CAEPIA 2024

XX Conferencia de la Asociación Española para la Inteligencia Artificial

A Coruña

19-21 Junio

CAEPIA 2024

XX Conferencia de la Asociación Española para la Inteligencia Artificial

A Coruña, 19 al 21 de Junio de 2024

PATROCINA

//ABANCA

XX Conferencia de la Asociación Española para la Inteligencia Artificial

CAEPIA 2024

19-21 de junio de 2024

A Coruña, España

Amparo Alonso, Bertha Guijarro, Óscar Fontenla, Noelia Sánchez, Beatriz Pérez, David Camacho, Juan R. Rabuñal, Manuel Ojeda-Aciego, Jesús Medina, José Riquelme, Alicia Troncoso, Eva Onaindia, Alberto Bugarín, José Antonio Gámez, María José Del Jesús Díaz, Luis Martínez, Francisco Bellas, Sara Guerreiro, Alejandro Rodríguez, José Alberto Benítez, María del Mar Marcos

Editores

Amparo Alonso
Universidade da Coruña
A Coruña, España

Bertha Guijarro
Universidade da Coruña
A Coruña, España

Óscar Fontenla
Universidade da Coruña
A Coruña, España

Noelia Sánchez
Universidade da Coruña
A Coruña, España

Beatriz Pérez
Universidade da Coruña
A Coruña, España

David Camacho
Universidad Politécnica de Madrid
Madrid, España

Juan R. Rabuñal
Universidade da Coruña
A Coruña, España

Manuel Ojeda-Aciego
Universidad de Málaga
Málaga, España

Jesús Medina
Universidad de Cádiz
Cádiz, España

José Riquelme
Universidad de Sevilla
Sevilla, España

Alicia Troncoso
Universidad Pablo de Olavide
Sevilla, España

Eva Onaindia
Universitat Politècnica de València
Valencia, España

Alberto Bugarín
Universidade de Santiago de Compostela
A Coruña, España

José Antonio Gámez
Universidad de Castilla la Mancha
Albacete, España

María José Del Jesús Díaz
Universidad de Jaén
Jaén, España

Luis Martínez
Universidad de Jaén
Jaén, España

Francisco Bellas
Universidade da Coruña
A Coruña, España

Sara Guerreiro
Universidade da Coruña
A Coruña, España

Alejandro Rodríguez
Universidad Politécnica de Madrid
Madrid, España

José Alberto Benítez
Universidad de León
León, España

María del Mar Marcos
Universitat Jaume I
Castelló de la Plana, España

ISBN: 978-84-09-62724-0

Presentación de CAEPIA 2024

La Inteligencia Artificial se ha convertido en una revolución, no sólo es un cambio tecnológico, ya que la economía, el escenario de empleo, las formas de relacionarnos en sociedad, los equilibrios geopolíticos, etc. están también sufriendo cambios importantes. Las disrupciones durante los últimos años han sido continuas, con la explosión de la IA generativa y la reciente publicación de la regulación europea de Inteligencia Artificial, en pos de una IA más confiable, sostenible y al servicio de las personas.

CAEPIA (Conferencia de la Asociación Española para la Inteligencia Artificial, con periodicidad bianual), es el foro en el que la comunidad científica española presenta y discute sus últimos avances en IA y desde sus orígenes está abierto a investigadores de todo el mundo. Este volumen contiene un conjunto de artículos seleccionados de entre los enviados a CAEPIA 2024, la XX Conferencia de la Asociación Española de Inteligencia Artificial, celebrada en A Coruña, del 19 al 21 de Junio de 2024. La conferencia comenzó a celebrarse en 1985, con ediciones anteriores que se celebraron en Alicante, Málaga, Murcia, Gijón, San Sebastián, Santiago de Compostela, Sevilla, La Laguna, Madrid, Albacete, Salamanca, Granada y Málaga. Esta edición nos permite celebrar una efeméride especial, que es el 40 aniversario de la Asociación Española para la IA (AEPIA), ente organizador del congreso. Los investigadores asistentes a la conferencia han optado por cinco tipos de contribuciones: trabajos inéditos de investigación en idioma inglés para un volumen en la serie Lecture Notes in Artificial Intelligence de Springer, trabajos inéditos de investigación para estas actas, trabajos destacados ya publicados (Keyworks), proyectos de doctorado (Doctoral Consortium y Frances Allen, este último para las tesis doctorales realizadas por una mujer), y desarrollos de aplicaciones móviles. La conferencia acoge tanto investigación teórica como metodológica, técnica o aplicada

Dentro de CAEPIA, junto con las sesiones generales, se celebran diversos congresos y talleres federados. En particular, en esta edición, se celebraron el XXI Congreso Español Sobre Tecnologías y Lógica Fuzzy (ESTYLF), el XV Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB), el XI Simposio de Teoría y Aplicaciones de la Minería de Datos (TAMIDA), el I Congreso Español en Sistemas de Recomendación y varios Talleres/workshops: I Workshop de la Sociedad Española de Inteligencia Artificial en Biomedicina (IABiomed 2024), I Taller de IA para la Educación (TIAE), y un tutorial en Large Scale Data Analytics (LSDA). También se organiza el Doctoral Consortium (DC), un foro para que estudiantes de doctorado puedan interactuar con otros investigadores discutiendo sobre sus avances y planes de tesis. Por último, para subrayar la importancia práctica de la IA a nivel de aplicaciones móviles, en CAEPIA 2024 tuvo lugar la 5ª Competición de Apps con técnicas de IA.

CAEPIA tiene como objetivo ser reconocida como uno de los congresos de referencia en IA, manteniendo los estándares de alta calidad de ediciones previas y el modelo de revisión por pares de todos los trabajos. A partir de estas revisiones, los responsables de área, presidentes de congresos y organizadores de talleres y sesiones especiales propusieron un número final de artículos que fueron analizados y aprobados por los editores de este volumen. En total, se recogen en estas actas un total de 100 trabajos de investigación originales, además de 59 Keyworks, y 24 contribuciones a los premios Doctoral Consortium, Frances Allen y el concurso de Apps.

CAEPIA 2024 invitó a dos destacados investigadores a impartir una charla plenaria. Nuestros dos ponentes plenarios fueron el Prof. José Mª Lassalle ("IA y autenticidad humana") y Amparo Alonso Betanzos ("Hacia una IA más sostenible"). Además, ESTYLF contó con otro investigador invitado de relevancia, Francesc Esteva ("Algunas reflexiones sobre investigación y sociedad"). Tras la conferencia de apertura, tuvo lugar una mesa redonda sobre Retos empresariales de la IA, en la que tuvimos el honor de contar con destacadas empresas influyentes en el sector como Inditex, Abanca, NTT Data, Minsait, el Clúster TIC Galicia y DataSpartan.

Nuestra conferencia se celebró como un gran evento dentro de uno aún mayor: el VII Congreso Español de Informática (CEDI), que también contó con charlas plenarias muy interesantes, en el

campo de la IA y la Robótica. AEPIA y los organizadores de CAEPIA 2024 reconocieron las mejores tesis doctorales y artículos originales en eventos federados escritos tanto por investigadores consolidados como por estudiantes. Como en ediciones anteriores, y con el objetivo de promover la presencia de mujeres en la investigación de IA, el premio Frances Allen también reconoció las dos mejores tesis doctorales defendidas por una mujer durante los dos últimos años. Los editores de este volumen quieren agradecer a las numerosas personas su contribución al éxito de CAEPIA 2024: autores, miembros de los comités científicos y los comités de programa, ponentes invitados, organizadores de eventos, gestores de medios electrónicos, etc. Quisiéramos agradecer asimismo el trabajo incansable del comité organizador, de nuestros patrocinadores (Abanca), del equipo de Springer y de AEPIA por su apoyo. Por último, pero no menos importante, en nombre de los participantes de CAEPIA 2024, Amparo Alonso Betanzos, Bertha Guijarro Berdiñas (co-presidentas), así como Noelia Sánchez Maroño y Beatriz Pérez Sánchez (responsables de este volumen) dan las gracias a la organización de CEDI, la Universidad de A Coruña (sede local de la conferencia) y a toda la comunidad española que trabaja en IA (y sus numerosos colaboradores extranjeros) por hacer de este evento, una vez más, y en esta fecha tan especial de 40 aniversario, un verdadero éxito.

Amparo Alonso Betanzos

Bertha Guijarro Berdiñas

Co-presidentas de CAEPIA 2024

Presentación de la Presidenta de la Asociación Española para la Inteligencia Artificial

Es para mi un placer presentar esta nueva edición de la Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA) que se celebra en junio de 2024 en A Coruña, en este año tan especial en el que la Asociación Española para la Inteligencia Artificial (AEPIA) celebra su 40 aniversario. Un grupo de investigadores españoles, que trabajaban en diversos campos de la Inteligencia Artificial (IA), coinciden en Karlsruhe (Alemania) en agosto de 1983, durante la celebración del congreso International *Joint Conference on Artificial Intelligence* (IJCAI1983). En ese momento se estaba gestando una asociación de IA a nivel europeo, y este grupo de investigadores pioneros considera necesario la organización a nivel español, para promover la comunicación entre los investigadores de IA en España y para tener una representación en la asociación que se estaba gestando a nivel europeo. En 1984, estos investigadores organizaron unas Jornadas Técnicas de IA en Buitrago (Madrid), en las cuales nace AEPIA, y que fueron el embrión de nuestra actual CAEPIA. Hoy, mis palabras son de sincera felicitación, pero también de gratitud a los investigadores de muchas generaciones, que han contribuido desde aquel lejano 1984 hasta nuestros días a la historia de AEPIA, construyendo una asociación con muchos logros y muchos retos aún por delante.

En esta XX CAEPIA están representadas las diferentes líneas de investigación de los grupos de investigación en IA en España, tanto en sus sesiones generales como en sus talleres y en las diferentes conferencias que integra: XXI Congreso Español de Tecnologías y Lógica Difusa (ESTYLF), XV Congreso Español de Metaheurística y Algoritmos evolutivos y bioinspirados (MAEB), el XI Simposium de Teoría y Aplicaciones de Minería de Datos (TAMIDA) y el I Congreso Español en Sistemas de Recomendación. AEPIA siendo consciente de la importancia que tienen los jóvenes investigadores, incluye en CAEPIA actividades que contribuyan al desarrollo del talento emergente como es el caso del Doctoral Consortium, los premios a los jóvenes investigadores autores de los mejores artículos, o a las mejores apps y videos divulgativos de IA. En AEPIA también contribuimos con acciones de género positivas, mediante los premios Frances Allen a las mejores tesis doctorales realizadas por mujeres, en un intento de reducir la significativa brecha de género existente en la IA e incorporar ese talento. Además, en esta edición de CAEPIA se ha organizado por primera vez una mesa redonda de retos empresariales de la IA, con la que pretendemos conocer de primera mano las necesidades y problemas a los que se enfrentan las empresas, acercar la IA al mundo empresarial, y establecer colaboraciones para el desarrollo conjunto de proyectos de IA.

Las dos conferencias plenarias de la XX CAEPIA se enmarcan dentro de dos retos de actualidad para la IA, como es la IA ética y la IA sostenible. Precisamente con respecto a la IA ética, estamos en un momento histórico, ya que en mayo de 2024 la Unión Europea ha aprobado la ley de inteligencia artificial que se aplicará de forma progresiva hasta 2026, que será cuando entre finalmente en vigor. En esta ley se etiquetan las aplicaciones basadas en IA en función del riesgo que genera para las personas, identificando así los sistemas de alto riesgo que solo se podrán usar si respetan los derechos fundamentales, y prohibiendo algunos sistemas como aquellos basados en biometría, puntuación de personas en función de su comportamiento, etc. En cuanto a la IA sostenible, dado el alto consumo energético que tiene el entrenamiento de algunos modelos de IA surge la necesidad de desarrollar nuevos modelos, que sean mucho más eficientes computacionalmente, e instaurar la eficiencia energética como medida de evaluación, con el objeto de caminar hacia una IA responsable con el medio ambiente.

El futuro de la IA se nos presenta apasionante, pudiendo contribuir a mejorar la calidad de vida de los ciudadanos, combatir el cambio climático y avanzar hacia una sociedad más sostenible y justa. Para lograrlo se necesita talento y estamos seguros de que la AEPIA aportará su grano de arena.

Alicia Troncoso Lora
Presidenta de AEPIA

Ponentes Plenarios

José María Lassalle

José María Lassalle (Santander, 1966) es doctor en Derecho. Inició su trayectoria profesional como investigador y profesor en la Universidad de Cantabria y en la Universidad Carlos III de Madrid. Fue coordinador científico del Centro de Estudios Hispánicos e Iberoamericanos de la Fundación Carolina y más tarde, director de esta institución. En 2004 inició su actividad política como diputado en el Congreso por Cantabria. Compatibilizó su dedicación a la política con la actividad docente en la Universidad San Pablo-CEU y en la Universidad Rey Juan Carlos de Madrid. En 2011 fue nombrado secretario de Estado de Cultura y en 2016 de Agenda Digital. En julio de 2018 abandonó la política.

Es autor de numerosos ensayos y publicaciones académicas sobre pensamiento político y filosofía anglosajona y escribe en El País y La Vanguardia. También es colaborador de Las mañanas de RNE y la SER. Actualmente es consultor privado y analista; Director del Foro de Humanismo Tecnológico de ESADE; miembro del Open internet Governance Institute de ESADE, vocal de la Junta Directiva del Cercle d´Economia de Barcelona, miembro del Consejo Asesor de AMETIC, patrono de la Biblioteca Nacional y de la Fundación Hermes. Es profesor de Filosofía del Derecho en la Universidad Pontificia de Comillas (ICADE).

Entre sus últimos ensayos destacan El liberalismo herido. Reivindicación de la libertad frente a la nostalgia del autoritarismo y Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital, ambos con Arpa editorial.

IA y autenticidad humana

La reflexión que planteo es la necesidad de desarrollar una política pública que defina una ética de la autenticidad humana que, partiendo del principio de responsabilidad de Hans Jonás, atribuya a los seres humanos la tarea de gobernar la IA desde un propósito de preservación de las bases antropológicas y culturales que, según Hannah Arendt, definen la condición humana.

Amparo Alonso Betanzos
Universidade da Coruña, España

Amparo Alonso Betanzos es Catedrática de Universidad en el área de Ciencias de la Computación e Inteligencia Artificial, en la Universidade da Coruña (UDC). Sus líneas de investigación actuales son la Inteligencia Artificial Confiable y Explicable, el desarrollo de modelos de aprendizaje máquina sostenibles, y los modelos basados en agentes para la sostenibilidad, entre otros. Ha sido Vicedecana y Coordinadora Erasmus, directora del Departamento de Computación, Coordinadora de la Especialidad de Sistemas Inteligentes del Master en Informática y Coordinadora del Máster Universitario en Bioinformática para Ciencias de la Salud, en la Facultad de Informática de la UDC, así como adjunta al Rector para la Inteligencia Artificial. Como investigadora ha publicado más de 270 artículos en revistas científicas y congresos internacionales, libros y capítulos de libros. Ha participado en más de 30 proyectos de investigación y transferencia competitivos autonómicos, nacionales y europeos. Fue presidenta de la Asociación Española de Inteligencia Artificial desde 2013-2021. Ha participado como miembro del Grupo de Trabajo en Inteligencia Artificial, del Ministerio de Ciencia, Innovación y Universidades para la redacción de la Estrategia Española de I+D+I en Inteligencia Artificial en 2018. En la actualidad es miembro de CAIA, la Comisión Asesora del Gobierno en Inteligencia Artificial, desde 2020, así como Miembro del Comité Español de Ética de la Investigación, desde 2023. Es Académica correspondiente de la Real Academia de Ciencias Exactas, Físicas y Naturales de España desde 2023. Es Senior Member de IEEE y de ACM y miembro del Advisory Board of the Chair of AI and Democracy de la UE. Ha recibido varios premios, entre ellos el Helena Rubinstein-UNESCO "Women in Science" en España y finalista europea (1998), Premio Galicia TIC a la Innovación Digital (2004), Premio Galicia TIC a la trayectoria profesional (2019), Premio Josefa Wonenburger Planells de la Xunta de Galicia, (2020), Premio Gallega del año 2020, Grupo Correo Gallego y Colegiada de Honor del Colegio de Ingenieros Informáticos de Galicia en 2023.

Hacia una IA más sostenible

El éxito de la Inteligencia Artificial (IA) hasta ahora ha dependido del desarrollo de modelos cada vez más precisos, pero también más complejos, con un mayor número de parámetros a estimar. La transparencia y explicabilidad de los modelos son menores, y el costo energético resultante de entrenarlos y ejecutarlos ha aumentado significativamente, con estimaciones que sugieren que para 2030 la IA podría ser responsable de más del 30% del consumo energético del planeta. En este contexto, surge la IA verde y responsable, caracterizada por huellas de carbono más reducidas, tamaños de modelo más pequeños, menor complejidad computacional y mayor transparencia. Existen diversas estrategias para lograrlo, como proporcionar algoritmos datos de mayor calidad, desarrollar modelos más eficientes o mejorar la eficiencia energética de los modelos. Estos temas se presentan como elementos clave para avanzar hacia una IA más ética y responsable, promoviendo la democratización de la tecnología, fortaleciendo la confianza de los ciudadanos en su uso y avanzando en el cumplimiento de las regulaciones de la UE.

Organización

Presidentas

Amparo Alonso	Universidade da Coruña, España
Bertha Guijarro	Universidade da Coruña, España

Presidenta del Comité de Premios

Eva Onaindía	Universidad Politécnica de Valencia, España

Presidente de Talleres y Trabajos Destacados

Óscar Fontenla	Universidade da Coruña, España

Presidente de la Sesión General de CAEPIA

Óscar Fontenla	Universidade da Coruña, España

Presidentes de Competiciones

Alberto Bugarín	Universidad de Santiago de Compostela, España
Juan Antonio Gámez	Universidad de Castilla La Mancha, España

Responsables de Actas LNAI

Elena Hernández	Universidade da Coruña, España
Verónica Bolón	Universidade da Coruña, España

Presidentes del Doctoral Consortium

José Riquelme	Universidad de Sevilla, España
María José Del Jesús Díaz	Universidad de Jaén, España

Responsables de Area de CAEPIA

Óscar Corcho	Universidad Politécnica de Madrid, España
Javier del Ser	TECNALIA, España
Juan Manuel Fernández Luna	Universidad de Granada, España
Jesús García Herrero	Universidad Carlos III de Madrid, España
Carlos Gómez Rodríguez	Universidade da Coruña, España
Serafín Moral Callejón	Universidad de Granada, España
José Luis Pérez de la Cruz Molina	Universidad de Málaga, España
Petia Radeva	Universidad de Barcelona, España
Camino Rodríguez Vela	Universidad de Oviedo, España

Presidentes de Conferencias Federadas

XXII Congreso Español sobre Tecnologías y Lógica Difusa (ESTYLF)

Manuel Ojeda-Aciego	Universidad de Málaga, España
Jesús Medina	Universidad de Cádiz, España

XV Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB)

Juan Ramón Rabuñal	Universidade da Coruña, España
David Camacho	Universidad Autónoma de Madrid, España

XI Simposio de Teoría y Aplicaciones de Minería de Datos (TAMIDA)

Alicia Troncoso	Universidad Pablo de Olavide, España
José Riquelme	Universidad de Sevilla, España
José Alejandro Fernández Cuesta	Universidad Complutense de Madrid, España

I Conferencia Española sobre Sistemas de Recomendación (SISREC)

Jesús Bobadilla	Universidad Politécnica de Madrid, España
Antonio Moreno	Universidad de Rovira i Virgili, España
Antonio Bahamonde	Universidad de Oviedo, España
Raciel Yera Toledo	Universidad de Jaén, España

Organizadores de Talleres

III Taller de Grupos de investigación españoles de IA en Biomedicina (IABiomed)

Alejandro Rodríguez González	Universidad Politécnica de Madrid, España
José Alberto Benítez Andrades	Universidad de León, España
María del Mar Marcos López	Universidad Jaume I, España

I Taller de Inteligencia Artificial para Educación (TIAE)

Francisco Bellas Bouza	Universidade da Coruña, España
Sara Guerreiro Santalla	Universidade da Coruña, España
Óscar Fontenla Romero	Universidade da Coruña, España
Noelia Sánchez Maroño	Universidade da Coruña, España

Comité de organización

Publicidad

Laura Morán	Universidade da Coruña, España

Web

Jorge Paz	Universidade da Coruña, España
Samuel Suárez	Universidade da Coruña, España

