sujeto-predicado-objeto*) provenientes de diferentes fuentes distribuidas. Para poder procesar la semántica de los recursos es necesario recurrir al uso de ontologías, unas herramientas que definen relaciones complejas entre los términos relevantes de un dominio determinado [17] [18].

Dado que la información en la Web está expresada principalmente en lenguaje natural, a veces es complejo gestionar problemas donde se trabaja con cierto grado de incertidumbre. En ese caso, el modelado lingüístico difuso [19] proporciona un conjunto de técnicas que son adecuadas para tratar aspectos cualitativos de determinados problemas. En concreto, el enfoque lingüístico ordinal se define de acuerdo con un conjunto finito S de etiquetas lingüísticas ordenadas en una escala de orden total y con cardinalidad impar (7 o 9 etiquetas):

$$\{s_i, i \in H = \{0,...T\}\} \quad (1)$$

El término central tiene un valor de aproximadamente 0.5 y el resto de términos se distribuyen simétricamente alrededor de él. La semántica de cada término lingüístico viene dada por la estructura ordenada del conjunto de

términos, considerando que cada término del par (s_i, s_{T-i}) ofrece la misma información. A cada etiqueta s_i se le asigna un valor difuso definido en el intervalo [0,1], que viene definido por una función lineal trapezoidal representada por la 4-tupla $(a_i, b_i, \alpha_i, \beta_i)$ (donde los dos primeros parámetros representan el intervalo donde el valor de la propiedad es 1.0 y los valores restantes definen lo límites izquierdo y derecho de la distribución). De manera adicional, es necesario definir las siguientes propiedades:

1. *El conjunto está ordenado:* $s_i \geq s_j \; if \; i \geq j$
2. *Negación:* $Neg(s_i) = s_j \; / \; j = T - i$
3. *Máximo:* $Max(s_i, s_j) = s_i \; if \; s_i \geq s_j$
4. *Mínimo:* $Min(s_i, s_j) = s_i \; if \; s_i \leq s_j$

Por último, también se definen operadores de agregación lingüísticos como, por ejemplo, el operador *Linguistic Weighted Averaging* (LWA) [20].

En nuestro modelo, además, aplicamos el modelado lingüístico difuso basado en 2-tuplas [21], un enfoque que reduce la pérdida de información que se produce en el enfoque ordinal (ya que la información se representa usando un modelo continuo en lugar de uno discreto). En este contexto, si como resultado de una agregación simbólica de información lingüística [22] obtenemos un valor $\beta \in [0, g]$ donde $\beta \notin \{0, \ldots, g\}$, entonces podemos definir una función de aproximación que expresa el valor obtenido como un elemento del conjunto *S*. El elemento clave de este enfoque es el concepto de *traducción simbólica*, que básicamente se define como la diferencia entre la información expresada por β y la etiqueta lingüística $s_i \in S$ más cercana [21].

Por último, para el desarrollo de nuestro modelo, recurrimos al análisis de coocurrencia de términos, que es una técnica ampliamente utilizada en diferentes dominios. Por ejemplo, en el campo de la recuperación de información se utiliza para relacionar palabras que aparecen en la misma frase o párrafo [23]; y en ciencias de la documentación esta técnica es fundamental para realizar estudios de co-autoría [24] o análisis geopolíticos basados en la manera en que diferentes organizaciones se citan unas a otras [25].

En nuestro caso, pretendemos crear una red de coocurrencia para determinar cómo se relacionan entre sí los términos que utilizan los usuarios de una comunidad científica para expresar sus intereses de investigación, y aprovechar esta información para generar recomendaciones sobre recursos de áreas tangencialmente relacionadas.