Comités de programa

Sesión general de CAEPIA

Amparo Alonso-Betanzos	Universidade da Coruña
Lourdes Araujo	Universidad Nacional de Educación a Distancia
Jose Maria Armingol Moreno	Universidad Carlos III de Madrid
Ivan Armuelles	Universidad de Panamá
Jaume Bacardit	Newcastle University
Antonio Bahamonde	Universidad de Oviedo
Alvaro Barreiro	Universidade da Coruña
Edurne Barrenechea	Universidad Pública de Navarra
Senén Barro	Universidade de Santiago de Compostela
Dena Bazazian	University of Bristol
Ana M. Bernardos	Universidad Politécnica de Madrid
Concha Bielza Lozoya	Universidad Politécnica de Madrid
Daniel Borrajo	Universidad Carlos III de Madrid
Juan Botia	King's College London
Alberto Bugarín	Universidade de Santiago de Compostela
Humberto Bustince	Universidad Pública de Navarra
Pedro Cabalar	Universidade da Coruña
José M. Cadenas	Universidad de Murcia
Iván Cantador	Universidad Autónoma de Madrid
Javier Carbó	Universidad Carlos III de Madrid
Manuel Carranza García	Universidad de Sevilla
Pedro A. Castillo	Universidad de Granada
Francisco Chicano	Universidad de Málaga
Christian Cintrano	Universidad de Málaga
Juan Manuel Corchado	Universidad de Salamanca
Rafael Corchuelo	Universidad de Sevilla
Oscar Cordon	Universidad de Granada
David Camilo Corrales Muñoz	Institut national de recherche pour l'agriculture l'alimentation et l'environnement
Zakaria Abdelmoiz Dahi	Universite Abdelhamid Mehri Constantine 2
Sergio Damas	Universidad de Granada
Andre de Carvalho	Universidade of São Paulo
Luis De La Ossa	Universidad de Castilla-La Mancha
Juan José Del Coz	Universidad de Oviedo
María José Del Jesús	Universidad de Jaén
Javier Del Ser Lorente	Tecnalia Research & Innovation
Francisco Javier Díez	Universidad Nacional de Educación a Distancia
Jose Dorronsoro	Universidad Autónoma de Madrid
Javier Echanobe	Euskal Herriko Unibertsitatea
Jorge Fandinno	University of Nebraska Omaha
Javier Faulín	Universidad Pública de Navarra
Andrea Fenández Martinez	AIMEN
Mariano Fernández López	Universidad CEU San Pablo
Juan Manuel Fernandez Luna	Universidad de Granada
Jesualdo Tomás Fernández-Breis	Universidad de Murcia
Antonio Fernández-Caballero	Universidad de Castilla-La Mancha
Javier Ferrer	Universidad de Málaga
Francesc J. Ferri	Universitat de València
José Manuel Galán	Universidad de Burgos
Jose Gamez	Universidad de Castilla-La Mancha
Pablo García Bringas	Universidad de Deusto
Mikel Garcia de Andoin	Euskal Herriko Unibertsitatea
Jesús García Herrero	Universidad Carlos III de Madrid

Raúl García-Castro	Universidad Politécnica de Madrid
Rodrigo Gil-Merino	Universidad de Málaga
Juan Gomez Romero	Universidad de Granada
Carlos Gómez-Rodríguez	Universidade da Coruña
Antonio Gonzalez	Universidad de Granada
Carlos González Val	AIMEN
Manuel Graña	Euskal Herriko Unibertsitatea
Bertha Guijarro-Berdiñas	Universidade da Coruña
Cesar Hervas	Universidad de Córdoba
José Antonio Iglesias Martínez	Universidad Carlos III de Madrid
Inaki Inza	Euskal Herriko Unibertsitatea
Luis Jimenez Linares	Universidad de Castilla-La Mancha
Salud María Jiménez-Zafra	Universidad de Jaén
Kaisar Kushibar	Universitat de Barcelona
Pedro Lara-Benítez	Universidad de Sevilla
Pedro Larranaga	Universidad Politécnica de Madrid
Agapito Ledezma Espino	Universidad Carlos III de Madrid
Juan Pedro Llerena	Universidad Carlos III de Madrid
Adolfo Lopez	Universidad de Valladolid
Victoria López	Colegio Universitario de Estudios Financieros
Beatriz López	Universitat de Girona
Pilar López-Úbeda	Universidad de Jaén
Ascension Lopez-Vargas	Universidad Carlos III de Madrid
Gabriel Luque	Universidad de Málaga
Lawrence Mandow	Universidad de Málaga
Felip Manyà	Instituto de Investigación en Inteligencia Artificial -CSIC
Mar Marcos	Universitat Jaume I
Pablo Marin	Universidad Carlos III de Madrid
Patricio Martínez Barco	Universidad de Alicante
Rafael Martinez Tomas	Universidad Nacional de Educación a Distancia
Eugenio Martinez-Camara	Universidad de Granada
Jose Massa	Universidad Nacional del Centro de la Provincia de Buenos Aires
José María Massa	-
Renzo Massobrio	University of Antwerp
Rafael Medina-Carnicer	Universidad de Córdoba
Marcos Mejia	Universitat de Barcelona
Belen Melian	Universidad de La Laguna
Pedro Meseguer	Instituto de Investigación en Inteligencia Artificial CSIC
Eva Millan	Universidad de Málaga
Salar Mohtaj	-
Jose M. Molina	Universidad Carlos III de Madrid
Miguel Molina Solana	Universidad de Granada
Serafín Moral Callejón	Universidad de Granada
José Angel Morell	Universidad de Málaga
Javier Muguerza	Euskal Herriko Unibertsitatea
Antonio Muñoz	Universidad de Málaga
Bhalaji Nagarajan	Universitat de Barcelona
Ismael Navas	Universidad de Málaga
Sergio Nesmachnow	Universidad de la República, Uruguay
Jose Fernando Nuñez	Universitat de Barcelona
Manuel Ojeda-Aciego	Universidad de Málaga
Jose Angel Olivas	Universidad de Castilla-La Mancha
Eva Onaindia	Universitat Politècnica de València
Sascha Ossowski	Universidad Rey Juan Carlos
Jose Palma	Universidad de Murcia
Daniel Pandolfi,	Universidad Nacional de la Patagonia Austral
C. Alejandro Parraga	Universitat Autònoma de Barcelona
Miguel Angel Patricio	Universidad Carlos III de Madrid
Juan Pavón Mestras	Universidad Complutense de Madrid
Diego Pedroza	Universidad de Málaga
Antonio Peregrin	Universidad de Huelva

Eduardo Perez	Instituto Maimónides de Investigación Biomédica de Córdoba
José Luis Pérez de la Cruz Molina	Universidad de Malaga
Giuseppe Pezzano	Universitat de Barcelona
Hector Pomares	Universidad de Granada
Jose M Puerta	Universidad de Castilla-La Mancha
Camino R. Vela	Universidad de Oviedo
Noelia Rico	Universidad de Oviedo
Ramon Rizo	Universidad de Alicante
Javier Rodenas	Universitat de Barcelona
Alejandro Rodriguez	Universidad Politécnica de Madrid
Rosa Rodríguez	Universidad de Jaén
Diego Gabriel Rossit	Universidad Nacional del Sur - CONICET
Elias Said Hung	Universidad Internacional de la Rioja
Antonio Salmeron	Universidad de Almería
Luciano Sanchez	Universidad de Oviedo
Jose Salvador Sanchez	Universitat Jaume I
Araceli Sanchis	Universidad Carlos III de Madrid
Estela Saquete	-
Encarna Segarra	Universitat Politècnica de València
M. Paz Sesmero Lorente	Universidad Carlos III de Madrid
Igor Škrjanc	Univerza v Ljubljani
Emilio Soria	Universitat de València
Maria Taboada	Universidade de Santiago de Compostela
Estefania Talavera	University of Groningen
Erik Torrontegui	Universidad Carlos III de Madrid
Alicia Troncoso	Universidad Pablo de Olavide
L. Alfonso Ureña-López	Universidad de Jaén
Rafael Valencia-Garcia	Universidad de Murcia
Alfredo Vellido	Universitat Politècnica de Catalunya
Sebastián Ventura	Universidad de Córdoba
Andrea Villagra	Universidad Nacional de la Patagonia Austral
José Ramón Villar	Universidad de Oviedo
Guobiao Zhang	-

ESTYLF

Jesús Alcalá-Fernández	Universidad de Granada
Cristina Alcalde	Universidad del País Vasco UPV/EHU
Sergio Alonso	Universidad de Granada
Jose M. Alonso	Universidad de Santiago de Compostela
Roberto G. Aragón	Universidad de Cádiz
Edurne Barrenechea	Universidad Pública de Navarra
Senén Barro	Universidad de Santiago de Compostela
María José Benítez-Caballero	Universidad de Cádiz
Fernando Bobillo	Universidad de Zaragoza
Alberto Bugarín-Diz	Universidade de Santiago de Compostela
Ana Burusco	Universidad Pública de Navarra
Humberto Bustince	Universidad Pública de Navarra
Inmaculada P. Cabrera	Universidad de Málaga
Francisco Javier Cabrerizo	Universidad de Granada
José M. Cadenas	Universidad de Murcia
Tomasa Calvo	Universidad de Alcalá
Pablo Carmona	Universidad de Extremadura
Juan Luis Castro	Universidad de Granada
Pablo Cordero	Universidad de Málaga
Óscar Cordón	Universidad de Granada
Maria Eugenia Cornejo Piñero	Universidad de Cádiz
Susana Cubillo	Universidad Politécnica de Madrid
Rocio De Andres	Universidad de Salamanca
Miguel Delgado	Universidad de Granada
Susana Díaz	Universidad de Oviedo

Jorge Elorza	Universidad de Navarra
Javier Fernández	Universidad Pública de Navarra
Mikel Galar	Universidad Pública de Navarra
José Luis García-Lapresta	Universidad de Valladolid
Lluis Godo	Instituto de Investigación en Inteligencia Artificial, IIIA - CSIC
Antonio González	Universidad de Granada
Francisco Herrera Triguero	Universidad de Granada
Enrique Herrera Viedma	Universidad de Granada
Luis Jiménez Linares	Universidad de Castilla-La Mancha
Maria Teresa Lamata	Universidad de Granada
David Lobo	Universidad de Cádiz
Bonifacio Llamazares	Universidad de Valladolid
Carlos López-Molina	Universidad Pública de Navarra
Nicolás Madrid	Universidad de Cádiz
Luis Magdalena	Universidad Politécnica de Madrid
María J. Martín-Bautista	Universidad de Granada
Luis Martínez	Universidad de Jaén
Sebastià Massanet	Universidad de the Balearic Islands
Francisco Mata	Universidad de Jaén
Jesús Medina	Universidad de Cádiz
Jose M. Molina	Universidad Carlos III de Madrid
Javier Montero	Universidad Complutense de Madrid
Susana Montes	Universidad de Oviedo
Juan Moreno García	Universidad de Castilla-La Mancha
Manuel Ojeda-Aciego	Universidad de Málaga
Jose Angel Olivas	Universidad de Castilla-La Mancha
Antonio Peregrín	Universidad de Huelva
Héctor Pomares	Universidad de Granada
Ana Pradera	Universidad Rey Juan Carlos
Eloísa Ramírez-Poussa	Universidad de Cádiz
Jordi Recasens	Universidad Politécnica de Cataluña
Juan Vicente Riera	Universidad de las Islas Baleares
Rosa Mª Rodríguez Domínguez	Universidad de Jaén
Francisco P. Romero	Universidad de Castilla-La Mancha
Luciano Sánchez	Universidad de Oviedo
Daniel Sánchez	Universidad de Granada
José Antonio Sanz Delgado	Universidad Pública de Navarra
Jesús Serrano-Guerrero	Universidad de Castilla-La Mancha
Miguel-Angel Sicilia	Universidad de Alcalá
Vicenc Torra	Universidad de Umea
Aida Valls	Universidad Rovira i Virgili
José Luis Verdegay	Universidad de Granada

MAEB

Ada Alvarez	Universidad Autónoma de Nuevo León
Ramón Álvarez-Valdés	Universitat de València
Jhon Edgar Amaya	Universidad Nacional Experimental del Táchira
Lourdes Araujo	Universidad Nacional de Educación a Distancia
Joaquín Bautista	Universitat Politècnica de Catalunya
Christian Blum	Consejo Superior de Investigaciones Científicas (CSIC)
Julio Brito	Universidad de La Laguna
Rafael Caballero	Universidad de Málaga
Carlos Camacho	Universidad Politécnica de Madrid
David Camacho	Universidad Politécnica de Madrid
Silvia Casado	Universidad de Burgos
Pedro Castillo	Universidad de Granada
Sergio Cavero Díaz	Universidad Rey Juan Carlos
Francisco Chávez	Universidad de Extremadura
Manuel Chica	Universidad de Granada
Carlos Coello Coello	Centro de Investigación y de Estudios Avanzados del

José Manuel Colmenar	Universidad Rey Juan Carlos
Feijoo Colomine D.	Universidad Nacional Experimental del Táchira
Juan Carlos Cortés-López	Universitat Politècnica de València
Carlos Cotta	Universidad de Málaga
Gretel De la Peña	-
Bernabe Dorronsoro	Universidad de Cádiz
Richard Duro	Universidade da Coruña
Jose A. Egea	Centro de Edafología y Biología Aplicada del Segura - CSIC
Javier Faulin	Universidad Pública de Navarra
Alberto Fernández	Universidad de Granada
Francisco Fernandez De Vega	Universidad de Extremadura
Antonio J. Fernández Leiva	Universidad de Málaga
Jose Gamez	Universidad de Castilla-La Mancha
Pablo García	Universidad de Granada
Carlos García-Martínez	Universidad de Córdoba
Nicolás García-Pedrajas	Universidad de Córdoba
Oscar Garnica	Universidad Complutense de Madrid
Sergio Gil Borrás	Universidad Politécnica de Madrid
Juan Gomez Romero	Universidad de Granada
Pedro Gonzalez	Universidad de Jaén
Antonio Gonzalez	Universidad de Granada
Antonio González-Pardo	Universidad Rey Juan Carlos
Pedro Antonio Gutierrez	Universidad de Córdoba
Alfredo G. Hernández-Díaz	Universidad Pablo de Olavide
Cesar Hervas	Universidad de Córdoba
Iñaki Hidalgo	Universidad Complutense de Madrid
Angel Juan	Universitat Oberta de Catalunya
Manuel Laguna	University of Colorado Boulder
Juan Lanchares	Universidad Complutense de Madrid
Isaac Lozano-Osorio	Universidad Rey Juan Carlos
Jose Maria Luna	Universidad de Córdoba
Francisco Luna	Universidad de Málaga
Mariano Luque	Universidad de Málaga
Gabriel Luque	Universidad de Málaga
Rafael Luque-Baena	Universidad de Extremadura
Luis Magdalena	Universidad Politécnica de Madrid
Raúl Martín Santamaría	Universidad Rey Juan Carlos
Alexander Mendiburu	Euskal Herriko Unibertsitatea
Pablo Mesejo	Universidad de Granada
Miguel Molina-Solana	Universidad de Granada
José-A. Moreno	Universidad de La Laguna
Antonio J. Nebro	Universidad de Málaga
Joaquín Pacheco	Universidad de Burgos
Angel Panizo Lledot	Universidad Politécnica de Madrid
Juanjo Peiró	Universitat de València
David Pelta	Universidad de Granada
Antonio Peregrin	Universidad de Huelva
Raul Perez	Universidad de Granada
Sergio Pérez-Peló	Universidad Rey Juan Carlos
Roger Z. Ríos-Mercado	Universidad Autónoma de Nuevo León
José C. Riquelme	Universidad de Sevilla
José L. Risco-Martín	Universidad Complutense de Madrid
Victor Manuel Rivas Santos	Universidad de Jaén
Juan Francisco Robles	Universidad de Granada
David Rodríguez	Universidad Nacional Experimental del Táchira
Alejandro Rosete	Instituto Superior Politécnico "José Antonio Echevarría" - CUJAE
Ruben Ruiz	Universitat Politècnica de València
Yago Sáez	Universidad Carlos III de Madrid
Luciano Sanchez	Universidad de Oviedo
Jesús Sánchez-Oro	Universidad Rey Juan Carlos
Thomas Stützle	Université Libre de Bruxelles

Leonardo Trujillo	Instituto Tecnológico de Tijuana
Angel Udias	Universidad Rey Juan Carlos
Miguel A. Vega-Rodríguez	Universidad de Extremadura
Jose Manuel Velasco	Universidad Complutense de Madrid
Sebastián Ventura	Universidad de Córdoba
José Luis Verdegay	Universidad de Granada
Rafael Villanueva	Instituto Universitario de Matemática Multidisciplinar - UPV
Pedro Villar	Universidad de Granada
Juan Villegas-Cortez	Universidad Autónoma Metropolitana
Gabriel Winter	Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería - ULPGC
Javier Yuste	Universidad Rey Juan Carlos
Amelia Zafra Gómez	Universidad de Córdoba

TAMIDA

Alicia Troncoso Lora	Universidad Pablo de Olavide
José C. Riquelme	Universidad de Sevilla
Salvador García López	Universidad de Granada
José del Campo	Universidad de Málaga
Emilio Corchado	Universidad de Salamanca
Francesc J. Ferri	Universidad de Valencia
Mª José Ramírez	Universidad Politécnica de Valencia
Francisco Herrera	Universidad de Granada
César Hervás	Universidad de Córdoba
Pedro González	Universidad de Jaén
José A. Lozano	Universidad del País Vasco
Juan J. Rodríguez	Universidad de Burgos
José Hernández Orallo	Universidad Politécnica de Valencia
Juan J. del Coz	Universidad de Oviedo
Karina Gibert	Universidad Politécnica de Cataluña
Alberto Cano	Virginia Commonwealth University
Miquel Sànchez i Marrè	Universidad Politécnica de Cataluña
Mª José del Jesús	Universidad de Jaén
Nicolás García Pedrajas	Universidad de Córdoba
Pedro A Gutierrez	Universidad de Córdoba
Olatz Arbelaitz	Universidad del País Vasco
Antonio Bahamonde	Universidad de Oviedo
Jorge Díez	Universidad de Oviedo
Óscar Cordón	Universidad de Granada
Jorge Casillas	Universidad de Granada
María Pérez-Ortíz	University College London
Pedro Larrañaga	Universidad Politécnica de Madrid
Sebastián Ventura	Universidad de Córdoba
Francisco Martínez Álvarez	Universidad Pablo de Olavide
María Martínez Ballesteros	Universidad de Sevilla
Verónica Bolón	Universidade da Coruña
José A. Gámez	Universidad de Castilla-La Mancha
Pablo Casas Gómez	Universidad Pablo de Olavide
Ángela Troncoso García	Universidad Pablo de Olavide
Manuel J. Jiménez Navarro	Universidad de Sevilla

SISREC

S. Alonso	Universidad de Granada
Y. Atrio	Universidad de Santiago de Compostela
M. Barranco	Universidad de Jaén
A. Bellogín	Universidad Autónoma de Madrid
F.J. Cabrerizo	Universidad de Granada
I. Cantador	Universidad Autónoma de Madrid
M. Caro-Martínez	Universidad Complutense de Madrid

P. Castells	Universidad Autónoma de Madrid
M.J. Cobo	Universidad de Granada
B. Díaz-Agudo	Universidad Complutense de Madrid
J. Díez	Universidad de Oviedo
F. Díez	Universidad Autónoma de Madrid
L M. de Campos	Universidad de Granada
B. Dutta	Universidad de Jaén
J.M. Fernández	Universidad de Granada
D. García	Universidad de Jaén
P.A. González	Universidad Complutense de Madrid
S. Heras	Universidad Politécnica de Valencia
J. Herce	Universidad de Granada
A. Hernández Carnerero	Universidad Politécnica de Cataluña
E. Herrera	Universidad de Granada
J. Huete	Universidad de Granada
N. Jakubiak	Universidad Politécnica de Cataluña
G. Jimenez-Díaz	Universidad Complutense de Madrid
A. Labella	Universidad de Jaén
O. Luaces	Universidad de Oviedo
J.M. Morales	Universidad de Granada
I. Neira	Universidad de Santiago de Compostela
J.A. Olivas Varela	Universidad de Castilla La Mancha
E. Onaindía	Universidad Politécnica de Valencia
F. Ortega	Universidad Politécnica de Madrid
P. Pérez	Universidad de Oviedo
F. Pascual Romero Chicharro	Universidad de Castilla La Mancha
E. Peis	Universidad de Granada
C. Porcel Gallego	Universidad de Granada
D. Ranchal	Universidad de Granada
J.A. Recio-García	Universidad Complutense de Madrid
B. Remeseiro	Universidad de Oviedo
R.M. Rodríguez	Universidad de Jaén
M. Sánchez-Marré	Universidad Politécnica de Cataluña
E.M. Sánchez Vila	Universidad de Santiago de Compostela
P. Saavedra	Universidad de Santiago de Compostela
M. Salamó	Universidad de Barcelona
L. Sebastiá	Universidad Politécnica de Valencia
J. Serrano-Guerrero	Universidad de Castilla La Mancha
S. Valero	Universidad Politécnica de Valencia
A. Valls	Universidad Rovira i Virgili

Taller IABiomed

Alfredo Vellido	Universitat Politècnica de Catalunya
Ernestina Menasalvas	Universidad Politécnica de Madrid
Sebastian Ventura	Universidad de Córdoba
Isabel Segura	Universidad Carlos III de Madrid
Francisco Javier Díez	Universidad Nacional de Educación a Distancia
Jesualdo Tomás Fernández-Breis	Universidad de Murcia
Berta Guijarro	Universidade da Coruña
Jose Manuel Juarez	Universidad de Murcia
Javier Muguerza	Universidad del País Vasco
Jose Palma	Universidad de Murcia
José Ramón Villar	Universidad de Oviedo
Victor Manuel González	Universidad de Oviedo
Beatriz Lopez	Universitat de Girona
Joaquin Dopazo	Fundacion Progreso y Salud
Antonio Lopez-Martinez Carrasco	Universidad de Murcia
Begoña Martinez Salvador	Universitat Jaume I
Maria Teresa Garcia	Universidad de Leon
Caroline König	Universitat Politècnica de Catalunya

Isaias Garcia	Universidad de León
Lucia Prieto Santamaria	Universidad Politécnica de Madrid
Jose Aveleira Mata	Universidad de León

Taller TIAE

Daniel González Peña	Universidad de Vigo
Miguel Reboiro Jato	Universidad de Vigo
Félix de La Paz	UNED
Francisco J. Rodríguez Lera	Universidad de León
Alejandro Paz López	Universidade da Coruña
Alma Mallo Castelo	Universidade da Coruña
Beatriz Pérez Sánchez	Universidade da Coruña
Laura Morán Fernández	Universidade da Coruña

Jurado de Competiciones

Alberto Bugarín	Universidad de Santiago de Compostela
Jose Antonio Gámez	Universidad de Castilla la Mancha
Oscar Luaces	Universidad de Oviedo
Eva Onaindia	Universidad Politécnica de Valencia
Alicia Troncoso	Universidad Pablo de Olavide

Comité evaluador Premio Frances Allen

Óscar Cordón García	Universidad de Granada
Eva Onaindía de la Rivaherrera	Universitat Politècnica de València
Camino Rodríguez Vela	Universidad de Oviedo

Índice general

XX Conferencia de la Asociación Española para la Inteligencia Artificial

Sesión General

Trabajos Destacados

XXI Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)

Sesión General

Trabajos Destacados

XV Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB)

Sesión General

Trabajos Destacados

XI Simposio de Teoría y Aplicaciones de Minería de Datos (TAMIDA)

Sesión General

Trabajos Destacados

I Congreso Español en Sistemas de Recomendación (SISREC)

Sesión General

Trabajos Destacados

III Taller de Grupos de investigación españoles de IA en Biomedicina (IABiomed)

Sesión General

Trabajos Destacados

Taller de Inteligencia Artificial para Educación (TIAE)

Sesión General

Premios APP Videos

Doctoral Consortium

Premios Frances Allen

Índice de Autores

XXI Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)

XXI Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)

Sesión General

Álgebras de Heyting débiles: una generalización para retículos no distributivos

1st Francisco Pérez-Gámez
Departamento de Matemática Aplicada
Universidad de Málaga
Málaga, España
franciscoperezgamez@uma.es

2nd Carlos Bejines
Departamento de Matemática Aplicada
Universidad de Málaga
Málaga, España
cbejines@uma.es

***Resumen*—En este artículo se presentan las álgebras de Heyting débiles. Estas álgebras constituyen una extensión del álgebra de Heyting adaptada a retículos no distributivos. Fijado un retículo, se enumeran condiciones que garantizan la existencia de estas álgebras. Además, se caracterizan en función de los operadores de implicación y se acota su rango.**

***Index Terms*—Álgebras de Heyting, Implicación, retículo no distributivo**

I. Introducción

Las álgebras de Heyting son estructuras algebraicas que extienden las álgebras de Boole y tienen una amplia aplicación en distintas áreas como son la lógica, la topología o la ciencia de computación. En este trabajo, estudiamos una nueva variedad de álgebra de Heyting que se puede considerar en algunos retículos no distributivos a las que bautizamos como álgebras de Heyting débiles en [3]. Esta nueva variedad surgió de la necesidad para demostrar la completitud y la correctitud en la lógica de simplificación en el caso de retículos no distributivos. Analizaremos en profundidad estas nuevas estructuras, estudiando sus propiedades más notables y caracterizando su existencia.

II. Resultados iniciales

Las álgebras de Heyting se definen sobre un retículo acotado $(L, \vee, \wedge, 0, 1)$ al que se le adjunta una operación binaria $\to$ que cumple la conocida propiedad de adjunción: para todo $a, b, c \in L$,

$$a \wedge b \leq c \text{ si y solo si } a \leq b \to c.$$

Fijado un retículo acotado $(L, \vee, \wedge, 0, 1)$, para poder definir la operación binaria $\to$ que cumpla la propiedad de adjunción es necesario que el retículo sea distributivo. Diferentes autores han estudiado distintas variedades de las álgebras de Heyting con el objetivo de utilizarlas en estructuras más amplias. Por ejemplo, las semi álgebras de Heyting que fueron estudiadas por H.P. Sankappanavar en [5] o la noción de las álgebras débilmente de Heyting definidas por S. Celani y R. Jansana en [2]. Sin embargo, ambos casos, requieren que el retículo con el que se trabaja sea distributivo.

Queremos agradecer la ayuda económica recibida de los proyectos del Ministerio de Ciencia e Innovación PID2022-140630NB-I00, PID2021-124063OB-I00 y PID2021-127870OB-I00.

En este artículo trabajaremos con retículos completos, por tanto, las álgebras de Heyting con las que trabajaremos serán álgebras de Heyting completas. No obstante, por simplificar el nombre, nos referiremos a ellas como álgebras de Heyting sin mencionar el atributo completa. Un álgebra de Heyting puede ser caracterizado acorde al siguiente resultado.

Proposición 2.1 ([4]): Sea $(L, \vee, \wedge, 0, 1)$ un retículo y $\to: L \times L \to L$ una operación. Entonces, $(L, \vee, \wedge, \to, 0, 1)$ es un álgebra de Heyting si y solo si $(L, \vee, \wedge, 0, 1)$ es un retículo completo y se cumplen las siguientes propiedades:

[H1] $a \wedge (a \to b) = a \wedge b$ para todo $a, b \in L$.
[H2] $a \wedge (b \to c) = a \wedge \Big((a \wedge b) \to (a \wedge c)\Big)$ para todo $a, b, c \in L$.
[H3] $(a \wedge b) \to a = 1$ para todo $a, b \in L$.

La noción de álgebra de Heyting débil se presentó [3] de manera dualizada. Aquí presentamos la definición sin dualizar.

Definición 2.2: Decimos que $(L, \vee, \wedge, \to, 0, 1)$ es un álgebra de Heyting débil si $(L, \vee, \wedge, 0, 1)$ es un retículo completo y se cumplen las siguientes propiedades:

[wH1] $a \wedge b \neq 0$ implica que $a \leq b \to (a \wedge b)$ para todo $a, b \in L$.
[wH2] $a \to b \geq b$ para todo $a, b \in L$.
[wH3] $a \to b = 1$ si y solo si $a \leq b$ para todo $a, b \in L$.
[wH4] $a \wedge (a \to b) = a \wedge b$ para todo $a, b \in L$.

Como primer resultado, tenemos que la definición de álgebra de Heyting débil extiende a la de álgebra de Heyting.

Teorema 2.3: Si $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting, entonces $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil.

Demostración. Sea $(L, \wedge, \vee, \to, 0, 1)$ un álgebra de Heyting. Vamos a probar que cumple las propiedades de álgebra de Heyting débil.

1. Para empezar, [wH1] se cumple ya que si $a, b \in L$ cumplen que $a \wedge b \neq 0$, entonces, por [H2] sabemos que:

$$a \wedge \Big(b \to (a \wedge b)\Big) = a \wedge \Big((a \wedge b) \to (a \wedge a \wedge b)\Big),$$

$$a \wedge \Big((a \wedge b) \to (a \wedge a \wedge b)\Big) = a \wedge \Big((a \wedge b) \to (a \wedge b)\Big).$$

Por [H3] sabemos que $(a \wedge b) \to (a \wedge b) = 1$, por tanto,

$$a \wedge \Big(b \to (a \wedge b)\Big) = a \wedge 1 = a,$$

es decir, $a \leq b \to (a \wedge b)$.

2. Para probar [wH2], tomamos $a, b \in L$. Por [H2] sabemos que

$$b \wedge (a \to b) = b \wedge \Big((a \wedge b) \to (a \wedge a)\Big) = b \wedge \Big((a \wedge b) \to a\Big).$$

y por [H3], $(a \wedge b) \to a = 1$. En consecuencia,

$$b \wedge (a \to b) = b \wedge 1 = b,$$

Es decir, $b \leq a \to b$.

3. Si $a \leq b$,por [H3], $a \to b = 1$. Por otro lado, supongamos que $a \to b = 1$, tenemos que:

$$a \wedge (a \to b) = a \wedge 1 = a.$$

Por otro lado, por [H1], tenemos que

$$a \wedge (a \to b) = a \wedge b.$$

Es decir, $a = a \wedge b$ y, por tanto, $a \leq b$.

4. wH4 es igual a [H1].

□

III. Caracterización de las álgebras de Heyting débiles en términos del operador implicación

Fijado un retículo $(L, \wedge, \vee, 0, 1)$, el propósito de esta sección es determinar una condición necesaria y suficiente sobre un operador implicación $(\to)$ para construir un álgebra de Heyting débil $(L, \wedge, \vee, \to, 0, 1)$.

Para ello, necesitamos introducir una condición sobre $(L, \wedge, \vee, 0, 1)$ que es clave en la construcción de un álgebra de Heyting débil.

Definición 3.1: Sea $(L, \wedge, \vee, 0, 1)$ un retículo completo. Decimos que L es cerrado bajo supremos-$\wedge$ si para todo $a, b \in L$ cumpliendo $a \geq b \neq 0$, se satisface la siguiente igualdad

$$a \wedge (\sup_{c \in L}\{a \wedge c = b\}) = b, \tag{1}$$

es decir, el supremo del conjunto es un máximo.

Ejemplo 3.2: Consideremos los retículos L y L' en la Figura 1 y en la Figura 2.

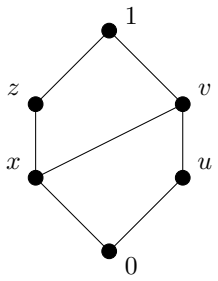


Figura 1. Retículo L.

veamos que L es cerrado bajo supremos-$\wedge$. Sean $a, b \in L$ con

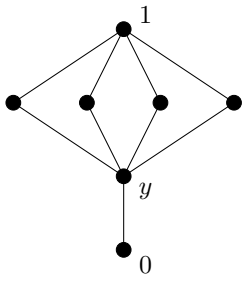


Figura 2. Retículo L'.

$a \geq b \neq 0$. Si $x = 1$ o $a = b$, la igualdad (1) es trivial. Por tanto, solo necesitamos comprobar los casos $b \in \{x, u\}$.

- Supongamos que $b = x$. Como $a > b$, sabemos que $a \in \{z, u, 1\}$. Si $a = 1$ es trivial. Si $a = z$,

$$z \wedge (\sup_{t \in L}\{z \wedge t = x\}) = z \wedge (\sup\{x, v\}) = z \wedge v = x.$$

Si $x = v$,

$$v \wedge (\sup_{t \in L}\{v \wedge t = x\}) = v \wedge (\sup\{x, z\}) = v \wedge z = x.$$

- Ahora supongamos que $b = u$. De $a > b$, sabemos que $a \in \{v, 1\}$. Si $a = 1$ el resultado es trivial. En el caso que $a = v$,

$$v \wedge (\sup_{t \in L}\{v \wedge t = u\}) = v \wedge (\sup\{u\}) = v \wedge u = u.$$

Veamos ahora que L' no es cerrado bajo supremos-$\wedge$. Este retículo pertenece a una familia de retículos conocida, estudiada en profundidad en [1]. Sea $a \in L'$ tal que $y < a < 1$. Tenemos que

$$a \wedge (\sup_{t \in L}\{a \wedge t = y\}) = a \wedge 1 = a \neq y.$$

De acuerdo con el siguiente resultado, dado un retículo completo $(L, \wedge, \vee, 0, 1)$, podemos construir un álgebra de Heyting débil sobre él siempre que el retículo sea cerrado bajo supremos-$\wedge$.

Proposición 3.3: Sea $(L, \wedge, \vee, 0, 1)$ un retículo completo tal que L es cerrado bajo supremos-$\wedge$. Consideremos el operador $\to$: $L \times L \to L$ definido por:

$$a \to b = \begin{cases} 1 & \text{si } a = b = 0, \\ 0 & \text{si } a \neq b = 0, \\ \sup\limits_{t \in L}\{a \wedge t = a \wedge b\} & \text{si } b \neq 0 \text{ con } a \wedge b \in \{a, b\}, \\ b & \text{en otro caso.} \end{cases}$$

Entonces, $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil.

Demostración. Simplemente tenemos que probar que $(L, \wedge, \vee, \to, 0, 1)$ cumple las propiedades de álgebra de Heyting débil.

1. Observemos que [wH1] se cumple ya que si $a, b \in L$ cumplen que $a \wedge b \neq 0$ entonces b es comparable con $(a \wedge b)$, por tanto,

$$b \to (a \wedge b) = \sup_{t \in L}\{b \wedge t = b \wedge (a \wedge b)\},$$

es decir, $a \leq b \to (a \wedge b)$.

2. De forma directa, por la definición de $\to$, tenemos que [wH2] se cumple.
3. Por definición, $0 \to 0 = 1$. Si $a \leq b \neq 0$, entonces

$$a \to b = \sup_{t \in L}\{a \wedge t = a \wedge b\} = \sup_{t \in L}\{a \wedge t = a\} = 1.$$

4. Dados $a, b \in L$, tenemos que comprobar que $a \wedge (a \to b) = a \wedge b$. Si $a = 0$ o $b = 0$, entonces la igualdad es trivial. Supongamos ahora que a es comparable con b, entonces:

$$a \wedge (a \to b) = a \wedge (\sup_{t \in L}\{a \wedge t = a \wedge b\}).$$

Si $0 \neq a < b$, la igualdad es directa. Si $a \geq b \neq 0$, tenemos que $a \wedge b = b$ y como L es cerrado bajo supremos-$\wedge$,

$$a \wedge (\sup_{t \in L}\{a \wedge t = a \wedge b\}) = a \wedge b.$$

Si a no es comparable con b, entonces la igualdad es trivial.

□

La Proposición 3.3 establece una condición suficiente sobre el retículo completo para poder construir un álgebra de Heyting débil. Como probamos en el siguiente teorema, la condición es también necesaria.

Teorema 3.4: Sea $(L, \wedge, \vee, 0, 1)$ un retículo completo. Existe una operación $\to$: $L \times L \to L$ tal que $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil si y solo si L es cerrado bajo supremos-$\wedge$.

Demostración. Si L es cerrado bajo supremos-$\wedge$, considerando la función $\to$ de la proposición 3.3, tenemos que $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil. Por otro lado, asumamos que L no es cerrado bajo surempos-$\wedge$, es decir, existen $a, b \in L$ tales que $a \geq b \neq 0$ cumpliendo que:

$$a \wedge (\sup_{t \in L}\{a \wedge t = b\}) \neq b. \tag{2}$$

Por reducción al absurdo, supongamos que existe una función $\to$ tal que $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil. Por la propiedad [wH4] tenemos que

$$a \wedge (a \to b) = a \wedge b = b,$$

por tanto, $(a \to b) \in \{t \in L \mid a \wedge t = b\}$. Sea $x \in \{t \in L \mid a \wedge t = b\}$, como $x \wedge a = b \neq 0$, por [wH1], obtenemos que $x \leq a \to (x \wedge a)$, que es equivalente a

$$x \leq (a \to b).$$

Es decir,

$$\sup_{t \in L}\{a \wedge t = b\} = a \to b,$$

que contradice la Condición 2. □

Nótese la importancia de la Proposición 3.3, pues no solo nos garantiza la existencia, sino que nos dice cómo deben ser los operadores implicación para su construcción. El Teorema 3.4 nos asegura que esa construcción captura todos los operadores implicaciones que pueden formar un álgebra de Heyting débil. Equivalentemente, cada uno de los operadores de implicación descritos en la proposición mencionada induce, y solo ellos, un álgebra de Heyting débil. Por ello, para acabar esta sección, daremos una caracterización de ellas en términos de su operación implicación.

Teorema 3.5: Sea $(L, \wedge, \vee, 0, 1)$ un retículo completo y un operador $\to$: $L \times L \longrightarrow L$. Entonces, $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil si y solo si el operador $\to$ cumple las siguientes condiciones:

$I)$ $0 \to 0 = 1$.

$II)$ Si $a \neq 0$, entonces $a \to 0 \in \{t \in L \mid a \wedge t = 0\}$.

$III)$ Si $b \neq 0$, además de $a \wedge b \in \{a, b\}$, entonces

$$a \to b = \sup\{t \in L \mid a \wedge t = a \wedge b\}.$$

$IV)$ Si $a \wedge b \notin \{a, b\}$, entonces

$$a \to b \in \{t \in L \mid a \wedge t = a \wedge b \text{ con } b \leq t\}.$$

Demostración. Supongamos que $(L, \wedge, \vee, \to, 0, 1)$ es un álgebra de Heyting débil. Entonces:

$I)$ es directo por [wH3].

$II)$ De [wH4] tenemos que

$$a \wedge (a \to 0) = a \wedge 0 = 0.$$

Por tanto, si $a \neq 0$, tenemos que necesariamente $a \to 0 \in \{t \in L \mid a \wedge t = 0\}$.

$c)$ Si $b \neq 0$ además de $a \wedge b \in \{a, b\}$ distinguiremos dos casos:
En el primer caso supondremos que $a \wedge b = a$. Por [wH3] tenemos que

$$a \to b = 1 = \sup\{t \in L \mid a \wedge t = a \wedge b\}.$$

En el segundo caso supondremos que $a \wedge b = b$. Por [wH4],

$$a \wedge (a \to b) = a \wedge b = b.$$

Por tanto, $a \to b \in \{t \in L \mid a \wedge t = a \wedge b\}$. Comprobemos que $a \to b$ es el máximo, tomamos $t \in L$ tal que $a \wedge t = b$. Como $t \wedge a = b \neq 0$, por [wH1], tenemos que

$$t \leq a \to (t \wedge a) = a \to b,$$

es decir, $a \to b = \sup\{t \in L \mid a \wedge t = a \wedge b\}$.

$d)$ Si $a \wedge b \notin \{a, b\}$, por [wH4] se cumple que $a \wedge (a \to b) = a \wedge b$, que implica que

$$a \to b \in \{t \in L \mid a \wedge t = a \wedge b\}.$$

Además, por [wH2], sabemos que $b \leq a \to b$.

Por otro lado, asumamos que $(L, \wedge, \vee, 0, 1)$ es un retículo completo y que la función $\to$: $L \times L \to L$ cumpla las propiedades $I), II), III)$ y $IV)$. Consideremos $a, b \in L$.

1. Si $a \wedge b \neq 0$, entonces tenemos que $a \neq 0$ y $b \neq 0$. Podemos comprobar que [wH1] se cumple, es decir, que $a \leq b \to (a \wedge b)$. Observemos que $0 \neq a \wedge b \leq b$, por tanto, podemos aplicar $III)$:

$$b \to (a \wedge b) = \sup\{t \in L \mid b \wedge t = b \wedge (a \wedge b)\}.$$

Es decir, para todo $t \in L$ que cumple $b \wedge t = a \wedge b$, tenemos que $t \leq b \to (a \wedge b)$. Terminando la prueba simplemente tomando el caso $t = a$.