4 ARQUITECTURA DEL MODELO

El sistema de recomendaciones que aquí presentamos es una mejora de un modelo previo [26] al que se han añadido nuevas funcionalidades y en el que se modifica la medida de similitud utilizada para generar las recomendaciones. El modelo actual presenta tres agentes software que se distribuyen en una arquitectura jerárquica de 5 niveles funcionales:

- *Nivel de usuario*, donde los individuos interactúan con el sistema recibiendo las recomendaciones y generando *feedback*.
- *Nivel de interfaz*, donde un agente de interfaz realiza operaciones simples de filtrado y actúa como intermediario entre los usuarios y el agente de tarea del siguiente nivel.
- *Nivel de tarea*, donde el agente de tarea genera las recomendaciones.
- *Nivel de agentes de información*, donde el agente de información (idealmente uno por usuario) accede al repositorio de documentos jugando el papel de mediador entre los recursos de información y el agente de tarea.
- *Nivel de recursos*, en el que se especifican los diferentes recursos de información a los que el sistema tiene acceso. En concreto se definen:
  1. Un conjunto de perfiles de usuario extraídos de la web *academia.edu*, una comunidad virtual de científicos donde sus miembros, entre otros datos, definen cuáles son sus intereses de investigación (ya estén relacionados con su área de especialización o con otros dominios).
  2. Un repositorio de recursos de carácter multidisciplinar.
  3. Un tesauro sobre "*Bibliotecas digitales*" desarrollado con el vocabulario semántico SKOS.

5 CÓMO SE GENERAN LAS RECOMENDACIONES

En trabajos previos, hemos abordado el problema de la satisfacción de necesidades especiales de información de la comunidad científica. En concreto, realizamos una aproximación a la recomendación de recursos de áreas afines a través de la reinterpretación de la medida de similitud y, para ello, propusimos una función μ de centrado gaussiana [27] (Fig. 1) que se define de la siguiente manera:

$$\mu\left(Sim\left(p_i, p_j\right)\right) = e^{\left[Sim\left(p_i, p_j\right) - k\right]^2} \quad (2)$$

donde $Sim\ (p_i, p_j)$ es la medida de similitud entre los recursos p_i y p_j , y k representa el valor central para el

cual la similitud es relevante para generar una recomendación (en este caso *k*=0'5).

Para generar las recomendaciones se hace una interpretación no lineal de la relevancia, de manera que los recursos recomendados están relacionados con los intereses del individuo, pero no encajan exactamente con estos. En otras palabras, las medidas de similitud extremas (muy grandes o muy pequeñas) son descartadas y son recomendados los recursos que presentan un valor medio de similitud.

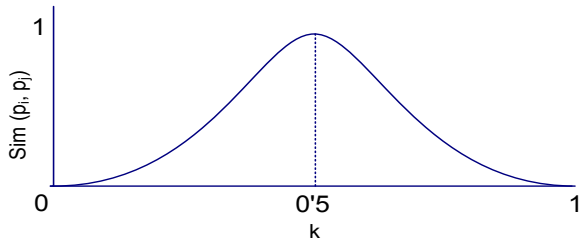


Figura 1: Función de centrado gaussiana

Abundando en esta idea, en este trabajo proponemos un nuevo enfoque que considera otros factores. En lugar de una reinterpretación de la medida de similitud, la nueva medida de relevancia se va a componer de dos factores: el primero de ellos va a ser la similitud semántica que ya proponíamos en el sistema (teniendo en cuenta la posición de los términos en el tesauro); y, el otro factor va a considerar la frecuencia de coocurrencia entre términos.

Partiendo de estas premisas, la nueva medida de similitud la expresamos de la siguiente manera:

$$Sim'(t_i, t_j) = Sim(t_i, t_j) + (1 - Sim(t_i, t_j))\, f \qquad (3)$$

donde $Sim(t_i, t_j)$ es el valor de similitud definida por Oldakowski y Bizer [28] y f es la frecuencia de coocurrencia normalizada al rango [0, 1]. De esta manera, cuando la similitud es baja pero el número de coocurrencias alto, el documento tiene posibilidades de ser recomendado. Además, en el caso de que ambos términos presenten un valor alto, siempre va a prevalecer la similitud semántica.

El valor de relevancia del perfil P_i con el recurso R_j viene definido por la siguiente función:

$$Rel(P_i, R_j) = \sum_{k=1}^{Min(N,M)} \frac{H_k(Sim'(t_i, t_j))}{Max(N, M)} \qquad (4)$$

donde H_k es una función que obtiene la *k* máxima similitud entre los términos $t_i \in P_i$ y $t_j \in R_j$; *N* es el número de preferencias definidas en el perfil de usuario y *M* el número de términos extraídos del análisis de la referencia bibliográfica del recurso, con las que son comparadas dichas preferencias.