2. Podemos comprobar que [wH2] se cumple comprobando que para todo $a, b \in L$ $a \to b$ viene dado por alguna de las expresiones de $I), II), III)$ y $IV)$ y en todas ellas se cumple que $a \to b \geq b$.
3. Supongamos que $a \leq b$. Si $b = 0$, por $I)$, tenemos que $a \to b = 1$. Si $b \neq 0$, como $a \wedge b = a$, por $III)$ tenemos que

$$a \to b = \sup\{t \in L \mid a \wedge t = a \wedge b = a\}$$

y es fácil comprobar que 1 es el supremo de ese conjunto.
Supongamos ahora que $a \to b = 1$. Si $a \to b$ viene dado por $I)$ entonces es claro que $a \leq b$. Si $a \neq 0$, $a \to 0 \neq 1$ y, por tanto, es imposible que $a \to b$ venga dado por $II)$. Si $a \to b$ viene dado por $III)$, entonces

$$1 = a \to b = \sup\{t \in L \mid a \wedge t = a \wedge b\},$$

concluyendo que $a \wedge 1 = a \wedge b$, que es equivalente a $a \leq b$. Para terminar, vemos que es imposible que $a \to b$ venga dado por $IV)$ por reducción al absurdo. Supongamos que $a \to b$ viene dado por $IV)$. Entonces,

$$1 = a \to b = \sup\{t \in L \mid a \wedge t = a \wedge b \text{ y } b \leq t\},$$

por tanto, $a \wedge 1 = a \wedge b$, pero si hemos aplicado $IV)$ es porque $a \wedge b \notin \{a, b\}$.
4. La demostración de [wH4] es análoga a la demostración de [wH2].

□

IV. Propiedades de las álgebras de Heyting débiles

En caso de existir el álgebra de Heyting, gracias al Teorema 2.3, sabemos que también es álgebra de Heyting débil. Teniendo en cuenta el Teorema 3.5, la siguiente proposición nos muestra que el operador implicación de un álgebra de Heyting débil es siempre igual o menor al operador de un álgebra de Heyting.

Proposición 4.1: Fijado un retículo completo L, supongamos que $(L, \wedge, \vee, \to_H, 0, 1)$ es un álgebra de Heyting y $(L, \wedge, \vee, \to_W, 0, 1)$ es un álgebra de Heyting débil. Entonces la implicación $\to_W$ está acotada superiormente por $\to_H$.

Demostración. Gracias al teorema 3.5, vamos a desglosar la demostración en los siguientes casos disjuntos. Sea a y b elementos del retículo L.

1. Si $a \leq b$, trivialmente

$$a \to_H b = 1 = a \to_W b.$$

2. Si $b = 0$ y $a \neq 0$, a partir del teorema, podemos afirmar que

$$a \to_H 0 = \text{máx}\{t \in L \mid t \wedge a = 0\}$$

y

$$a \to_W 0 \in \{t \in L \mid a \wedge t = 0\}.$$

Se concluye que $a \to_W 0 \leq a \to_H 0$.
3. Si $0 \neq b < a$, tenemos que

$$a \to_H b = \text{máx}\{t \in L \mid t \wedge a \leq b\}$$

y

$$a \to_W b \in \{t \in L \mid a \wedge t = b\}.$$

Teniendo en cuenta que $\{t \in L \mid a \wedge t = b\}$ está contenido en $\{t \in L \mid t \wedge a \leq b\}$, concluimos que $a \to_W b \leq a \to_H b$.
4. Por último, si a no es comparable a b, tenemos que

$$a \to_H b = \text{máx}\{t \in L \mid t \wedge a \leq b\}$$

y

$$a \to_W b \in \{t \in L \mid a \wedge t = a \wedge b \text{ y } b \leq t\}.$$

Como $\{t \in L \mid a \wedge t = a \wedge b \text{ y } b \leq t\}$ es un subconjunto de $\{t \in L \mid t \wedge a \leq b\}$, concluimos que $a \to_W b \leq a \to_H b$.

□

Además de tener el supremo de todas las álgebras de Heyting débiles desde el punto de vista de sus operadores de implicación, el cual es el operador de implicación $\to_H$ asociado al álgebra de Heyting, es posible construir el ínfimo de la siguiente forma.

Proposición 4.2: Sea $(L, \wedge, \vee, 0, 1)$ un retículo completo tal que L es cerrado bajo supremos-$\wedge$. La menor implicación $\rightsquigarrow$ tal que $(L, \wedge, \vee, \rightsquigarrow, 0, 1)$ es un álgebra de Heyting débil viene dada por la siguiente expresión:

$$a \rightsquigarrow b = \begin{cases} 0 & \text{si } b = 0, \\ b & \text{si no son compatibles,} \\ \sup\limits_{t \in L}\{a \wedge t = a \wedge b\} & \text{en otro caso.} \end{cases} \tag{3}$$

La demostración de esta proposición es inmediata usando el Teorema 3.5.

Por lo tanto, fijado un retículo completo $(L, \wedge, \vee, 0, 1)$ cerrado bajo $supremos - \wedge$, cualquier operador $\to_W$ que induzca un álgebra de Heyting débil está acotado inferiormente por $\rightsquigarrow$ y superiormente por $\to_H$, es decir, para todo $a, b \in L$, tenemos que

$$a \rightsquigarrow b \ \leq \ a \to_W b \ \leq \ a \to_H b,$$

V. Conclusiones y trabajos futuros

En [3] se introdujo las álgebras de Heyting débiles como una extensión de las álgebras de Heyting, adaptadas para considerarlas sobre retículos no distributivos. En esa línea, se han establecido condiciones sobre el retículo para garantizar la existencia de un álgebra de Heyting débil y se ha proporcionado una caracterización en términos del operador implicación. Además, hemos presentado un método para construir un operador implicación que forme un álgebra de Heyting débil siempre que el retículo sea cerrado bajo supremos-$\wedge$.

Para futuras investigaciones, se propone determinar el número de implicaciones que pueden formar un álgebra de Heyting débil dado un retículo cerrado bajo supremos-$\wedge$. Asimismo, se plantea investigar las propiedades que deben cumplir estas implicaciones para poder extenderlas punto a punto en el caso de productos de retículos, manteniéndose como álgebras de Heyting débiles. Además, se considera la exploración de aplicaciones de estas álgebras en campos distintos a la lógica de simplificación.

Referencias

[1] Bejines, C., and Bruteničová, M., and Chasco, M.J., and Elorza, J., and Janiš, V., “The number of t-norms on some special lattices”, Fuzzy sets and Systems, 4, vol. 408 pp. 26-43, 2021.

[2] Celani, S., and Jansana, R., A Treatise “ Bounded distributive lattices with strict implication”, Mathematical Logic Quarterly, vol. 51, number 3, pp. 219–246, Wiley Online Library, 2005.

[3] Pérez-Gamez, F., Cordero, P., Enciso, M., and Mora, A., “ Simplification logic for the management of unknown information”, Information Sciences, vol. 634, pp. 505–519, 2023.

[4] Raymond Balbes and Philip Dwinger, “Distributive Lattices”, Journal of Symbolic Logic, 4, vol. 42 pp. 587–588, 1977.

[5] Sankappanavar, Hanamantagouda P, “Semi-Heyting algebras: an abstraction from Heyting algebras”, Actas del IX Congreso Dr. A. Monteiro, pp. 33–66, 2007.

Nilpotent Minimum logic preserving non-falsity and consistency operators

Francesc Esteva
IIIA – CSIC
08193 Bellaterra, Spain
esteva@iiia.csic.es

Joan Gispert
Fac. Matemàtiques – Univ. Barcelona
08007 Barcelona, Spain
jgispertb@ub.edu

Lluís Godo
IIIA – CSIC
08193 Bellaterra, Spain
godo@iiia.csic.es

***Abstract*—Nilpotent Minimum logic (NML) is one of the fuzzy logics of the family of t-norm based logics that is particularly interesting because it enjoys two important properties: its residual negation is involutive and it satisfies a form of deduction theorem. In this paper we study a paraconsistent companion of NML that captures a weak notion of logical consequence that preserves non-zero truth-values from the premises to the conclusions. Moreover, we also consider its expansion with a consistency operator in the sense of the logics of formal inconsistency (LFI).**

***Index Terms*—Fuzzy logic; Nilpotent minimum logic; paraconsistent logic; non-falsity preserving companion; consistency operators.**

I. Introduction

Nilpotent Minimum logic (NML for short) is a distinguished member of the family of formal systems of mathematical fuzzy logic (MFL) [11], introduced by two of the authors of this paper in [6] as a particular extension of the so-called Monoidal t-norm based fuzzy logic (MTL), a very general logic whose equivalent algebraic semantis is the variety of prelinear (commutative, bounded, integral) residuated lattices, also known as MTL-algebras, that is generated by the subclass of algebras with domain the real unit interval $[0, 1]$ and defined by left-continuous t-norms[1], see [13]. In fact, the logic NM was originally defined in [6] as the axiomatic extension of MTL by the involutive negation axiom

$$\text{(INV)} \quad \neg\neg\varphi \to \varphi$$

and the (weak) nilpotent minimum axiom

$$\text{(WNM)} \quad (\psi * \varphi \to \bot) \vee (\psi \wedge \varphi \to \psi * \varphi).$$

NML is an algebraizable logic, as all the axiomatic extensions of MTL, and the corresponding variety of NM-algebras is generated by a single algebra on the real unit interval $[0, 1]$, called *standard NM-algebra*, defined by the Nilpotent minimum t-norm and its residuum, see Section 2.

The NM logic together with all its axiomatic and finitary extensions has been exhaustively studied by Gispert in [8], [9]. They are all explosive with respect to its residual negation $\neg\varphi = \varphi \to 0$. This means that any theory T containing or deriving both a formula φ and its negation $\neg\varphi$ is contradictory, and hence it can derive any formula. In other words, the explosion rule with respect to $\neg$:

$$\text{(Exp)} \quad \frac{\varphi \quad \neg\varphi}{\bot}$$

is valid in NML. From a semantical point of view, this is so, because the only designated value for NML is truth-value 1: a formula φ is a logical consequence of a theory T in NML if, under any evaluation, φ gets value 1 whenever all the premises in T get value 1 as well. It is clear then that no evaluation can assign the truth-value 1 to both φ and $\neg\varphi$.

When the (Exp) rule with respect to a negation $\neg$ is not valid for a logic, the logic is called $\neg$-*paraconsistent*, that is, when deduction from theory having a contradiction does not immediately trivialise. Paraconsistent variants of fuzzy logics have been already studied under the paradigm of the so-called *truth-preserving* logics. In these logics, φ is a consequence of T if, under any evaluation, φ gets a value at least as high as all the values got by the premises in T. In the case of NML, this variant, denoted NML$^{\leq}$ is paraconsistent since $\varphi \wedge \neg\varphi$ can get a value greater than 0. In this paper our main aim is to study and characterise another paraconsistent variant of NML obtained by taking the semi-open interval $(0, 1]$ as set of designated values, i.e. when we only exclude the *falsum* truth-value. In this logic, that will be denoted nf-NML (nf for non-falsity), φ follows from T if, under any evaluation, φ does not get value 0 whenever all the premises in T do not get value 0. The logic nf-NML is paraconsistent as well, and it lies between NML and NML$^{\leq}$.

II. Preliminaries: the NM logic

The *nilpotent minimum logic*, NML for short, was firstly introduced by two of the authors in [6] in order to formalize the logic of the nilpotent minimum t-norm, that was defined by Fodor in [7] as an example of an involutive left continuous t-norm which is not continuous.[2]

The language of NML consists of countably many propositional variables $p_1, p_2, \ldots$, binary connectives $\wedge$ (weak or lattice conjunction), $*$ (strong conjunction), $\to$ (implication), and the truth constant $\overline{0}$. Formulas, which will be denoted by

[1] A t-norm $*$ is a binary operation in $[0, 1]$ which is commutative, associative, non-decreasing and having 1 as neutral element and 0 as absorbent elements.

[2] Actually, Pei showed later in [15] that NML and NM-algebras are equivalent to Wang's $\mathcal{L}^*$ logic and R_0-algebras, respectively [16], [17].

lower case greek letters $\varphi, \psi, \chi, \ldots$, are recursively defined from propositional variables, connectives and truth-constant as usual. Further definable connectives and constants are as follows: $\neg\varphi$ stands for $\varphi \to \overline{0}$ and $\overline{1}$ stands for $\neg\overline{0}$.

As already mentioned in the previous section, NML is defined as the axiomatic extension of the monoidal t-norm logic MTL, also introduced in [6], by the axioms

$$\text{(INV)} \quad \neg\neg\varphi \to \varphi$$
$$\text{(WNM)} \quad (\psi * \varphi \to \bot) \vee (\psi \wedge \varphi \to \psi * \varphi).$$

Axiom (INV) forces the negation to be involutive, and axiom (WNM) forces the strong conjunction $*$ to coincide with the lattice or weak conjunction $\wedge$ wherever it does not vanish. It is worth observing that NML enjoys the following form of deduction theorem:

$$\Gamma \cup \{\varphi\} \vdash_{NM} \psi \text{ iff } \Gamma \vdash_{NM} \varphi^2 \to \psi,$$

where φ^2 is a shorthand for $\varphi * \varphi$. It is well known that NML is algebraizable and the class $\mathbb{NM}$ of all nilpotent minimum algebras is its equivalent algebraic quasivariety semantics [6].

A *nilpotent minimum algebra* (*NM-algebra*) $\mathbf{A} = \langle A, *, \to, \wedge, \vee, \mathbf{0}, \mathbf{1}\rangle$ is an involutive MTL-algebra (i.e. a bounded, commutative, integral, involutive, prelinear residuated lattice) that satisfies the following equation

$$\text{(WNM)} \quad (x * y \to \mathbf{0}) \vee (x \wedge y \to x * y) \approx \mathbf{1}.$$

We say that an NM-algebra is an NM-chain provided that its underlying lattice order (defined as $x \leq y$ if $x \to y = \mathbf{1}$) is total. Since the class $\mathbb{NM}$ of all NM-algebras is a proper subvariety of MTL-algebras, it inherits the subdirect representation of MTL-algebras, and thus each NM-algebra is representable as a subdirect product of NM-chains (see [6, Proposition 3]).

NM-chains can be easily characterised. Namely, given a bounded totally ordered set $(A, \leq)$, with upper bound $\mathbf{1}$ and lower bound $\mathbf{0}$, equipped with an involutive negation $\neg$, then defining for every $a, b \in A$,

$$a*b = \begin{cases} \mathbf{0}, & \text{if } b \leq \neg a \\ a \wedge b, & \text{otherwise} \end{cases} \quad a \to b = \begin{cases} \mathbf{1}, & \text{if } a \leq b \\ \neg a \vee b, & \text{otherwise,} \end{cases}$$

where $\wedge$ and $\vee$ denote meet and join in $(A, \leq)$, it follows that $\mathbf{A} = \langle A, *, \to, \wedge, \vee, \mathbf{0}, \mathbf{1}\rangle$ is an NM-chain. And moreover, every NM-chain is of this form. In particular, when $A = [0,1], \wedge = \min, \vee = \max$, and $\neg x = 1 - x$, then $\mathbf{A} = [\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$ is called the *standard NM-algebra*. It turns out that the variety $\mathbb{NM}$ is generated by the standard algebra $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$, and this means that the logic NML is sound and complete w.r.t. the semantics given by evaluations of formulas on $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$.

Theorem 2.1: [6] For any set of formulas $\Gamma \cup \{\varphi\}$, $\Gamma \vdash_{NM} \varphi$ iff, for every $[0,1]_{NM}$-evaluation e, if $e(\psi) = 1$ for all $\psi \in \Gamma$, then $e(\varphi) = 1$.

III. Non-falsity preserving companion of NML

We start by generalising the usual notion of 1-preserving logical consequence by considering the consequence relations $\models_{(a}$ and $\models_a$ that respectively preserve values above $a \in [0,1]$ both in a strict (for $a < 1$) and non-strict sense (for $0 > a$).

Definition 3.1: For any finite set of formulas $\Gamma \cup \{\varphi\}$, we define:

- $\Gamma \models_{(a} \varphi$ if, for any NM-evaluation e, if $e(\psi) > a$ for all $\psi \in \Gamma$, then $e(\varphi) > a$.
- $\Gamma \models_a \varphi$ if, for any NM-evaluation e, if $e(\psi) \geq a$ for all $\psi \in \Gamma$, then $e(\varphi) \geq a$.

Note that $\models_1$, by the above completeness result, coincides with $\vdash_{NM}$. Also, it is easy to check that $\models_{(a}$ is paraconsistent iff $a < 1/2$, while $\models_a$ is paraconsistent iff $a \leq 1/2$. Moreover, one can show that many of these logics colapse.

Proposition 3.1: The following properties hold:

1) $\models_a$, $\models_{(a}$ and $\models_1$ are the same logic for all $a \in (1/2, 1)$.
2) $\models_a$, $\models_{(a}$ and $\models_{(0}$ are the same logic for all $a \in (0, 1/2)$.

In the rest of the section we axiomatise the logic $\models_{(0}$, that we will refer to as the *non-falsity preserving* companion of NML. We borrow the terminology of 'non-falsity preserving logic' from Avron [1], where the author considers a similar companion for Łukasiewicz logic, although in fact the logic defined there is the non-falsity preserving companion of only the $\{\wedge, \vee, \neg\}$-fragment of Łukasiewicz logic, and the idea of the proof is totally different.

Next lemma is a key observation that tightly relates both logics $\models_1$ and $\models_{(0}$ through the negation connective.

Lemma 3.1: For every formula φ,

$$\psi \models_{(0} \varphi \text{ iff } \neg\varphi \models_1 \neg\psi \quad (\text{iff } \models_1 (\neg\varphi)^2 \to \neg\psi)$$

In particular, $\models_{(0} \varphi$ if, and only if, $\models_1 \neg(\neg\varphi)^2$.

Proof: By definition, $\psi \models_1 \varphi$ iff for every NM-evaluation e, if $e(\psi) = 1$ then $e(\varphi) = 1$; that is, if $e(\varphi) < 1$ then $e(\psi) < 1$, for all e; that is, $e(\neg\varphi) > 0$ then $e(\neg\psi) > 0$, for all e; iff $\neg\varphi \models_{(0} \neg\psi$. ■

Now we define an axiomatic system aimed to syntactically capture the logical consequence $\models_0$ that preserves the non-falsity.

Definition 3.2: The *non-falsity preserving companion* of NML, denoted nf-NML, is the logic defined by the following axioms and rules:

- Axioms: those of NML
- The rule of Adjunction (Adj): from φ and ψ derive $\varphi \wedge \psi$
- The rule (r-MP^2): from φ and $\varphi \to \neg(\neg\psi)^2$ derive ψ, whenever $\vdash_{NM} \varphi \to \neg(\neg\psi)^2$

Finally, using the above lemma and some further adjustments, we can prove that nf-NML is sound and complete w.r.t. the intended semantics. Details can be found in the paper [10] currently under submission.

Theorem 3.1: For any set of formulas $\Gamma \cup \{\varphi\}$, it holds that $\Gamma \vdash_{\text{nf-NM}} \varphi$ iff $\Gamma \models_{(0} \varphi$.

Since NM-algebras are *locally finite* (i.e. the NM-subalgebra generated by a finite set of elements of a given NM-algebra is finite), the logic enjoys the finite model property, and thus it

is decidable. Furthermore, due to the direct relation between $\models_{(0}$ and $\models_1$ shown in Lemma 3.1 above, the computational complexity of deciding whether a deduction $\Gamma \vdash_{\mathsf{nf\text{-}NM}} \varphi$ holds, with Γ finite, is the same as in the case of the NML logic, which is known to be **coNP**-complete, see e.g. [12], [14].

IV. The nf-NM logic and consistency operators

Among the plethora of paraconsistent logics proposed in the literature, the Logics of Formal Inconsistency (LFIs), proposed in [4] (see also [2]), play an important role, since they internalize in the object language the very notion of consistency by means of a specific unary connective (primitive or not), usually denoted as $\circ$.

Definition 4.1: Let L be a logic defined in a language L containing a negation $\neg$ and a unary operator $\circ$ whose deduction system is denoted by $\vdash$. L is an LFI (with respect to $\neg$ and $\circ$) if the following conditions hold:

(i) $\varphi, \neg\varphi \nvdash \psi$, for some formulas φ, ψ, i.e. L is not explosive w.r.t. $\neg$,
(ii) $\circ(\varphi), \varphi \nvdash \psi$, for some formulas φ and ψ,
(iii) $\circ(\varphi), \neg\varphi \nvdash \psi$, for some formulas φ and ψ, and
(iv) $\varphi, \neg\varphi, \circ(\varphi) \vdash \bot$, for every formula φ.

A consistency operator in an LFI logic can be primitive or it can be defined from other connectives of the language.

However, the nf-NM logic, although it is paraconsistent, it is not an LFI. Obviously, the consistency operator $\circ$ is not a primitive connective, but as we will show below it is not definable either. Anyway, similarly to what was done in the case of fuzzy logics preserving degrees of truth [5], we can expand nf-NML with a consistency operator $\circ$, that is, a unary operator such that the expanded logic satisfies the above properties (i)-(iv) and hence it becomes an LFI. This will be done in the rest of this section.

In the following, we will denote by $\mathcal{L}_\circ$ the expansion of the language of NML with $\circ$. And given a unary operator $\circ : [0,1] \to [0,1]$ we will denote by $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}^\circ}$ the the expansion of the standard algebra $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}}$ with $\circ$[3] and by $\models^\circ_{(0}$ the consequence relation that preserves non-falsity (defined as in Def. 3.1 for $a = 0$) but over the expanded language $\mathcal{L}_\circ$ and where evaluations interpret formulas on the expanded algebra $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}^\circ}$.

We start by considering the most general semantical conditions on $\circ$ such that the logic $\models^\circ_{(0}$ is an LFI, that is, such that the following conditions are satisfied:

- $\circ\varphi, \varphi, \neg\varphi \models^\circ_{(0} \bot$
- $\varphi, \circ\varphi \not\models^\circ_{(0} \bot$
- $\neg\varphi, \circ\varphi \not\models^\circ_{(0} \bot$

It immediately follows that these conditions are satisfied if, and only if, in the algebra $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}^\circ}$ the following conditions are in turn satisfied:

- for all $x \in [0,1]$, $x \wedge \neg x \wedge \circ x = 0$,
- there exists $x \in [0,1]$, such that $x \wedge \circ x \neq 0$,
- there exists $x \in [0,1]$, such that $\circ x \wedge \neg x \neq 0$,

which can be equivalently replaced by the next three conditions on $\circ$:

(C0) $\circ x = 0$ for all $x \in (0,1)$,
(1-NZ) $\circ 1 > 0$,
(0-NZ) $\circ 0 > 0$.

Definition 4.2: A unary operator $\circ : [0,1] \to [0,1]$ that satisfies conditions (C0), (1-NZ) and (0-NZ) will be called a *basic* consistency operator.

As we anticipated, such basic consistency operators are not definable in $[0,1]_{NM}$, and more generally in any NM-algebra. An argument for this claimn is the following. Since the 2-element Boolean $\mathbf{2}$ algebra over $\{0,1\}$ is a subalgebra of any NM-chain, if $\circ$ were definable (by a unary term), the only consistency operator that could be definable would be the one where $\circ(0) = \circ(1) = 1$, since this is the only compatible possibility when restricting $\circ$ to $\mathbf{2}$. Now, consider the NM-homomorphism $h : [0,1]_{NM} \to NM_3$, where NM_3 is the NM-subalgebra of $[0,1]_{NM}$ on the set $\{0, 1/2, 1\}$, defined as $h(x) = 1$ if $x > 1/2$, $h(1/2) = 1/2$ and $h(x) = 0$ if $x < 1/2$. Then it should be $h(\circ(x)) = \circ(h(x))$, but if $1/2 < x < 1$ or $0 < x < 1/2$, we have $h(\circ(x)) = 0$ while $\circ(h(x)) = 1$, contradiction.

In the following, we will call an element $x \in [0,1]$ *stricly positive* (SP) if $1/2 < x < 1$ and *strictly negative* (SN) if $0 < x < 1/2$.

It turns out that one cannot distinguish in $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}}$ the case $\circ(0) = a$ from the case $\circ(0) = b$ if both a and b are SP or SN, because there exists an isomorphism f of $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}}$ such that $f(a) = b$. Therefore, from conditions (1-NZ) and (0-NZ) above, we are left only four significant cases to consider for the values $\circ(0)$ and $\circ(1)$, that can be characterized by equations and inequations in $[\mathbf{0},\mathbf{1}]_{\mathbf{NM}}$. The proof is easy and thus omitted.

Proposition 4.1: For $\mathsf{x} \in \{0,1\}$, the following conditions hold:

[x-1] $\circ(\mathsf{x}) = 1$ is equivalently characterized by the equation $\neg(\circ(\mathsf{x})) = 0$,
[x-SP] $\circ(\mathsf{x}) \in (1/2, 1)$ is characterized by the inequation $(\circ(\mathsf{x}))^2 \wedge \neg(\circ(\mathsf{x})) > 0$,
[x-fix] $\circ(\mathsf{x}) = 1/2$ is characterized by the inequation $(\circ(\mathsf{x}) \leftrightarrow \neg(\circ(\mathsf{x})))^2 > 0$,
[x-SN] $\circ(\mathsf{x}) \in (0, 1/2)$ is characterized by the inequation $\circ(\mathsf{x}) \wedge (\neg(\circ(\mathsf{x})))^2 > 0$.

Combining these four conditions for $\mathsf{x} = 1$ and $\mathsf{x} = 0$, we obtain sixteen types of basic consistency operators $\circ$. In particular, the operator satisfying the conditions [1-1] and [0-1] is the maximal consistency operator $\circ_{max}$, i.e. the one such that $\circ_{max}(0) = \circ_{max}(0) = 1$.

Proposition 4.2: Two interesting properties of consistency operators are the following:

[3]Without risk of confusion, we will use the same symbol $\circ$ to denote the connective and a generic unary operation on the unit real interval $[0,1]$.

(i) The operator $\circ_{max}$ and Baaz-Monteiro's projection operator[4] Δ are interdefinable.
(ii) The maximal consistency operator $\circ_{max}$ (and the Δ operator) is definable from any of the sixteen types of consistency operators except from the one defined by the pair of conditions [0-SN] and [1-SN].

Proof: (i) To prove the first item it is enough to check that $\Delta(x) = \circ_{max}(x) \wedge x$ and also that $\circ_{max}(x) = \Delta(x \vee \neg x)$.
(ii) The second item can be proved by checking the following cases:

- if both $\circ(0), \circ(1) \geq 1/2$, then $\circ_{max}(x) = \neg((\neg(\circ(x)))^2)$ and $\Delta(x) = \circ_{max}(x) \wedge x$.
- if $\circ(1) \geq 1/2$ and $\circ(0) \in (0, 1/2)$, then $\Delta(x) = \neg((\neg(\circ(x))^2)^2)$ and $\circ_{max}(x) = \Delta(x \vee \neg x)$.
- finally, if $\circ(1) \in (0, 1/2)$ and $\circ(0) \geq 1/2$, then $\Delta(x) = \neg((\neg(\circ(\neg x))^2)^2)$ and $\circ_{max}(x) = \Delta(x \vee \neg x)$.

■

Note that the converse of the previous results does not hold in the sense that if we add $\circ_{max}$ to the algebra $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$, it is not possible to recover the previous consistency operators, of course with the exception of $\circ_{max}$ itself, because Δ and $\circ_{max}$ are crisp operators (i.e. they only take values 0 or 1) and the operations of the algebra $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$ are classical when restricted to $\{0, 1\}$.

In the next section we focus our attention to the case of adding the maximal consistency operator $\circ_{max}$ to the logic nf-NML. The expansions of nf-NML with the remaining fifteen cases of consistency operators can be dealt in a similar way, except for the case of operators where both $\circ(0)$ and $\circ(1)$ are SN.

V. Expanding nf-NML with the maximal consistency operator $\circ_{max}$

In this final part of the paper we formally define and axiomatise the expansion of the logic nf-NM with the maximal consistency operator $\circ_{max}$, i.e. the basic consistency operator $\circ$ further satisfying:

[1-1] $\circ(1) = 1$
[0-1] $\circ(0) = 1$

As already noted before, the crucial observation is that, in this case, $\circ_{max}$ and the Baaz-Monteiro operator Δ are interdefinable: $\Delta(x) = \circ_{max}(x) \wedge x$, and $\circ_{max}(x) = \Delta(x \vee \neg x)$.

We start by axiomatising first the 1-preserving logic $\models_1^{\circ_{max}}$ and then, based on that, we will axiomatise the non-falsity preserving logic $\models_{(0}^{\circ_{max}}$. In the following we introduce the following abbreviation: $\Delta\varphi := (\circ\varphi) \wedge \varphi$.

Definition 5.1: $\mathrm{NML}_\circ^{\max}$ is the logic defined by the following axioms and rules:

- Axioms of NML
- Consistency Axioms:
  (C0) $\neg(\circ\varphi \wedge \varphi \wedge \neg\varphi)$
  ($\top$-1) $\circ\top$
  ($\bot$-1) $\circ\bot$
- Modus ponens: (MP) $\dfrac{\varphi, \quad \varphi \to \psi}{\psi}$
- Congruence rule:
  (Cong) $\dfrac{(\varphi \leftrightarrow \psi) \vee \chi}{(\circ\varphi \leftrightarrow \circ\psi) \vee \chi}$

Observe that it is easy to check that the following three inference rules

$$\frac{\varphi}{\circ\varphi}, \quad \frac{\neg\varphi}{\circ\varphi}, \quad \frac{\varphi}{\Delta\varphi}$$

are derivable in $\mathrm{NML}_\circ^{\max}$ from the axioms ($\top$-1) and ($\bot$-1) and the rule (Cong). Moreover, one can also check that the formula $\circ\varphi \vee \neg\circ\varphi$, stating that $\circ$ is a Boolean operator, can be proved to be a theorem of the logic as well: by applying the (Cong) rule to the axiom (C0), equivalently expressed as $\varphi \vee \neg\varphi \vee \neg\circ\varphi$, one gets $\circ\varphi \vee \neg\varphi \vee \neg\circ\varphi$, and applying (Cong) again, one gets $\circ\varphi \vee \circ\varphi \vee \neg\circ\varphi$, which is equivalent to $\circ\varphi \vee \neg\circ\varphi$.

Theorem 5.1: $\mathrm{NML}_\circ^{\max}$ is a sound and complete axiomatisation of $\models_1^{\circ_{max}}$.

Proof: (Sketch) First of all, note that $\mathrm{NML}_\circ^{\max}$ is an expansion of NM with axioms plus the (Cong) inference rule, so the logics keeps being algebraizable, and hence it is strongly complete with respect to the class (quasivariety) of $\mathrm{NML}_\circ^{\max}$-algebras. Moreover, the (Cong) inference rule is closed by disjunctions (thanks to the addition of the clause '$\vee\chi$' in the premise and in the conclusion of the rule). Then, by results in [3], the quasi-variety of $\mathrm{NM}_\circ^{\max}$-algebras is semilinear, that is, it is generated by its linearly ordered members. Hence, if an equation fails in an $\mathrm{NML}_\circ^{\max}$-algebra, it also fails in a $\mathrm{NML}_\circ^{\max}$-chain. The final observation is the fact that every embedding from a countable NM-chain into $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$ (which always exists) easily extends to an embedding from a $\mathrm{NM}_\circ^{\max}$-chain into the standard algebra $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}_\circ^{\max}}$, hence if an equation fails in a $\mathrm{NM}_\circ^{\max}$-chain it will aso fail in the standard chain $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}_\circ^{\max}}$. Therefore, if $\Gamma \not\vdash \varphi$ there will always exist an evaluation over an evaluation e on $[\mathbf{0}, \mathbf{1}]_{\mathbf{NM}}$ such that $e(\Gamma) = 1$ and $e(\varphi) < 1$. ■

It is worth noticing that, from this completeness result, it follows that the set of axioms for the Δ operator (defined above as $\Delta\varphi := (\circ\varphi) \wedge \varphi$), as proposed e.g. in [11] to syntactically characterizing it, are provable in $\mathrm{NML}_\circ^{\max}$, since they are obviously valid formulas for $\models_1^{\circ_{max}}$.

Now we move from the logic $\models_1^{\circ_{max}}$ to the paraconsistent logic $\models_{(0}^{\circ_{max}}$. Note that $\models_{(0}^{\circ_{max}}$ can be described in terms of $\models_1^{\circ_{max}}$ by using the Δ connective. Namely, it holds that

$$\begin{aligned} &\{\varphi_1, \ldots, \varphi_n\} \models_{(0}^{\circ_{max}} \psi \\ &\quad \text{iff } \{\nabla\varphi_1, \ldots, \nabla\varphi_n\} \models_1^{\circ_{max}} \nabla\psi, \\ &\quad \text{iff } \models_1^{\circ_{max}} (\nabla\varphi_1 \wedge \ldots \wedge \nabla\varphi_n) \to \nabla\psi, \\ &\quad \text{iff } \models_1^{\circ_{max}} \nabla(\varphi_1 \wedge \ldots \wedge \varphi_n) \to \nabla\psi, \end{aligned}$$

where $\nabla = \neg\Delta\neg$. Indeed, for any evaluation e, it holds that $e(\varphi) > 0$ iff $e(\neg\Delta\neg\varphi) = 1$.

Now we introduce an axiomatic system for the logic $\models_{(0}^{\circ_{max}}$.

[4] Recall that the Baaz-Monteiro unary operator Δ on the unit interval $[0, 1]$ is defined as $\Delta(1) = 1$ and $\Delta(x) = 0$ for $x < 1$.

Definition 5.2: nf-NML$_\circ^{\max}$ is the logic defined by the following axioms and rules:

- Axioms of NML$_\circ^{\max}$
- Rule of Adjunction: (Adj) $\dfrac{\varphi, \quad \psi}{\varphi \wedge \psi}$
- Restricted Modus Ponens:
  (r-MP) $\dfrac{\varphi, \quad \varphi \to \psi}{\psi}$, if $\vdash_{\text{NML}_\circ} \varphi \to \psi$
- Restricted Congruence:
  (r-Cong) $\dfrac{(\varphi \leftrightarrow \psi) \vee \chi}{(\circ\varphi \leftrightarrow \circ\psi) \vee \chi}$, if $\vdash_{\text{NML}_\circ} (\varphi \leftrightarrow \psi) \vee \chi$
- Reversed Necessitation for ∇:
  (r-∇Nec) $\dfrac{\nabla\varphi}{\varphi}$

Observe that the rule of necessitation for ∇:

$$(\nabla\text{Nec}) \quad \frac{\varphi}{\nabla\varphi},$$

which is the reverse of (r-∇Nec), is derivable. In fact, it is a direct consequence of the fact that, by definition, $\neg\Delta\neg\varphi = (\neg\circ\neg\varphi)\vee\varphi$. On the other hand, from (r-∇Nec) it easily follows that the rule

$$\frac{\neg\circ\neg\varphi}{\varphi},$$

is also derivable since clearly $\neg\circ\neg\varphi \to (\neg\circ\neg\varphi) \vee \varphi$ is a theorem of NML$_\circ^{\max}$.

Theorem 5.2: nf-NML$_\circ^{\max}$ is a sound and complete axiomatisation of $\models_{(0}^{\circ_{max}}$.

Proof: Suppose $\{\varphi_1, \ldots, \varphi_n\} \models_{(0}^{\circ_{max}} \psi$. Then, as observed above, this holds iff $\models_1^{\circ_{max}} \nabla(\varphi_1 \wedge \ldots \wedge \varphi_n) \to \nabla\psi$, and by completeness of NML$_\circ^{\max}$, iff $\vdash_{NM_\circ^{\max}} \nabla(\varphi_1 \wedge \ldots \wedge \varphi_n) \to \nabla\psi$. Therefore, in NML$_\circ^{\max}$ there is a proof

$$\Pi_1, \ldots, \Pi_r = \nabla(\varphi_1 \wedge \ldots \wedge \varphi_n) \to \nabla\psi,$$

where each Π_i (with $1 \leq i < r$) is either an axiom of NML$_\circ^{\max}$, it has been obtained from a previous Π_k by the (Cong) rule, or has been obtained from previous Π_k, Π_j $(k, j < r)$ by the application of Modus ponens rule. Then, in order to get a proof of φ from $\psi_1, \ldots, \psi_n$ in nf-NML$_\circ^{\max}$ we only need do the following:

(i) add two previous steps Π_0^1 and Π_0^2, where
  - $\Pi_0^1 = \varphi_1 \wedge \ldots \wedge \varphi_n$, obtained from the premises by the (Adj) rule,
  - $\Pi_0^2 = \nabla(\varphi_1 \wedge \ldots \wedge \varphi_n)$, obtained from Π_0^1 by the (∇Nec) rule

(ii) add two final steps Π_{r+1} and Π_{r+2}, where
  - $\Pi_{r+1} = \neg\Delta\neg\psi$, obtained by the application of the (r-MP) rule to Π_0 and Π_r, and
  - $\Pi_{r+2} = \psi$, obtained by applying the rule (r-ΔNec) to Π_{r+1}.

Therefore, the sequence $\Pi_0^1, \Pi_0^2, \Pi_1, \ldots \Pi_r, \Pi_{r+1}, \Pi_{r+2}$ is a proof of ψ from $\{\varphi_1, \ldots, \varphi_n\}$ in the logic nf-NML$_\circ^{\max}$, with the proviso that the applications of the modus ponens and the (Cong) rules in the original proof $\Pi_1, \ldots \Pi_r$ in NML$_\circ^{\max}$ have to be replaced now by applications of the corresponding restricted rules (r-MP) and (r-Cong). ■

VI. Conclusions

The Nilpotent Minimum logic NML is an axiomatic extension of the Monoidal t-norm based fuzzy logic MTL that enjoys nice properties. In this paper we have explored the definition and axiomatisation of the logic nf-NML, the non-falsity preserving companion of the Nilpotent Minimum logic. nf-NML is a $\neg$-paraconsistent logic, but it does not belong to the family of well-behaved paraconsistent logics known as Logics of Formal Inconsistency. These logics are characterised by having in its language a unary connective $\circ$ (primitive or definable) by which one can internalize in the object language the notion of consistency. To remedy this problem we have considered expanding nf-NML with a consistency operator, and have presented a complete axiomatic system for this expansion in the particular case of the maximum consistency operator $\circ_{\max}$. Nevertheless, let us notice that the same kind of approach could be used to define the logics corresponding to each of the remaining fourteen basic consistency operators described in Proposition 4.1 for which the Δ operator is definable, see (ii) of Proposition 4.2. To do this, in Definition 5.1 one has to:

(1) Replace axioms ($\top$-1) and ($\bot$-1) respectively by suitable axioms corresponding to any pair of conditions [x-SP], [x-fix], [x-SN].
(2) Change the working definition of Δ in terms of $\circ$ in Definition 5.1 (i.e. $\Delta\varphi := (\circ\varphi) \wedge \varphi$) in each case according to the three expressions shown in the proof of Proposition 4.2.

As future work we plan to generalise the approach to other axiomatic extensions of MTL with a (primitive or definable) involutive negation.

Acknowledgments: The authors thank the anonymous reviewers for their comments and suggestions. The authors acknowledge support by the MOSAIC project (EU H2020-MSCA-RISE- 2020 Project 101007627). Esteva and Godo acknowledge partial support by the Spanish project LINEXSYS (PID2022-139835NB-C21), and Gispert acknowledges partial support by the Spanish project SHORE (PID2022-141529NB-C21), both projectsbeing's g funded by MCIN/AEI/10.13039/501100011033. Gispert also acknowledges the project 2021 SGR 00348 funded by AGAUR.

References

[1] A. Avron. Paraconsistent fuzzy logic preserving non-falsity. *Fuzzy Sets and Systems* 292, 75-84, 2016.
[2] W. Carnielli, M.E. Coniglio. *Paraconsistent Logic: Consistency, Contradiction and Negation*. Vol. 40 of Logic, Epistemology, and the Unity of Science, Springer, 2016.
[3] P. Cintula, C. Noguera C. A general framework for mathematical fuzzy logic. In: Cintula P, Hájek P, Noguera C (eds) *Handbook of Mathematical Fuzzy Logic - Volume 1*. Studies in logic, Mathematical Logic and Foundations, vol 37. College Publications, London, pp. 103-207.