5.1. PROCEDIMIENTO

La estrategia para la implementación del sistema es añadir a los términos, que originariamente forman parte del tesauro, nuevos términos que cumplan dos condiciones:

1. Deben aparecer conjuntamente con el concepto "*Digital libraries*" en el perfil de los usuarios.
2. Deben aportar algún tipo de información relacional.

Para ello, es necesario llevar a cabo un análisis de coocurrencia y obtener la frecuencia con la que estos términos aparecen mencionados simultáneamente como "*temas de interés*" en los perfiles de los usuarios. De esta manera, podemos definir para el tesauro nuevas relaciones asociativas ponderadas que no son estáticas, ya que dependen de las preferencias definidas por la comunidad en un momento determinado.

Como fuente para obtener los datos de análisis, hemos elegido la web *academia.edu*, en la cual existe una comunidad de usuarios pertenecientes al mundo académico y científico de diferentes áreas. Cada uno de los miembros de la comunidad dispone de un perfil en el que puede expresar cuáles son los temas de investigación que más le interesan (no necesariamente de su dominio). En una primera aproximación al problema hemos decidido analizar la comunidad de usuarios que entre sus preferencias definen el término "*Digital libraries*", que es el que vamos a considerar como dominio de nuestro sistema. El valor de la información que podemos extraer de los perfiles reside precisamente en que disponemos de los intereses de investigadores de diferentes áreas que comparten un interés común por las bibliotecas digitales, lo cuál nos puede dar una idea bastante certera de la relación que existe entre este dominio y otras áreas afines.

Para la obtención del conjunto de datos de análisis, se ha desarrollado una aplicación JAVA que recorre los perfiles de los usuarios del sitio web, para recuperar aquellos en cuyos intereses aparece "*digital libraries*". Una vez hemos discriminado los perfiles, se procede a capturar todos los intereses de investigación que aparecen simultáneamente con "*digital libraries*" para obtener el conjunto de términos que caracteriza al domino. Este listado, aún sin refinar, está formado por alrededor de 6500 términos diferentes, los cuales se someterán a un proceso de normalización para eliminar términos superfluos o redundantes y aquellos que ofrezcan poca información relacional (es decir, los que presenten una baja frecuencia de coocurrencia con "*Digital libraries*"). Con el listado resultante, se procederá a calcular la matriz de coocurrencia, cuyas frecuencias serán un elemento determinante en el paso siguiente.

Éste consiste en añadir los términos del listado, ya normalizados, a la estructura del tesauro. Para ello, se procederá, en primera instancia, a equiparar estos términos con los conceptos del tesauro para, a continuación, refinar los resultados de una manera semiautomática, mediante

técnicas de *ontology learning* y el trabajo de categorización de expertos en el dominio. Cada equiparación positiva o cada nuevo término incorporado añadirá a la estructura del tesauro, además, todos los términos con los que mantiene una relación asociativa (definida a partir del análisis de coocurrencia). Esta relación asociativa, a su vez, estará caracterizada por la frecuencia de coocurrencia *f* que existe entre los dos términos y que servirá como factor de ponderación para calcular la similitud entre términos, tal y como hemos explicado en el apartado anterior.

Cuando ingresan en el sistema nuevos recursos, se procede a analizar los elementos con mayor capacidad informativa de la referencia bibliográfica (título, resumen y palabras clave) y se equiparan semánticamente con los del perfil del usuario activo. Aquellos recursos cuyo valor de relevancia sobrepase el umbral mínimo requerido (definido por defecto en 0.5) serán recomendados a los usuarios.

Estos serán presentados al usuario por orden de relevancia y tendrán asociado un valor lingüístico difuso basado en 2-tuplas [21], para facilitar la tarea de diferenciar de una manera precisa los recursos más pertinentes a los intereses definidos en su perfil.

6 CONCLUSIONES Y TRABAJO FUTURO

En este trabajo presentamos el estado actual de la investigación sobre el desarrollo de un sistema de recomendaciones para la comunidad científica e investigadora que permite satisfacer necesidades especiales de información, proporcionando los medios para sugerir recursos de áreas emergentes, en los que encontrar nuevas oportunidades de colaboración con personas o grupos de investigación de áreas afines. Dicho con otras palabras, queremos ofrecer al investigador una herramienta que le permita descubrir tendencias en lo referente a las líneas de investigación multidisciplinares que involucran su área de especialización. Por esta razón, el análisis de coocurrencias debe realizarse con una cierta periodicidad para detectar estas variaciones en los intereses de la comunidad (y que estos se reflejen en el tesauro del sistema).

El presente modelo se basa en un sistema previo al que se han añadido algunas mejoras, como la utilización de redes de coocurrencia para modificar la función que permite calcular la similitud de términos y, de esta forma, poder incluir entre las recomendaciones recursos pertenecientes a áreas tangencialmente relacionadas con el dominio del sistema. El tesauro, así, se convierte en una herramienta dinámica que no sólo describe un dominio, sino que también ofrece información relacional que representa una instantánea de cómo se interrelaciona dicho dominio con áreas afines.

En este momento nos encontramos en la fase de normalización de términos para generar la red de coocurrencia y proceder a la extensión del tesauro. Una vez completada esta fase, se procederá a la evaluación del sistema para comprobar su rendimiento. Esta evaluación se hará de acuerdo a términos de precisión [29] y exhaustividad [30] y F1 [31], medidas usadas tradicionalmente para la evaluación de la calidad del conjunto de recursos sugeridos por los SR y de filtrado. Posteriormente, estos resultados serán comparados con los indicadores obtenidos para el sistema previo en el que se basa esta propuesta y, de esta manera, determinar qué aproximación es más eficiente.
Dado que el análisis de coocurrencia toma en consideración los intereses reales expresados por la propia comunidad de usuarios (en lugar de una mera estimación de la relevancia), esperamos que los resultados de este nuevo enfoque supongan una mejora considerable con respecto al modelo anterior.

Agradecimientos

Este trabajo está financiado por los programas de *"Formación de doctores"* y de *"Perfeccionamiento de doctores"* del plan propio de la Universidad de Granada y desarrollado en el marco del proyecto Semantic Selective Dissemination of Information (TIN2010-22145-C02-01), financiado por el Ministerio de Ciencia e Innovación.

Referencias

[1] "Research and innovation: Building Europe's future," *Europa: Gateway to the European Union*. [Online]. Available: http://europa.eu/pol/rd/. [Accessed: 31-Oct-2011].

[2] W. M. Cohen and D. A. Levinthal, "Absorptive capacity: a new perspective on learning and innovation," *Administrative science quarterly*, pp. 128–152, 1990.

[3] C. L. Borgman, *Scholarship in the digital age: Information, infrastructure, and the Internet*. The MIT Press, 2007.

[4] C. L. Palmer, L. C. Teffeau, and C. M. Pirmann, "Scholarly information practices in the online environment," *Report commissioned by OCLC Research. Published online at: www. oclc. org/programs/publications/reports/2009-02. pdf*, 2009.

[5] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-commerce recommendation applications," *Data mining and knowledge discovery*, vol. 5, no. 1, pp. 115–153, 2001.

[6] A. Popescul, L. H. Ungar, D. M. Pennock, and S. Lawrence, "Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments," 2001.

[7] F. Long, Y. F. Zhang, and F. Hu, "A Personalized Recommender Systems Framework Based on Ontology and Bayesian Network in E-Commerce,"

Advanced Materials Research, vol. 143, pp. 961–965, 2011.

[8] L. Zhuhadar, O. Nasraoui, R. Wyatt, and E. Romero, "Multi-model ontology-based hybrid recommender system in e-learning domain," in *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03*, 2009, pp. 91–95.

[9] C. Porcel, A. G. LóPez-Herrera, and E. Herrera-Viedma, "A recommender system to promote collaborative research groups in an academic context," 2008, pp. 847-852.

[10] C. Porcel, A. Tejeda-Lorente, M. A. Martínez, and E. Herrera-Viedma, "A hybrid recommender system for the selective dissemination of research resources in a Technology Transfer Office," *Inf. Sci.*, vol. 184, no. 1, pp. 1–19, Feb. 2012.

[11] E. Herrera-Viedma, L. Martinez, F. Mata, and F. Chiclana, "A consensus support system model for group decision-making problems with multigranular linguistic preference relations," *Fuzzy Systems, IEEE Transactions on*, vol. 13, no. 5, pp. 644–658, 2005.