[4] W. Carnielli, M.E. Coniglio, and J. Marcos. Logics of Formal Inconsistency. In Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic* (2nd. edition), volume 14, pages 1–93. Springer, 2007.
[5] M.E. Coniglio, F. Esteva, L. Godo. Logics of formal inconsistency arising from systems of fuzzy logic, *Logic Journal of the IGPL* 22, n. 6: 880–904, 2014.
[6] F. Esteva and L. Godo. Monoidal t-norm based logic: towards a logic for left-continuous t-norms. *Fuzzy Sets and Systems*, 124, 271–288, 2001.
[7] J. Fodor, *Nilpotent minimum and related connectives for fuzzy logic*, Proc. FUZZ–IEEE '95, 1995, pp. 2077–2082.
[8] J. Gispert. Axiomatic extensions of the nilpotent minimum logic. *Reports on Mathematical Logic* 37: 113-123, 2003
[9] J. Gispert. Finitary Extensions of the Nilpotent Minimum Logic and (Almost) Structural Completeness. *Studia Logica* 106(4): 789-808 (2018)
[10] J. Gispert, F. Esteva, L. Godo. On Nilpotent Minimum logics defined by lattice filters. Submitted
[11] P. Hájek. *Metamathematics of Fuzzy Logic*. Vol. 4 of Trends in Logic - Studia Logica Library. Kluwer, 1998.
[12] Z. Haniková. Computational complexity of propositional fuzzy logics. Chapter X of *Handbook of Mathematical Fuzzy Logic*, Volume 2, College Publications, pp. 793–851, 2011.
[13] S. Jenei and F. Montagna. A completeness proof of Esteva and Godo's MTL logic. *Studia Logica* 70: 183-192, 2002.
[14] E. Marchioni. On computational complexity of semilinear varieties. *Journal of Logic and Computation*, 18(6): 941-958, 2008.
[15] D. Pei. On equivalent forms of fuzzy logic systems NM and IMTL. *Fuzzy Sets and Systems* 138, 187-195, 2003.
[16] G.J. Wang. A formal deductive system for fuzzy propositional calculus, *Chinese Sci. Bull.* 42, 1521-1526, 1997.
[17] G.J. Wang. On the logic foundation of fuzzy reasoning, *Inform. Sci.* 117, 47-88, 1997.

FLocalX - De Explicaciones Difusas Locales a Globales para Clasificadores de Caja Negra

Guillermo Fernández
Laboratorio de Sistemas Inteligentes y Minería de Datos
Universidad de Castilla-La Mancha
Albacete, España
Guillermo.Fernandez@uclm.es

Jose A. Gámez, Jose M. Puerta
Departamento de Sistemas Informáticos
Universidad de Castilla-La Mancha
Albacete, España
{Jose.Gamez,Jose.Puerta}@uclm.es

Juan A. Aledo
Departamento de Matemáticas
Universidad de Castilla-La Mancha
Albacete, España
JuanAngel.Aledo@uclm.es

Riccardo Guidotti, Mattia Setzu
Universidad de Pisa
Pisa, Italia
{riccardo.guidotti,mattia.setzu}@unipi.it

Fosca Giannotti
Scuola Normale Superiore
Pisa, Italia
fosca.giannotti@sns.it

***Resumen*—La necesidad de explicación para complejos algoritmos de aprendizaje automático ha provocado el auge del campo de la Inteligencia Artificial Explicable (*eXplainable Artificial Intelligence*, XAI). Se puede distinguir entre *explicaciones locales*, que se enfocan en explicar la clasificación de una instancia en particular, y *explicaciones globales*, que tratan de mostrar una visión global del funcionamiento interno del modelo. En este trabajo proponemos FLocalX, una metodología que permite construir una explicación global, expresada en términos de reglas difusas, a partir de un conjunto de explicaciones locales. Esta explicación global se obtiene mediante un proceso de optimización guiado por una metaheurística, que se ha instanciado a un algoritmo genético. Hemos realizado una evaluación experimental considerando distintos conjuntos de datos y métricas (fidelidad, complejidad, etc.) habituales en las propuestas de XAI. Los resultados obtenidos muestran que FLocalX es capaz de generar explicaciones globales compactas y comprensibles que resumen fielmente el comportamiento del clasificador que explican.**

***Index Terms*—XAI, Optimización, Sistemas Basados en Reglas Difusas, Explicaciones Locales, Explicaciones Globales**

I. Introducción

En los últimos tiempos, la gran cantidad disponible de datos y nuevos paradigmas de aprendizaje automático (*deep learning*, *ensembles*, etc.) ha fomentado la obtención de nuevos (y complejos) modelos que han sido incorporados a tareas previamente inabordables [5], [24]. Sin embargo, la creciente complejidad normalmente viene de la mano de una pérdida en la interpretabilidad del modelo [2], lo que puede ser una gran desventaja en campos críticos como la medicina, la legislación, la aviación, etc. La propia legislación europea trata este tema a través del *derecho a la explicación* incluido en el Reglamento General de Protección de Datos [16], que afecta tanto a humanos como a técnicas de inteligencia artificial. La Inteligencia Artificial Explicable (XAI) [3], [9] trata de promover el uso de la interpretabilidad y la explicabilidad para comprender los modelos de caja negra, y facilitar su uso en contextos sensibles y áreas críticas.

Una distinción importante en XAI, es si un método genera explicaciones *locales* o *globales*. Las *explicaciones locales* razonan las decisiones tomadas ante instancias individuales, mientras que las *explicaciones globales* proporcionan una visión general del comportamiento completo del modelo. Un tipo común de explicaciones locales son las *explicaciones factuales y contrafactuales* [6], [8]. Las explicaciones factuales tratan sobre el razonamiento que soporta una decisión, mientras que las explicaciones contrafactuales reflejan los cambios necesarios para revertir dicha decisión. LORE (LOcal Rule-based Explainer) [8] es un conocido método de XAI que genera tanto explicaciones factuales como contrafactuales en forma de reglas de clasifiación. Para ello, genera un vecindario (dataset) cercano a la instancia que es usado para inducir un árbol de decisión *crisp*, del que se extraen las reglas usadas como explicaciones. FLARE [7] trabaja en la misma línea de investigación, pero empleando un árbol de decisión *difuso*.

En trabajos previos podemos ver que las explicaciones locales se han tomado como punto de partida para la construcción de explicaciones globales, difuminando la línea entre ambas. En [11] se convierten valores locales Shapley en explicaciones globales a través de descomposición funcional. Otros métodos fusionan ambos tipos de explicaciones mediante la importancia de las variables [15], relevancia de conceptos [17] y resúmenes de estrategia [13]. GLocalX [19], en el cual está inspirado este trabajo, fusiona explicaciones locales formadas por reglas *crisp* para construir una *teoría de explicación global*.

En este trabajo, introducimos *FLocalX*, una metodología agnóstica al modelo subyacente para explicar clasificadores de caja negra mediante la obtención de teorías de explicación globales, consistentes en un sistema de reglas *difusas* obtenido a partir de *explicaciones locales difusas*. Estas teorías de explicación globales replican el comportamiento de los clasificadores de caja negra subyacentes, y se pueden usar para generar explicaciones factuales a nuevas instancias, a la vez que proporcionan un entendimiento global del modelo.

Así, un usuario puede entender cómo funciona el clasificador y cómo se comportará ante nuevas instancias, en lugar de generar explicaciones nuevas para cada instancia desconocida. La construcción de una teoría global con reglas difusas, en lugar de crisp, resulta más comprensible, flexible y fiel al modelo de caja negra. Las reglas difusas aprovechan las etiquetas lingüísticas, que aumentan su legibilidad asociando cada premisa con conceptos lingüísticos de alto nivel, y han sido utilizadas previamente para diseñar sistemas explicables [1], [22], así como para generar explicaciones locales [20].

Este trabajo se estructura de la siguiente manera. La Sección II presenta el problema e identifica los elementos relevantes. La Sección III ilustra el flujo de trabajo de nuestra propuesta. La Sección IV muestra los experimentos realizados y analiza el rendimiento de FLocalX. Finalmente, la Sección V presenta las conclusiones e indica algunas líneas de trabajo futuro.

II. Conceptos previos

En un problema de clasificación, una instancia $x = (x_1, \ldots, x_n) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$, donde $\mathcal{X}_1, \ldots, \mathcal{X}_n$ son n *variables de entrada*, se asocia con una decisión $y \in \mathcal{Y} = \{y_1, \ldots, y_m\}$ mediante una función (clasificador) $f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \rightarrow \mathcal{Y}$. Escribimos $f(x) = y$ para expresar la clasificación y dada a x. Decimos que n_{cont} (resp. n_{disc}) es el número de variables continuas (resp. discretas) en $\mathcal{X}$, t.q. $0 \leq n_{cont}, n_{disc} \leq n$, $n_{cont} + n_{disc} = n$.

Asumimos que asociada a cada variable de entrada $\mathcal{X}_i$, existe una variable difusa (lingüística) $\mathcal{F}_i = \{v_{i,1} \ldots, v_{i,k_i}\}$ definida por una partición de Ruspini [1] de k_i conjuntos difusos ordenados (ver Figura 1)[1]. Con v_{i,z_i} nos referimos indistintamente tanto al conjunto difuso como a su etiqueta lingüística asociada. Un conjunto difuso triangular se define por una terna de puntos reales: (comienzo, pico, fin), p. ej. *teen* = (15, 15, 25) y *young* = (15, 25, 45) en la Figura 1a. Si conocemos los valores mínimos y máximos de $dom(\mathcal{X}_i)$, la partición se puede especificar usando $k_i - 2$ valores. Dado un valor $\delta \in dom(\mathcal{X}_i)$, sea $\mu_i(\delta) = (\mu_{i,1}(\delta), \ldots, \mu_{i,k_i}(\delta))$ el vector de grados de pertenencia de δ a los k_i conjuntos difusos de $\mathcal{F}_i$. Es decir, $\mu_{i,z_i}(\delta)$ es el grado de pertenencia de δ al conjunto v_{i,z_i}. Un modificador lingüístico es una función que altera la función de pertenencia (y la forma) de un conjunto difuso (ver Figura 1b). En este trabajo, utilizamos dos de los modificadores más comunes, *muy* y *algo*, siendo $\mu^{muy}_{i,z_i}(x_i) = (\mu_{i,z_i}(x_i))^2$ y $\mu^{algo}_{i,z_i}(x_i) = \sqrt{\mu_{i,z_i}(x_i)}$. Finalmente, para las variables discretas, podemos interpretar cada valor como una etiqueta lingüística cuyos conjuntos difusos asociados tienen un grado de pertenencia 1 en caso de que la instancia tome dicho valor y 0 en caso contrario.

Sea $b()$ un clasificador (posiblemente un modelo de caja negra) entrenado con un conjunto de datos $TR = \{(x^t_1, \ldots, x^t_n, y^t)\}^T_{t=1}$ cuyo proceso de toma de decisiones necesita ser explicado. Sea $e = \{r_1, \ldots, r_e\}$ una explicación multi-regla formada por una (o más) reglas de decisión difusas. Cada regla $r = P(r) \rightarrow y(r)$ consiste en un conjunto de premisas en forma conjuntiva $P(r) = p_{s_1} \wedge \cdots \wedge p_{s_r}$ y un consecuente $y(r) \in \mathcal{Y}$. Cada premisa $p_i = \langle \mathcal{F}_i, v_{i,z_i} \rangle$ es un par atributo-valor. Para las variables continuas, $\mathcal{F}_i$ es una variable difusa y v_{i,z_i} es uno de sus conjuntos difusos. Para las variables discretas, $\mathcal{F}_i = \mathcal{X}_i$ y v_{i,z_i} es un valor de su dominio. Un ejemplo sería la siguiente explicación para una solicitud de préstamo de 30k € de un usuario $x = \{(edad = 30), (trabajo = Contable), (ingresos = 20k)\}$:

$$e = \{(r_1 : edad\ es\ joven \wedge trabajo\ es\ Contable \rightarrow acepta),$$
$$(r_2 : edad\ es\ adulto \wedge ingresos\ son\ medios \rightarrow acepta)\}$$

Cabe destacar que dada una explicación multi-regla e que explica una instancia x, entonces $y(r) = b(x)$ para todo $r \in e$. Las reglas difusas se diferencian de las reglas crisp en que, mientras que una regla crisp tiene un emparejamiento binario (0 o 1) con la instancia x, una regla difusa r tiene un *grado de emparejamiento* con la instancia, $md(r, x)$, definido como:

$$md(r, x) = \min_{i \in \{s_1, \ldots, s_r\}} \{\mu_{i,z_i}(x_i)\} \in [0, 1].$$

Una teoría de explicación $E = e_1 \cup \cdots \cup e_q$ consiste en la unión de explicaciones que pueden tener distintas conclusiones.

Así, dado un clasificador $b()$, un conjunto de instancias $X = \{x^1, \ldots, x^q\}$ y sus explicaciones locales $\{e_1, \ldots, e_q\}$, el Problema de pasar de Explicaciones Locales a Globales consiste en obtener una teoría de explicación global $E_G = e'_1 \cup \cdots \cup e'_{q'}$ que agrega las explicaciones locales para resumir la lógica de b.

III. FLocalX: una metodología para obtener una explicación global a partir de explicaciones locales difusas

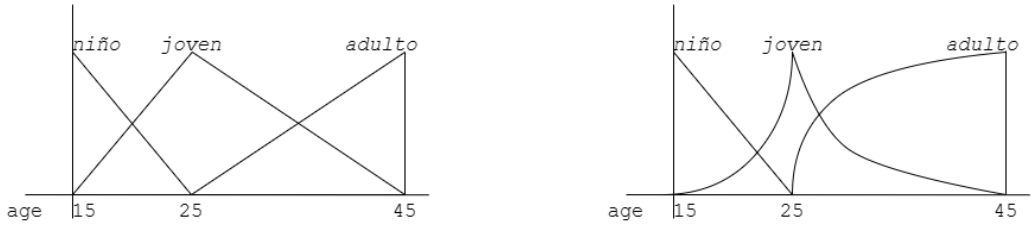
En este artículo proponemos FLocalX, una metodología para obtener una explicación global de un clasificador a partir de un conjunto de explicaciones locales difusas relativas al mismo. FLocalX toma como entrada un conjunto de instancias X y una teoría de explicación $E_L = e_1 \cup \cdots \cup e_q$, formada por la unión de las explicaciones de cada instancia en X, y genera una teoría de explicación global E_G aplicando los siguientes pasos:

A. *Transformar* (mapear) los conjuntos difusos locales $\mathcal{F}^j$ definidos para cada $e_j \in E_L$ a una definición de conjuntos difusos comunes $\mathcal{F}^C$ obteniendo una teoría de explicación con conjuntos comunes E_C.
B. *Codificar* E_C en una representación simple y única que será la configuración (explicación) inicial C^{E_C} del proceso de optimización.
C. *Generar* la teoría de explicación global E_G desde C^{E_C} mediante un proceso de optimización.

III-A. Mapeo de los Conjuntos Difusos Locales a Globales

Las explicaciones locales podrían no compartir las mismas definiciones de variables difusas, y la misma etiqueta lingüística podría tener un significado distinto en cada explicación.

[1]Empleamos funciones triangulares de pertenencia por simplicidad y conveniencia. Esta metodología permite otros tipos de funciones de pertenencia (Gaussianas, trapezoidales, etc.) para representar los conjuntos difusos, siempre que cubran el dominio completo de las variables.



(a) Variable difusa *age*

(b) Modificadores: *muy joven* and *algo adulto*.

Figura 1: Partición difusa fuerte para una variable difusa *age*

Para homogeneizar las definiciones locales $\mathcal{F}_i^1, \ldots, \mathcal{F}_i^{|E_L|}$ que existen para cada variable $\mathcal{X}_i$, establecemos una definición global $\mathcal{F}^C$ dividiendo el dominio de las variables numéricas en conjuntos de igual anchura, si bien se puede partir de otros conjuntos, p.e. definidos por expertos.

Dados dos conjuntos difusos $v_{i,z_i} \in \mathcal{F}_i$ y $v'_{i,z'_i} \in \mathcal{F}'_i$, calculamos su similaridad [23] como

$$S(v_{i,z_i}, v'_{i,z'_i}) = A(v_{i,z_i} \cap v'_{i,z'_i})/A(v_{i,z_i} \cup v'_{i,z'_i})$$

donde $A(v)$ es el área del conjunto v. Como es habitual en la literatura, utilizamos min para modelar la intersección y max para la unión. Así, dada una variable $\mathcal{X}_i$, definimos el mapeo

$$M(v'_{i,z'_i}, \mathcal{F}_i) = \arg \max_{v_{i,z_i} \in \mathcal{F}_i} S(v_{i,z_i}, v'_{i,z'_i}), \qquad (1)$$

que toma un conjunto difuso $v'_{i,z'_i} \in \mathcal{F}'_i$ y devuelve el conjunto $v_{i,z_i} \in \mathcal{F}_i$ con la mayor similaridad. Se obtiene E_C aplicando la Eq. 1 a cada premisa de cada $e_i \in E_L$.

III-B. Codificación de la Teoría de Explicación

En FLocalX, redefinimos el objetivo de construir una teoría de explicación global como el proceso de optimizar el Sistema Basado en Reglas Difusas (SBRD) formado por las reglas de decisión difusas de E_C. Para ello necesitamos una codificación de E_C, es decir, una representación de una solución potencial al problema que usará el proceso de optimización.

El objetivo de la optimización de FLocalX es doble: *(i)* mantener o mejorar la capacidad del SBRD de imitar el clasificador y *(ii)* hacer el SBRD tan compacto como sea posible (en número de reglas) para favorecer la interpretabilidad [9]. Inspirados por [4], utilizamos un algoritmo genético como proceso de optimización del SBRD. Nos centramos en optimizar dos elementos del SRBD:

- **Estructura superficial**. Define la regla como la relación entre las variables de entrada y salida. Consta de dos elementos clave: *(i)* los modificadores lingüísticos, y *(ii)* las premisas de una regla. La optimización puede alterar el modificador lingüístico correspondiente a una premisa $p = \langle \mathcal{F}_i, v_{i,z_i} \rangle$, el conjunto difuso v_{i,z_i} asociado con p, y si $\mathcal{F}_i$ aparece o no en una regla.
- **Estructura profunda**. Es una descripción más específica que expande la estructura superficial con las definiciones de las funciones de pertenencia. La optimización sólo afecta a dichas funciones. Cabe destacar que el uso de un algoritmo metaheurístico puede reducir notablemente la explicabilidad de un sistema en pos de una mayor precisión. Controlamos esto mediante la preservación de la forma de las particiones difusas (particiones triangulares de Ruspini) como se explica en la Sección II.

Configuración. Cada configuración C del proceso de optimización representa una teoría de explicación E, mostrada gráficamente en la Figura 2. Para este propósito, usaremos una configuración de cuatro partes ($C_F + C_R + C_H + C_U$):

- C_F es la codificación de las variables difusas. Asumimos que los valores mínimo y máximo del $dom(\mathcal{X}_i)$ son conocidos. Como ejemplo, en la Figura 1a conocemos $k_i = 3$, $min = 15$ y $max = 40$, luego tenemos un valor libre (25) para codificar los tres conjuntos difusos: $\{(15, 15, 25); (15, 25, 40); (25, 40, 40)\}$. Tan solo cambiando el valor 25 por 20, modificamos la semántica de la variable, obteniendo una nueva partición: $\{(15, 15, 20); (15, 20, 40); (20, 40, 40)\}$. Así, C_F tiene una longitud de $(\sum_{i=1}^{n_{cont}} k_i) - 2$ elementos, todos números reales.
- C_R es la codificación de las reglas. Tiene una longitud de $n \cdot |E|$ elementos, donde $|E|$ es el número de reglas en la teoría de explicación. Cada n elementos consecutivos codifican una regla tomando valores del conjunto $\{0, \ldots, k_i\}$, donde 0 representa que la i-ésima variable no aparece en la regla y 1 hasta k_i identifican cada conjunto difuso o valor de la variable $\mathcal{F}_i$, dependiendo de si $\mathcal{X}_i$ es numérica o categórica.
- C_H es la codificación de los modificadores lingüísticos. Tiene una longitud de $n_{cont} \cdot |E|$ elementos, que pertenecen al conjunto $\{-1, 0, 1\}$ representando que no hay modificador lingüístico (-1), muy (0) o $algo$ (1), para la variable continua (difusa) asociada a esa posición.
- C_U es la codificación de las reglas usadas. Tiene $|E|$ elementos que indican si una regla se usa en el SBRD (1) o no (0).