[12] Á. García-Crespo, J. L. López-Cuadrado, R. Colomo-Palacios, I. González-Carrasco, and B. Ruiz-Mezcua, "Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain," *Expert systems with applications*, vol. 38, no. 10, pp. 13310-13319, 2011.

[13] I. E. Liao, W. C. Hsu, M. S. Cheng, and L. P. Chen, "A library recommender system based on a personal ontology model and collaborative filtering technique for English collections," *Electronic Library, The*, vol. 28, no. 3, pp. 386–400, 2010.

[14] T. Berners-Lee, J. Hendler, and O. Lassila, "A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.

[15] J. Hendler, "Agents and the semantic web," *Intelligent Systems, IEEE*, vol. 16, no. 2, pp. 30–37, 2001.

[16] G. Klyne and J. J. Carroll, "Resource Description Framework (RDF): Concepts and Abstract Syntax," 2004.

[17] T. R. Gruber and others, "Toward principles for the design of ontologies used for knowledge sharing," *International Journal of Human Computer Studies*, vol. 43, no. 5, pp. 907–928, 1995.

[18] N. Guarino, "Formal Ontology in Information Systems," in *Formal Ontology in Information Systems*, Amsterdam: IOS Press, 1998, pp. 3-17.

[19] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning–I," *Information sciences*, vol. 8, no. 3, pp. 199–249, 1975.

[20] F. Herrera and E. Herrera-Viedma, "Aggregation operators for linguistic weighted information," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 27, no. 5, pp. 646–656, 1997.

[21] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *Fuzzy Systems, IEEE Transactions on*, vol. 8, no. 6, pp. 746–752, 2000.

[22] F. Herrera, E. Herrera-Viedma, and J. Verdegay, "Direct approach processes in group decision making using linguistic OWA operators," *Fuzzy Sets and Systems*, vol. 79, no. 2, pp. 175–190, 1996.

[23] L. B. Doyle, "Indexing and abstracting by association," *American Documentation*, vol. 13, no. 4, pp. 378–390, 1962.

[24] M. E. J. Newman, "Scientific collaboration networks. I. Network construction and fundamental results," *Physical Review E*, vol. 64, no. 1, p. 016131, 2001.

[25] V. Herrero-Solana and J. M. Morales-del-Castillo, "Mapas 'geopolíticos' de internet: aplicación de las nuevas técnicas de representación de la información," *Ciência da informação*, vol. 33, no. 3, pp. 69-75, 2005.

[26] J. M. Morales-del-Castillo, E. Peis, J. M. Moreno, and E. Herrera-Viedma, "D-Fussion: A semantic selective disssemination of information service for the research community in digital libraries," *Information Research*, vol. 14, no. 2, 2009.

[27] R. R. Yager, "Centered OWA operators," *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, vol. 11, no. 7, pp. 631–639, 2007.

[28] B. Oldakowski, R. C., "SemMF: A Framework for Calculating Semantic Similarity of Objects. Represented as RDF Graphs," presented at the 4th International Semantic Web Conference (ISWC 2005), Galway, Ireland, 2005, pp. 2-4.

[29] Y. Cao and Y. Li, "An intelligent fuzzy-based recommendation system for consumer electronic products," *Expert Systems with Applications*, vol. 33, no. 1, pp. 230–240, 2007.

[30] C. W. Cleverdon, J. Mills, and M. Keen, *Factors determining the performance of indexing systems: Vol.2, Test Results*. Cranfield: ASLIB Cranfield Project, 1966.

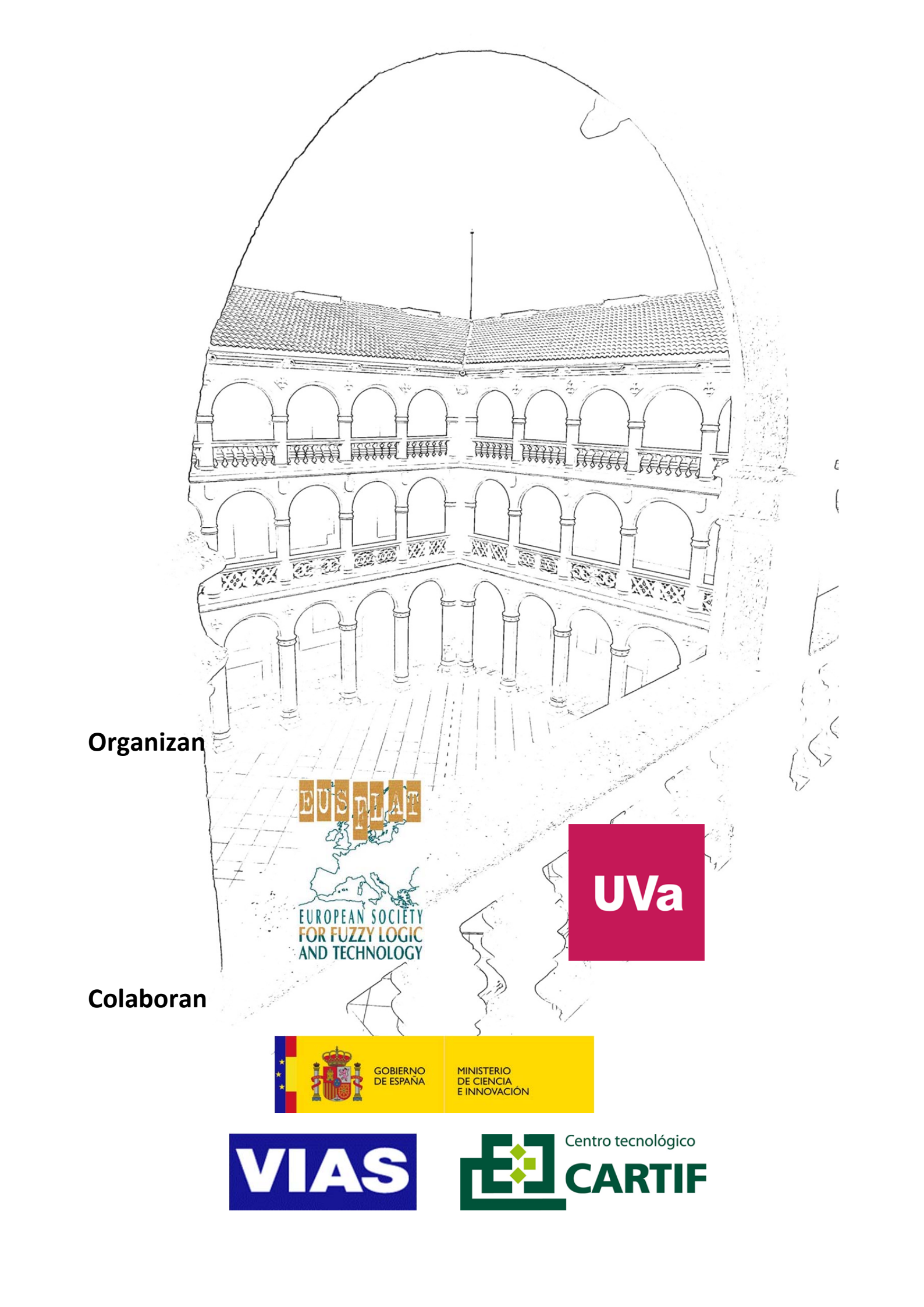
[31] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e-commerce," in *Proceedings of the 2nd ACM conference on Electronic commerce*, 2000, pp. 158–167.