III-C. Generación de la Teoría de Explicación Global

Al tener simultáneamente que *(i)* simular el clasificador $b()$ de forma precisa, y que *(ii)* obtener un SBRD compacto, la función objetivo debe tener en cuenta ambos aspectos. Medimos el primer objetivo como la curva bajo el área ROC (AUC) para tener en cuenta los conjuntos de datos no balanceados, y medimos el segundo objetivo como el número

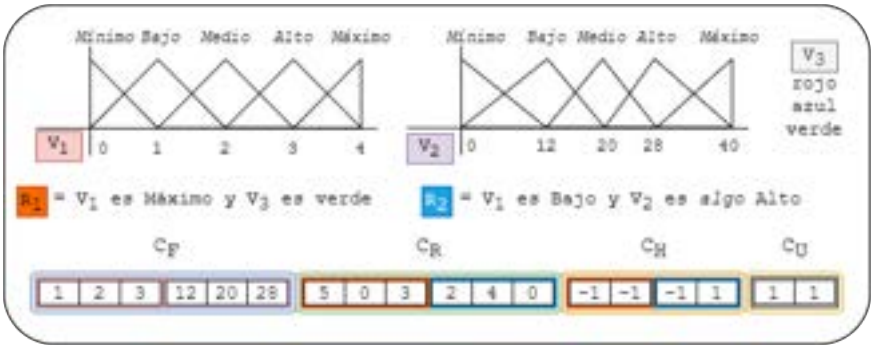


Figura 2: Representación de una codificación de un SBRD

de reglas utilizado en el sistema. Así, la función objetivo $f(C)$ a maximizar es la siguiente:

$$f(C) = \alpha \cdot (1 - \frac{\sum_{i=0}^{|C_U|-1} C_U[i]}{|C_U|}) + (1 - \alpha) \cdot AUC$$

usando α para balancear los dos valores a optimizar. Otras posibles funciones objetivo tendrían cabida en la metodología propuesta.

FLocalX puede usar cualquier algoritmo metaheurístico como optimizador para obtener la teoría de explicación global E_G. En este trabajo, hemos utilizado un algoritmo genético [12], cuyo diseño se muestra a continuación.

- **Población inicial.** La población inicial de tamaño $\rho+1$ se genera de manera informada alterando una configuración conocida (C^{E_C}). Dada una configuración inicial que representa un SBRD, C_F, C_R, C_H y C_U, se generan $\lceil \rho/4 \rceil$ cromosomas aplicando el operador de mutación a cada parte (detallado a continuación). La configuración original también se incluye en la población inicial.
- **Mutación.** Se seleccionan los cromosomas que se van a mutar usando probabilidad p_{mut}. La mutación en cada parte se realiza aplicando una operación a una única posición $C[i]$ de cada parte del cromosoma:
  - Para C_F, el valor se genera de manera aleatoria muestreando un número real de una distribución uniforme en el rango de la variable continua.
  - Para C_R, el valor se obtiene aleatoriamente del conjunto $\{0, \cdots, k_i\} \setminus C[i]$.
  - Para C_H, el valor se obtiene aleatoriamente del conjunto $\{-1, 0, 1\} \setminus C[i]$.
  - Para C_U se obtiene como $1 - C[i]$ (negación binaria).
- **Cruce.** Se seleccionan pares de cromosomas que se cruzan con una probabilidad p_{cross}. Debido a la codificación adoptada, el cromosoma se divide en dos, y se aplican diferentes operadores a cada una de las partes:
  - Se aplica un *cruce aritmético-min-max* [10] a la parte C_F, generando cuatro hijos.
  - Se aplica un *cruce de seis puntos* al resto del cromosoma, eligiendo dos puntos de cruce en cada parte (es decir, dos puntos para C_R, dos para C_H, y dos para C_U). Este proceso de cruce genera dos hijos.

  Tras recombinar ambas partes, se generan ocho hijos, de los que se seleccionan los dos mejores para mantener constante el tamaño de población.
- **Selección.** Se utiliza una selección basada en *ranking*.
- **Reemplazo.** Se utiliza un reemplazo con elitismo (manteniendo la mejor configuración previa).
- **Criterio de Parada.** El algoritmo genético se detiene cuando el fitness del mejor individuo no ofrece suficiente mejora (de acuerdo a un umbral ϵ) durante un periodo de κ iteraciones consecutivas.

IV. Experimentos

Hemos evaluado *FLocalX* utilizando tres conocidos datasets multi-clase: *Iris*[2], *Wine*[3] y *Beer*[4]. Se utilizan datasets pequeños, ya que el objetivo de este trabajo es mostrar una metodología general de generación de explicaciones globales, no la implementación específica de la misma. Dado que los algoritmos metaheurísticos son procesos costosos tanto en recursos como en tiempo, generalmente requieren ajustes y optimizaciones específicas para cada caso y algoritmo para abordar diversos problemas. Al utilizar datasets más simples, podemos centrarnos en ilustrar las capacidades de la metodología, como trabajar con distintos tipos de explicaciones locales (i.e. FLARE, LORE) e imitar una variedad de clasificadores. Este es el primer paso en una línea de investigación de una experimentación más compleja, donde se utilizarán múltiples algoritmos metaheurísticos optimizados para esta metodología y así abordar problemas mucho más complejos. FLocalX ha sido programado en Python 3.10, utilizando librerías como `numpy` y `scikit-learn` para gestionar las estructuras de datos y generar las explicaciones eficientemente. La implementación de FLocalX está disponible en GitHub[5].

Detalles de la experimentación. Hemos adoptado las siguientes métricas para evaluar el rendimiento de FLocalX y el resto de clasificadores utilizados como referencia:

[2]https://archive.ics.uci.edu/dataset/53/iris
[3]https://archive.ics.uci.edu/dataset/109/wine
[4]https://gitlab.citius.usc.es/jose.alonso/xai
[5]https://github.com/Kaysera/flocalx. Para garantizar la reproducibilidad, todos los experimentos se han publicado también en un repositorio separado https://github.com/Kaysera/ida2024-experiments.

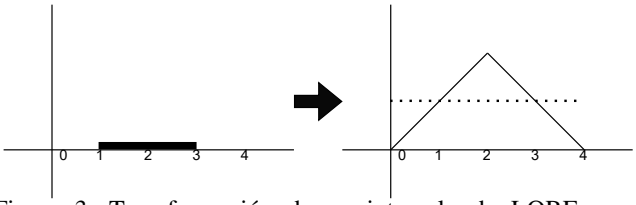


Figura 3: Transformación de un intervalo de LORE a un conjunto difuso

- *Precisión*. Medimos la precisión del clasificador original (Acc-B), de la teoría de explicación formada por la unión de las explicaciones locales (Acc-U), y de la teoría de explicación global después de aplicar FLocalX (Acc-F).
- *Fidelidad*. Mide la capacidad de la teoría de explicación global para imitar al clasificador [8]. Medimos la fidelidad de la teoría de explicación formada por la unión de las explicaciones locales (Fid-U) y la de la teoría de explicación global resultante de aplicar FLocalX (Fid-F).
- *Número de Reglas*. Mide el número total de reglas del sistema. Más reglas indican un sistema más complejo y por tanto menos interpretable. Medimos el número de reglas antes (R) y después (R_F) de aplicar FLocalX.
- *Número de premisas*. Mide el número de premisas en el antecedente de las reglas. Más premisas se perciben algunas veces (erróneamente) como más útiles [14], así que acortar las reglas junto con una comunicación correcta de la importancia de los atributos es una buena práctica. Medimos el número de premisas antes (P) y después (P_F) de aplicar FLocalX.

Se utiliza una partición entrenamiento-validación-test (60 %-30 %-10 %) para la experimentación. Con la partición de entrenamiento se entrenan los clasificadores con los hiperparámetros por defecto. Con la partición de validación se buscan los hiperparámetros de los métodos de explicación locales, y se extraen las explicaciones locales (E_L). La partición de test se ha utilizado para medir la precisión para todos los algoritmos. El algoritmo genético se ha repetido 20 veces, alterando la semilla aleatoria, y se muestra la media de los resultados. Los parámetros se han seleccionado de manera empírica. Para estos datasets, un gran tamaño de población ρ = 128 (y múltiplo de 4) muestra mejores resultados y una baja presión del tamaño α = 0.1 permite una rápida convergencia y alta precisión. Los parámetros se han fijado a valores estándar: número de conjuntos difusos $k_i = 5$; #iteraciones $\kappa = 20$; umbral de parada $\epsilon = 0.01$; p_{mut}= 0.15 y p_{cross}=0.8. Los conjuntos difusos para Iris y Wine han sido obtenidos usando particiones de igual anchura, mientras que los conjuntos difusos para Beer se han tomado de [21].

Se consideran las siguientes alternativas en la experimentación:

- *Modelos de caja negra (BB)*: Se escogen *SVM*, *Redes Neuronales (NN)* y *Random Forest (RF)* como clasificadores de referencia utilizando su implementación en *scikit-learn*.
- *Modelos basados en reglas*: Algoritmos de los que se puede extraer un conjunto de reglas y pueden ser usados para predicción y explicación. Se utilizan como sistemas de explicación globales. Los algoritmos utilizados son *Árbol de Decisión Difuso (FDT)* [18], *LORE (LO)* [8] y *FLARE (FL)* [7].
- *Métodos de transformación Local a Global*: Utilizan explicaciones locales para fusionarlas en una teoría de explicación global que es capaz de predecir y explicar instancias del dataset:
  - *FLocalX + LORE*: Se usa LORE para extraer las explicaciones locales y después aplicar FLocalX. Como FLocalX necesita reglas difusas, se generan conjuntos difusos considerando los intervalos como un α-corte de 0.5. Por ejemplo, el intervalo $[1, 3]$ se transforma en el conjunto difuso $(0, 2, 4)$ en la Figura 3.
  - *FLocalX + FLARE*: Se usa FLARE para extraer las explicaciones locales y después aplicar FLocalX.

Tabla I: Rendimiento y Explicabilidad de los distintos modelos

	Met.	BB	Fid-U	Fid-F	Acc-B	Acc-U	Acc-F	R	$\mathbf{R_F}$	P	$\mathbf{P_F}$
Iris	FDT	-	-	-	-	1.00	-	12.00	-	1.42	-
	FL	NN	**0.97**	0.93	1.00	0.95	0.91	32.00	5.05	1.31	**1.27**
		RF	0.94	0.91	0.93	0.94	0.91	26.00	**4.16**	1.81	1.70
		SVM	0.93	**0.95**	1.00	0.95	0.92	19.00	4.47	1.26	1.32
	LO	NN	0.93	0.94	1.00	0.94	**0.93**	45.00	4.79	1.96	1.55
		RF	**0.97**	0.93	0.93	**0.97**	**0.93**	45.00	**4.16**	2.07	1.85
		SVM	0.92	**0.95**	1.00	0.94	0.92	45.00	4.37	1.58	1.36
Wine	FDT	-	-	-	-	0.94	-	36.00	-	3.00	-
	FL	NN	0.86	0.76	0.89	0.82	0.77	48.00	8.68	1.42	1.54
		RF	0.61	0.61	1.00	0.61	0.61	41.00	**2.89**	1.39	**1.23**
		SVM	**0.99**	0.73	0.67	0.71	0.76	17.00	4.95	1.00	1.32
	LO	NN	0.77	0.78	0.89	0.76	0.76	54.00	4.79	2.52	1.94
		RF	0.90	**0.88**	1.00	**0.90**	**0.88**	54.00	6.21	3.19	3.02
		SVM	0.93	0.76	0.67	0.68	0.71	52.00	5.05	1.02	1.56
Beer	FDT	-	-	-	-	1.00	-	69.00	-	2.42	-
	FL	NN	0.69	0.71	0.80	0.67	0.79	128.00	20.42	1.85	**1.78**
		RF	0.87	0.88	1.00	0.87	**0.88**	129.00	26.68	2.34	2.18
		SVM	0.86	0.77	0.85	0.85	0.82	99.00	15.21	1.96	1.93
	LO	NN	0.74	0.76	0.80	0.70	0.80	119.00	**13.42**	2.01	1.98
		RF	**0.92**	**0.89**	1.00	**0.92**	**0.88**	119.00	14.63	2.54	2.60
		SVM	0.78	0.82	0.85	0.67	0.86	119.00	15.58	2.02	2.07

Resultados. Comparamos las teorías de explicación global generadas por FLocalX para los dos métodos de explicación local con la unión trivial de las explicaciones locales, y usamos un método de caja blanca basado en reglas (FDT) como referencia. La Tabla I muestra los resultados de la experimentación. En negrita se resalta el mejor resultado por base de datos para la fidelidad (**Fid-U** y **Fid-F**), la precisión (**Acc-U** y **Acc-F**) y el número de reglas ($\mathbf{R_F}$) y premisas ($\mathbf{P_F}$) después de aplicar FLocalX.

Como un objetivo del proceso de optimización es la reducción del tamaño del sistema basado en reglas, comprobar su impacto en el rendimiento es necesario. En algunos problemas con Acc-U muy alta (> 0.9), Acc-F disminuye ligeramente, posiblemente porque la mayoría de las reglas son necesarias para alcanzar ese grado de precisión. Por otro lado, en problemas más complejos donde el punto de partida no es tan bueno (el dataset Beer para FLARE y NN, o LORE

y SVM, son claros ejemplos), el algoritmo es incluso capaz de mejorar la precisión del punto de partida. Esto sugiere que una aproximación metaheurística, aunque costosa en tiempo, beneficia problemas difíciles de resolver. Finalmente, las reglas de LORE suelen ser un mejor punto de partida para FLocalX que las reglas de FLARE. Este descubrimiento sugiere que quizás más premisas proporcionan un mejor punto de partida (lo que ocurre en el caso crisp), aunque consideramos que es necesario ampliar la experimentación para explorarlo con detenimiento.

En cuanto a la complejidad de las explicaciones, podemos observar que se obtiene una reducción de entorno al $85\,\% - 90\,\%$ de las reglas de partida. Beer muestra las teorías de explicación más grandes (unas 20 reglas para FLARE y unas 14 para LORE), que aun así son legibles por humanos. Además, hay una gran reducción con respecto al clasificador de referencia, siendo necesarias un $40\,\%$ de las reglas en datasets más simples y sobre un $20\,\% - 30\,\%$ para problemas más complejos. El número de reglas generado por el FDT aumenta con la complejidad de los problemas, es decir, es válido para Iris y Wine, pero es demasiado largo (70 reglas) para Beer. Por otro lado, mirando el número de premisas, generalmente nos encontramos con 1-3 premisas por regla, muy aceptable para lectores humanos. $\#P_F$ solo es ligeramente mayor que $\#P$, ya que $f(C)$ no tiene en cuenta la longitud de la regla. Finalmente, observamos que las explicaciones globales de LORE generalmente tienen menos reglas que las de FLARE, aunque con alguna premisa más por regla.

Los resultados de esta experimentación preliminar, con un único algoritmo de optimización (el algoritmo genético) y datasets pequeños, muestran la flexibilidad de la metodología, el cual puede generar explicaciones globales compactas y fiables que pueden resultar útiles para un lector humano.

V. Conclusiones y trabajo futuro

Este trabajo introduce FLocalX, una metodología para extraer una teoría de explicación global agnóstica al modelo partiendo de una serie de explicaciones locales difusas mediante un proceso de optimización. Hemos llevado a cabo una experimentación utilizando un algoritmo genético como proceso de optimización, y observamos que FlocalX es capaz de generar teorías de explicaciones globales cortas y precisas, mejores que la unión trivial de las explicaciones locales de partida y que el clasificador de caja blanca de referencia. Las futuras líneas de trabajo pasan por un mayor estudio experimental de los hiperparámetros, operadores y funciones objetivo para el algoritmo genético. También planteamos un estudio de otros algoritmos metaheurísticos como procesos de optimización. Finalmente, la diferencia de rendimiento mostrada entre el uso de FLARE y el de LORE para generar la teoría de explicación local, da pie a experimentar con otros explicadores locales.

Agradecimientos

Este trabajo ha sido financiado por los proyectos: SBPLY/21/180225/000062 (Gobierno de Castilla-La Mancha y fondos ERDF); PID2019–106758GB–C33, and FPU19/02930 (MCIN/AEI/10.13039/501100011033 y ERDF Next Generation EU); 2022-GRIN-34437 (Universidad de Castilla-La Mancha y fondos ERDF); programa EU NextGenerationEU PNRR-PE-AI FAIR (Future Artificial Intelligence Research); PNRR-SoBigData.it - Prot. IR0000013, H2020-INFRAIA-2019-1: Res. Infr. G.A. 871042 *SoBigData++*, ERC-2018-ADG G.A. 834756 *XAI*, y CHIST-ERA-19-XAI-010 SAI.

Referencias

[1] J. M. Alonso *et al.*, "Explainable fuzzy systems: Paving the way from interpretable fuzzy systems to explainable AI systems," *SCI*, vol. 970, 2021.

[2] P. P. Angelov *et al.*, "Explainable artificial intelligence: an analytical review," *WIREs Data Mining Knowl. Discov.*, vol. 11, no. 5, 2021.

[3] A. B. Arrieta *et al.*, "Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, 2020.

[4] J. Casillas *et al.*, "Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set," *IEEE TFS*, vol. 13, no. 1, pp. 13–29, 2005.

[5] Z. Dai *et al.*, "Coatnet: Marrying convolution and attention for all data sizes," pp. 3965–3977, 2021.

[6] G. Fernández *et al.*, "Factual and counterfactual explanations in fuzzy classification trees," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 12, pp. 5484–5495, 2022.

[7] G. Fernández *et al.*, "Flare: Fuzzy local agnostic rule-based explanations for black box classifiers (submitted)," 2023, preprint available at https://dsi.uclm.es/descargas/technicalreports/DIAB-24-02-1/FLARE_Tech_Rep.pdf.

[8] R. Guidotti *et al.*, "Factual and counterfactual explanations for black box decision making," *IEEE Intelligent Systems*, vol. 34, no. 6, pp. 14–23, 2019.

[9] R. Guidotti *et al.*, "A survey of methods for explaining black box models," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 93:1–93:42, 2019.

[10] F. Herrera *et al.*, "Fuzzy connectives based crossover operators to model genetic algorithms population diversity," *Fuzzy Sets Syst.*, vol. 92, no. 1, pp. 21–30, 1997.

[11] M. Hiabu *et al.*, "Unifying local and global model explanations by functional decomposition," in *AISTATS*, vol. 206. PMLR, 2023, pp. 7040–7060.

[12] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and AI*. MIT press, 1992.

[13] T. Huber *et al.*, "Local and global explanations of agent behavior: Integrating strategy summaries with saliency maps," *Artif. Intell.*, vol. 301, p. 103571, 2021.

[14] T. Kliegr *et al.*, "A review of possible effects of cognitive biases on interpretation of rule-based machine learning models," *Artif. Intell.*, vol. 295, p. 103458, 2021.

[15] S. M. Lundberg *et al.*, "From local explanations to global understanding with explainable ai for trees," *Nature machine intelligence*, vol. 2, no. 1, pp. 56–67, 2020.

[16] G. D. P. Regulation, "General data protection regulation (GDPR)," *Intersoft Consulting, Accessed in October 24*, vol. 1, 2018.

[17] J. Schrouff *et al.*, "Best of both worlds: local and global explanations with human-understandable concepts," *CoRR*, vol. abs/2106.08641, 2021.

[18] A. Segatori *et al.*, "On distributed fuzzy decision trees for big data," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 1, pp. 174–192, 2017.

[19] M. Setzu *et al.*, "Glocalx-from local to global explanations of black box ai models," *Artificial Intelligence*, vol. 294, p. 103457, 2021.

[20] I. Stepin *et al.*, "Generation and evaluation of explanations for decision trees and fuzzy rule-based classifiers," in *FUZZ*. IEEE, 2020, pp. 1–8.

[21] I. Stepin *et al.*, "Factual and counterfactual explanation of fuzzy information granules," *Interpretable AI: A Perspective of Granular Computing*, pp. 153–185, 2021.

[22] A. K. Varshney *et al.*, "Literature review of the recent trends and applications in various fuzzy rule-based systems," *IJFS*, pp. 1–24, 2023.

[23] X. Wang, B. De Baets, and E. Kerre, "A comparative study of similarity measures," *Fuzzy sets and systems*, vol. 73, no. 2, pp. 259–268, 1995.

[24] S. Zhang *et al.*, "The diversified ensemble neural network," *Advances in Neural Information Processing Systems*, vol. 33, pp. 16 001–16 011, 2020.