Índice de autores

Alcalá, R. 265,295
Alcalá-Fdez, J. 265,325,534
Alcalde, C. 8
Almonacid, M. 360
Alonso, J. M. 516,585
Alonso, S. 193
Alonso-Martínez, I. 438
Álvarez-Álvarez, A. 614
Andújar-Márquez, J. M. 354
Arca, B. 498
Arenas-Parra, M. 132
Arnedo-Fdez, J. 668
Ausín, T. 684
Aznar-Torres, M. J. 354
Barragán-Piña, A. .J. 354
Barranco, M. 259
Barrenechea, E. 283,504
Barrero, P. R. 674
Barrientos, F. 102, 390
Bedregal, B. 230
Beliakov, G. 545
Bello, R. 77
Benítez, J. M. 96,102
Bergmeir, C. 96
Bilbao-Terol, A. 132
Boixader, D. 43
Bonis-Martín, V. 126
Bordons,C. 108
Bragard, J. 54
Bronselaer, A. 163
Brox-Jiménez, M. 426
Bugarín, A. 644,701
Burillo, P. 54
Burusco, A. 8
Bustince, H. 230,289,486,504,545
Cables Pérez, E. 199
Cabrerizo, F. J. 193
Cadenas, J. M. 414
Calvo, T. 552
Cano-Izquierdo, J. M. 360
Capdevila, C. A. 175
Carmona, C. J. 271
Carmona, P. 596
Caro, R. 205
Casado-Molina, A. M. 223
Casasús, T. 366
Castellanos, D. 241
Castillo, E. 456
Castillo-Ortega, R. 620
Castro, A. 498
Castro, J. 259
Castro, J. L. 596
Cerrón, J. 283
Chamorro-Martínez, J. 480
Chávez, F. 265
Chiclana, F. 235
Contreras, R. 384
Contreras-Bárcena, D. 438
Cordón, O. 342,585
Cornelis, C. 638
Couso, I. 336,528
Cruz, C. 120
Cubillo, S. 48
de Andrés Calle, R. 187
de Baets, B. 504
de la Cal, E. 348
de Pedro Lucio, T. 372
de Soto, A. R. 90, 175
de Tré, G 163
del Jesús, M. J. 271,313
del Rosario Vázquez, M. R. 632
del Val, C. 656,668
Delgado-López, J.A. 721
Derrac, J. 534
Díaz, J. C. 14,156
Díaz, P. 20
Díaz-Hermida, F. 644,701
Diaz-Valenzuela, I. 450
Doña-Fernández, J. 223
Eciolaza, L. 608
Elorza, J. 54
Escaño, J. M. 108
Esteva, F. 144
Falcó-Díaz de Cerio, E. 181
Fdez-Viñas, E. 444
Fernández, A. 289,313,534

Fernández, F. 265
Fernández, J. 230,545
Fernández-Lanza, S. 689
Fernández-Reyes, F. C. 462
Ferreira-Satler, M. 468
Franco, C. A. 84,114,579
Fuente, M. J. 384
Fuentes-González, R. 8,54
Gacto, M. J. 271,295
Galar, M. 283,545
Galende, M. 384
García, D. 277
García, J. 247
García, S. 534
García-Cascales, M. S. 199
García-Honrado, I. 31,37,522
García-Lapresta, J. L. 181
García-Pajares, R. 102,390
García-Tamargo, M. 348
Garmendia Salvador, L. 60,71,126,205
Garmendia-Salvador, A. 205,126
Garrido, L. 331
Garrido, M. C. 414
Godo, L. 144
Godoy, J. 301
Gómez, D. 253,492,564
Gómez-Romero, J. 247
Gómez-Skarmeta, A. 420
González Valdez, M. 408
González, A. 277
González, P. 271
González-Arteaga, T. 187
González-del-Campo, R. 60
González-Hidalgo, M. 510
Gonzalez-Jaime, L. 486
Gutiérrez-Avilés, D. 650
Guzmán, J. L. 402
Harari, O. 656
Hernández, P. 48
Herrera, F. 217,265,289,295,313,602,638
Herrera-Viedma, E. 193
Herrero, J. L. 596
Herreros López, A. 396
Hopp, E. 662
Ibarrola, J. 360
Iglesias, T. 20,26
Janis, V. 1, ... 20,26
Jiménez, M. 132
Jirón, J. 169
Julián-Iranzo, P. 695
Junco, L. 336
Jurio, A. 230,486
Kolesárová, A. 230
Lamata Jiménez, M. T. 199
Lau, R. 462
León, S. 456
Leyva, E. 277
Leyva, E. C. 462
Liern, V. 366
Liu, J. 211
López, V. 313
López-Molina, C. 504
Lucca, M. F. 662
Luengo, J. 534,602
Magdalena, L. 516,585
Marín, N. 620
Márquez, A. 307
Márquez, F. A. 307
Martín, J. 552
Martín, R. 205
Martín-Bautista, M. J. 450
Martínez, L. 187,211,217,259
Martínez, N. 253
Martínez, R. 414
Martínez-Álvarez, F. 650
Martínez-Jiménez, P. 480
Martínez-Moreno, J. 319
Martín-Toral, S. 378
Mas, M. 569
Masegosa, A. 241
Massanet, S. 150,510,558
Mata, F. 235
Mattioli, G. 65
Mayor, G. 552
Medina, J. 14,156
Menéndez-Arduengo, L. 31,37
Menéndez-Domínguez, V. H. 444
Mesiar, R. 230,545
Milanés, V. 301
Mir, A. 510
Molina, C. 620
Molina, J. M. 247
Monserrat, M. 569
Montero, J. 84,114,253,492,564,579,626
Montes, I. 1,26
Montes, S. 1,20,26
Montoto, P. 701
Moraga, C. 591

Morales-del-Castillo, J.M.……721
Moreno-Velo, F. J 432
Murer, M. G. 678
Nachidi, M. 402
Olivas, J. A. 444,462,468,474,632
Onieva, E. 301
Orduna, R. 283
Otero, J. 528
Pagola, M. 486
Palacios, A. M. 325,336
Palomares, I. 211
Pancho, D. P. 37,516,585
Patricio, M. A. 247
Peís, E,……721
Peláez-Sánchez, J. I. 223
Peregrín, A. 307
Pereira-Fariña, M. 608,701
Pérez, J. C. 366
Pérez, I. J. 193
Pérez, R. 277
Pérez-Rastelli, J. 372
Pinto, D. 456
Plá, L. 366
Pons, J. E. 163
Pons, O. 163
Pradera, A. 486,545
Quirin, A. 342,585
Ramírez-Mendoza, M. 408
Ramos-Guajardo, A. B. 522
Ramos-Soto, A. 644
Recasens, J. 43,65,71
Rey, A. 498
Riquelme, J. C. 650
Riquelme, L. 678
Rodríguez, F. 402
Rodríguez, R. 14,156
Rodríguez, R.M. 217
Rojas, K. 564
Roldán, A. 319
Roldán, C. 319
Romero-Chicharro, F. P. ... 438,444,468,474,632
Romero-Zaliz, R. 662,668
Roselló Saurí, L. 181
Rubio-Escudero, C. 650
Rubio-Manzano, C. 695
Ruiz-Aguilera, D. 569
Sáez, J. A. 602
Sáinz-Palmero, G. I. 102,378,384,390
Salvador, A. 71
San José-Alonso, S. 378
Sánchez, L. 336,342,528
Sánchez-Lozano, J. 199
Sánchez-Solano, S. 426
Sánchez-Valdés, D. 480,614,620
Santano-Ruiz, R. 432
Santos-Peñas, M. 372
Sanz, J. A. 289
Sanz-Villanueva, J. C. 396
Sedano, J. 348
Seising, R. 707
Serrano, C. 307
Serrano, M. A. 247
Serrano-Guerrero, J. 444,468,474
Shang-Ming Zhou 235
Silva, R. C. 120
Sobrino, A. 715
Sols, J. 575
Soto-Hidalgo, J. 480
Soto-Villaverde, A. 468,474
Tadeo, F. 402
Terroso-Sáenz, F. 420
Tinguaro-Rodríguez, J. 84,114,564,579,626
Torrens, J. 138,150, 569
Torres-Blanc, C. 48
Tovar, M. 456
Trawinski, K. 342
Triguero, I. 96
Trillas, E. 31,540
Triviño, G. 608,614
Urrutia, A. 169
Valdés-Vela, M. 420
Valero, O. 558
Velasco, F. 96
Verbiest, N. 638
Verdegay, J. L. 77,120
Vicente-Riera, J. 138
Vidal-Aguiar, J. C. 701
Vila, M. A. 450
Vilariño, D. 456
Villacorta, P. 241
Villagrá, J. 301
Villar, J. R. 348
Vitoriano, B. 626
Yang Xu 211
Yankilevich, P. 674,678
Yáñez, J. 492
Zarrazola, E. 492
Zwir, I. 656,668,674



Organizan

EUSFLAT

EUROPEAN SOCIETY FOR FUZZY LOGIC AND TECHNOLOGY

UVa

Colaboran

GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN

VIAS

Centro tecnológico CARTIF