Un estudio sobre el Modus Ponens generalizado para implicaciones derivadas de uninormas discretas

Isabel Aguiló*†, Pilar Fuster-Parra†, Juan Vicente Riera*†
†*Departamento de Ciencias Matemáticas e Informática*
**Grupo de investigación en Soft Computing, Procesamiento de Imágenes y Agregación (SCOPIA)*
Universidad de las Islas Baleares, 07122 Palma, España
†*Instituto de Investigación Sanitaria de las Islas Baleares (IdISBa), 07010 Palma, España*
{isabel.aguilo, pilar.fuster, jvicente.riera}@uib.es

***Resumen*—En este trabajo se investigan las funciones de implicación discreta que satisfacen el Modus Ponens con respecto a una uninorma discreta. De este estudio, y de manera análoga al caso de funciones de implicación definidas en el intervalo unidad, se obtiene que las uninormas a considerar deben de ser conjuntivas. En esta investigación, se estudiarán las funciones de implicación residuadas derivadas de una uninorma conjuntiva discreta. En particular, se analizará el caso de las derivadas de uninormas de la familia de $U_{\text{mín}}$ y de la familia de uninormas idempotentes discretas U_{ide}.**

***Palabras Clave*—Uninormas discretas, Funciones de implicación discreta, U-Modus Ponens, Reglas de inferencia**

I. Introducción

Las funciones de implicación borrosas [1], [7] son operadores clave en el control borroso y en el razonamiento aproximado, así como también en los diferentes campos donde se aplican estas teorías. Esto es debido principalmente a que estas funciones de implicación se utilizan para modelizar los condicionales borrosos de la forma "Si ... entonces", así como para modelar reglas de inferencias clásicas como es el caso del Modus Ponens y del Modus Tollens. Muchos de los estudios realizados con funciones de implicación tratan sobre propiedades adicionales que en cada caso o contexto puedan ser deseables. La mayoría de ellas provienen de tautologías de la lógica clásica que, traducidas al ámbito de la lógica borrosa, se convierten en ecuaciones funcionales. En este marco, la propiedad Modus Ponens [14] se vuelve esencial para realizar inferencias directas y es bien sabido que esta regla de inferencia está garantizada cuando la conjunción y la función de implicación utilizadas en el proceso satisfacen la desigualdad funcional correspondiente. Esta desigualdad [13], [14] se ha estudiado ampliamente para muchos tipos de funciones de implicación cuando la conjunción se modela mediante una t-norma

$$T(x, I(x, y)) \leq \text{y} \quad \text{para todo } x, y \in [0, 1]. \tag{1}$$

Isabel Aguiló y Juan Vicente Riera estan parcialmente subvencionados por el proyecto R+D+i PID2020-113870GB-I00-"Desarrollo de herramientas de Soft Computing para la Ayuda al Diagnóstico Clínico y a la Gestión de Emergencias (HESOCODICE)", financiado por MCIN/AEI/10.13039/501100011033/ P. Fuster-Parra está parcialmente subvencionada por el Ministerio de Ciencia, Innovación y Universidades con el proyecto PID2022-139248NB-I00 subvencionado por MCIN/AEI/10.13039/501100011033 y "ERDF A way of making Europe".

Sin embargo, el uso de uninormas conjuntivas para modelar conjunciones es una opción cada vez más extendida en sistemas difusos y por ello se ha investigado el Modus Ponens respecto de una uninorma conjuntiva U en lugar de una t-norma T (ver [5], [11]), dando lugar a la propiedad llamada U-Modus Ponens (o también U-condicionalidad). Lo primero que se deriva de estos trabajos es que las implicaciones usuales, como las R, (S, N), QL y D implicaciones derivadas de t-normas y t-conormas o las implicaciones de Yager, no satisfacen la U- condicionalidad. Así los candidatos posibles para satisfacer el U-Modus Ponens aparecen entre las implicaciones derivadas de uninormas. Se ha estudiado y resuelto ya para el caso de RU-implicaciones en [4], [11] y también para el caso de (U,N)-implicaciones [9].

Por otro lado, el estudio de operaciones definidas en escalas finitas, habitualmente denotadas como $L_n = \{0, \ldots, n\}$, es un área de creciente interés. Principalmente, porque permite tratar con familias finitas de etiquetas lingüísticas evitando interpretaciones numéricas (necesarias en el enfoque de lógica difusa). Las implicaciones discretas [5], [6] muestran nuevas posibilidades para el razonamiento aproximado con familias finitas de etiquetas lingüísticas, con sus consecuentes aplicaciones en la computación con palabras.

El estudio del Modus Ponens ha sido estudiado para implicaciones discretas derivadas de t-normas discretas [8], [10]. Por el contrario, no se ha realizado un estudio en el caso de uninormas discretas similar al elaborado para uninormas definidas en el intervalo unidad. En este trabajo, se analiza en profundidad esta regla del inferencia para el caso de uninormas conjuntivas discretas, que seguiremos llamando U-Modus Ponens y expresado funcionalmente como

$$U(x, I(x, y)) \leq y \quad \text{para todo } x, y \in L_n. \tag{2}$$

Análogamente al caso del intervalo unidad, se investiga en detalle el U-Modus Ponens respecto de una uninorma discreta. En particular, se estudia el caso de implicaciones residuales derivadas de uninormas de U_{min} y de la familia de uninormas idempotentes U_{ide}.

II. Preliminares

En esta sección solo recordaremos los resultados y definiciones necesarias para que el trabajo sea lo más autocontenido

posible. Trabajos sobre operadores de agregación discreta se pueden encontrar en [2], [3], [12]. Es sabido (ver [12]) que para el estudio de funciones de agregación binarias así como para el estudio de funciones de implicación todas las cadenas finitas con el mismo número de elementos son equivalentes. Por ello, utilizaremos la más simple de ellas con $n+1$ elementos:

$$L_n = \{0, 1, 2, \ldots, n\}$$

y, para todo $a, b \in L_n$ con $a \leq b$, utilizaremos la notación $[a, b]$ para denotar la subcadena dada por $[a, b] = \{x \in L_n \mid a \leq x \leq b\}$.

Definición 1: [12]

- Una función $f : L_n \to L_n$ se dice que es *suave* cuando $\mid f(x) - f(x-1) \mid \leq 1$ para todo $x \in L_n$ con $x \geq 1$.
- Una operación binaria F sobre L_n se dice que es *suave* cuando sus secciones, vertical y horizontal, lo son.

La importancia de la condición de suavidad radica en el hecho de que generalmente esta característica es usada en el caso discreto de manera equivalente a la continuidad en el intervalo $[0, 1]$, propiedad equivalente a la de divisibilidad (para una t-norma T, $x \leq y$ si y solo si existe $z \in L_n$ tal que $T(y, z) = x$), ver también [12].

Proposición 2: [12] La única negación suave (equivalentemente fuerte o estrictamente decreciente) sobre L_n es la negación clásica dada por

$$N(x) = n - x \quad \text{para todo} \quad x \in L_n.$$

Definición 3: [2], [3] Una *uninorma discreta* es una aplicación $U : L_n^2 \longrightarrow L_n$ asociativa, commutativa, creciente en cada variable y tal que existe un elemento $e \in L_n$, llamado *elemento neutro*, tal que $U(e, x) = x$ para todo $x \in L_n$.

Evidentemente, una uninorma con elemento neutro $e = n$ es una t-norma discreta y una uninorma con elemento neutro $e = 0$ es una t-conorma. Para cualquier otro valor $e \in L_n \setminus \{0, n\}$ la operación se comporta como una t-norma en $[0, e]^2$, como una t-conorma en $[e, n]^2$ y toma valores entre el mínimo y el máximo en el conjunto $A(e)$ dado por

$$A(e) = [0, e[\times]e, n] \cup]e, n] \times [0, e[.$$

Denotaremos de forma habitual una uninorma con elemento neutro e, t-norma y t-conorma subyacentes T y S, respectivamente, por $U \equiv \langle T, e, S \rangle$. Cualquier uninorma discreta satisface que $U(0, n) \in \{0, n\}$; cuando $U(n, 0) = 0$, se dice que la uninorma U es *conjuntiva*, mientras que si $U(n, 0) = n$, se dice que es *disyuntiva*. A continuación, se recuerda la estructura de las dos familias de uninormas discretas que serán utilizadas en este trabajo.

Las uninormas discretas de la familia de $U_{\text{mín}}$ han sido caracterizadas, tal como muestra el siguiente resultado.

Teorema 4: [3] Una operación binaria U de L_n con elemento neutro $0 < e < n$ es una uninorma discreta de $U_{\text{mín}}$ si y sólo si su expresión viene dada por

$$U(x, y) = \begin{cases} T(x, y) & \text{si} \quad x, y \in [0, e] \\ S(x, y) & \text{si} \quad x, y \in [e, n] \\ \text{mín}(x, y) & \text{en caso contrario,} \end{cases}$$

siendo T una t-norma discreta sobre la cadena $[0, e]$ y S es una t-conorma discreta sobre la cadena $[e, n]$.

La familia de uninormas idempotentes discretas, es decir, aquellas que verifican $U(x, x) = x$ para todo $x \in L_n$ fueron caracterizadas en [2] de la siguiente forma.

Teorema 5: [2] Una operación binaria U definida en L_n con elemento neutro $0 < e < n$ es una uninorma idempotente discreta si y sólo si existe una función decreciente $g : [0, e] \to [e, n]$ con $g(e) = e$ tal que

$$U(x, y) = \begin{cases} \text{mín}(x, y) & \text{si } y \leq \overline{g}(x) \quad \text{y} \quad x \leq \overline{g}(0) \\ \text{máx}(x, y) & \text{en otro caso,} \end{cases}$$

donde $\overline{g}$ es la única extensión de g que es simétrica con respecto a la diagonal principal y viene dada por la expresión

$$\overline{g}(x) = \begin{cases} g(x) & \text{si} \quad x \leq e \\ \text{máx}\{z \in [0, e] \mid g(z) \geq x\} & \text{si} \quad e < x \leq g(0) \\ 0 & \text{si} \quad x > g(0). \end{cases}$$

En estos casos, se dice que g es la función asociada a la uninorma idempotente U, y será denotada por $U = (e, g)$, siendo $0 < e < n$ su elemento neutro.

Recordemos que las familias de uninormas conjuntivas discretas son las uninormas de $U_{\text{mín}}$ y las uninormas idempotentes discretas tales que g verifica la condición $g(0) = n$.

Definición 6: [5] Un operador binario $I : L_n \times L_n \to L_n$ se denomina *función de implicación discreta*, o *implicación*, si satisface:

(I1) $I(x, z) \geq I(y, z)$ cuando $x \leq y$, para todo $z \in L_n$.
(I2) $I(x, y) \leq I(x, z)$ cuando $y \leq z$, para todo $x \in L_n$.
(I3) $I(0, 0) = I(n, n) = n$ y $I(n, 0) = 0$.

Notar que, a partir de la definición, se verifica que $I(0, x) = n$ y además se tiene que $I(x, n) = n$ para todo $x \in L_n$ mientras que los valores simétricos $I(x, 0)$ y $I(n, x)$ no se pueden deducir a partir de la definición.

Hay muchas otras propiedades que se exigen a las funciones de implicación dependiendo del contexto, entre ellas queremos destacar:

- El *principio de neutralidad*,

$$I(n, y) = y \quad \text{para todo} \quad y \in L_n. \qquad (NP)$$

- El *principio de identidad*,

$$I(x, x) = n \quad \text{para todo } x \in L_n. \qquad (IP)$$

- $I(x, y) = n$ si y sólo si $x < y$.

Definición 7: Sea I una función de implicación discreta. La función N_I definida como $N_I(x) = I(x, 0)$ para todo $x \in L_n$, es siempre una negación borrosa, denominada *negación natural* de I .

En la literatura [5], [6] existen diferentes métodos de construcción de funciones de implicación discreta. Entre ellas, la implicación residuada derivada de una t-norma discreta T y construidas a partir del operador

$$I_T(x, y) = \text{máx}\{z \in L_n \mid T(x, z) \leq y\}, \text{ x, y} \in L_n,$$

Una generalización de este operador se puede considerar si, en vez de utilizar una t-norma discreta, se utiliza una uninorma discreta U con elemento neutro $0 < e < n$.

Definición 8: [4] Sea $U : L_n^2 \to L_n$ una uninorma discreta, su operador residual $I_U : L_n^2. \to L_n$, vendrá dado por

$$I_U(x, y) = \text{máx}\{z \in L_n \mid U(x, z) \leq y\}, \text{ x, y} \in L_n.$$

En este caso, el operador resultante no siempre verifica las condiciones de ser una función de implicación discreta y dependerá de la elección de la uninorma.

Proposición 9: [4] Dada una uninorma discreta $U : L_n^2 \to L_n$, su operador residual I_U es una función de implicación discreta si y sólo si U es conjuntiva, esto es, si y sólo si $U(n, 0) = 0$.

Este familia de funciones de implicación han sido estudiadas y caracterizadas para el caso de uninormas discretas de $U_{\text{mín}}$ y para el caso de uninormas idempotentes discretas tales que su generador verifica la condición $g(0) = n$ (para más detalles ver [4]).

III. Modus Ponens respecto de una uninorma discreta

En esta sección vamos a investigar el Modus Ponens respecto de una uninorma U discreta, o simplemente el U-Modus Ponens, para una función de implicación discreta I. Para ello, empecemos dando la definición formal del U-Modus Ponens.

Definición 10: Sea I una implicación borrosa discreta y U una uninorma discreta. Diremos que I verifica el Modus Ponens respecto de U (o simplemente el U-Modus Ponens), o también que I es un U-condicional si

$$U(x, I(x, y)) \leq y \text{ para todo } x, y \in L_n. \tag{3}$$

Un primer paso será investigar qué condiciones debe de satisfacer la función de implicación discreta I así como la uninorma discreta U para que I sea U-condicional.

De manera semejante al caso de funciones de implicación y uninormas definidas en el intervalo unidad [11], se tienen los siguientes resultados.

Lema 11: Sea I una función de implicación discreta y U una uninorma discreta. Si I es U-condicional entonces U es conjuntiva.

El resultado anterior, nos dice qué tipo de uninormas discretas deberemos considerar. A continuación se estudiarán algunas propiedades necesarias que deberán verificar la uninormas conjuntivas así como las funciones de implicación discreta.

Proposición 12: Consideremos una uninorma conjuntiva discreta U con elemento neutro $0 < e < n$ y supongamos que I es una función de implicación discreta tal que I es U-condicional. En estas condiciones, se verifican las siguientes propiedades:

1. $I(e, y) \leq y$ para todo $y \in L_n$.
2. U verifica $U(x, N_I(x)) = 0$ para todo $x \in L_n$.
3. La negación natural N_I verifica $N_I(x) = 0$ para todo $x \geq e$ y $N_I(x) < e$ para todo $0 < x < e$. En particular, N_I no puede ser suave.
4. $I(x, y) < e$ para todo $x > y \geq e$. En particular, $I(n, y) < e$ para todo $y < n$.
5. $U(n, I(n, y)) \leq y$ para todo $y \in L_n$.

Este último resultado nos da condiciones necesarias que deberá cumplir la uninorma U así como la función de implicación I para que se verifique que I sea U-condicional. En primer lugar, notar que la propiedad 4 es incompatible con la propiedad $I(n, y) = y$ para todo $y \in L_n$, propiedad que verifican entre otras, las cuatro clases más usuales de funciones de implicación discreta, a saber, R, (S, N), QL o D implicaciones (ver por ejemplo [6], [8]). Sin embargo, no es el caso de las implicaciones residuadas derivadas de uninormas discretas ya que satisfacen $I_U(e, y) = y$ para todo $y \in L_n$ (ver proposición 2 en [4]). Por ello, y a partir de ahora, toda nuestra investigación se centrará en considerar este tipo de funciones de implicación.

Un primer resultado que se obtiene del hecho de considerar funciones de implicación derivadas de uninormas conjuntivas discretas es la relación existente entre sus elementos neutros.

Proposición 13: Sean U, U_0 dos uninormas discretas conjuntivas cuyos elementos neutros son $e, e_0 \in L_n \setminus \{0, n\}$ respectivamente y consideremos I_{U_0} la implicación residuada discreta derivada de U_0. Si I_{U_0} es U-condicional entonces $e_0 \leq e$.

Por otra parte, sabemos (ver proposición 2 en [4]) que si U es una uninorma conjuntiva discreta, se tiene que I_U siempre es U-condicional. Por ello, se tiene el siguiente resultado.

Proposición 14: Sean U, U_0 dos uninormas conjuntivas discretas cuyos elementos neutros son $e, e_0 \in L_n \setminus \{0, n\}$ respectivamente. Si $U \leq U_0$ entonces se tiene que I_{U_0} es U-condicional.

Ejemplo 15: Consideremos U la menor uninorma con elemento neutro $e \in L_n \setminus \{0, n\}$, que puede verse representada en la figura 1, donde T_D representa la t-norma drástica en la región $[0, e] \times [0, e]$. Entonces, I_{U_0} es U-condicional para cualquier uninorma conjuntiva U_0 con el mismo elemento neutro e ya que $U \leq U_0$.

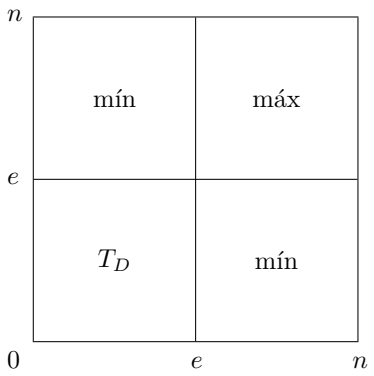


Figura 1. Estructura de la uninorma más pequeña con elemento neutro $e \in L_n \setminus \{0, n\}$.

En la siguiente sección se estudiará el caso particular de implicaciones residuadas derivadas de una uninorma discreta de la familia de $U_{\text{mín}}$.

IV. U-Modus Ponens para implicaciones residuales obtenidas de uninormas de $U_{\text{mín}}$

Consideremos una uninorma $U_0 \in U_{\text{mín}}$, y su implicación residual I_{U_0}; queremos estudiar cuando I_{U_0} es U-condicional con respecto a una uninorma conjuntiva U. Para ello, recordemos la estructura general de una implicación residual derivada de una uninorma discreta de $U_{\text{mín}}$.

Proposición 16: [4] Sea $U_0 \equiv< T_0, e_0, S_0 >_{\text{mín}}$ una uninorma discreta de $U_{\text{mín}}$ con elemento neutro $e_0 \in (0, n)$. Entonces su operador residual I_{U_0} es siempre una función de implicación y ésta vendrá dada por

$$I_{U_0}(x,y) = \begin{cases} n & \text{si} \quad 0 \le x < e_0 \text{ con } x \le y \\ I_T(x,y) & \text{si} \quad 0 \le x \le e_0 \text{ con } x > y \\ y & \text{si} \quad y < e_0 < x \\ e_0 - 1 & \text{si} \quad e_0 \le y < x \\ I_S(x,y) & \text{si} \quad e_0 \le x \le n, e_0 \le y \le n \\ & \text{con } x \le y. \end{cases}$$

Para esta clase de RU-implicaciones, el siguiente resultado nos da condiciones necesarias para que I_{U_0} sea U-condicional.

Proposición 17: Sea U una uninorma discreta con elemento neutro $e \in L_n \setminus \{0, n\}$ y $U_0 \equiv< T_0, e_0, S_0 >_{\text{mín}}$ una uninorma discreta de $U_{\text{mín}}$ con elemento neutro e_0 tal que $e_0 < e$. Si la RU-implicación I_{U_0} es U-condicional entonces se verifica $U(x,y) = \text{mín}(x,y)$ para todo $x, y \in R_0$ donde la región R_0 viene dada por

$$R_0 = [0, e_0] \times [e, n] \cup [e, n] \times [0, e_0].$$

Imponiendo algunas restricciones sobre las condiciones que deben de cumplir las componentes de las correspondientes uninormas U y U_0, se obtienen condiciones necesarias y suficientes, tal como se puede ver en el siguiente resultado.

Proposición 18: Sea $U \equiv< T_U, e, S_U >$ una uninorma discreta con elemento neutro $0 < e < n$ y sea $U_0 \equiv< T_0, e_0, S_0 >_{\text{mín}}$ una uninorma de $U_{\text{mín}}$ con elemento neutro e_0 tal que $e_0 < e$. Supongamos que $T_0 = \text{mín}$ y que $S_U = \text{máx}$. Entonces la RU-implicación I_{U_0} es U-condicional si y solo si $U(x,y) = \text{mín}(x,y)$ para todo $x, y \in R_0 = [0, e_0] \times [e, n] \cup [e, n] \times [0, e_0]$.

Ejemplo 19: Consideremos $U \equiv< T_U, e, \text{máx} >$ y sea $U_0 \equiv< \text{mín}, e_0, \text{máx} >_{\text{mín}}$, siendo $e_0 < e$. En este caso, y de acuerdo a la proposición anterior I_{U_0} es U-condicional. Se puede ver la estructura de la uninorma U así como la estructura de la función de implicación I_{U_0} en la figura 2.

Finalmente, si imponemos condiciones sobre los elementos neutros tendremos:

Teorema 20: Sea $U \equiv< T_U, e, S_U >$ una uninorma discreta con elemento neutro $0 < e \le n$ y sea $U_0 \equiv< T_0, e_0, S_0 >_{\text{mín}}$ una uninorma de $U_{\text{mín}}$ con elemento neutro e_0 tal que $e_0 < e$. Supongamos que $U(e_0, e_0) = e_0$ (en este caso, U viene dada por una t-norm T_1 en la región $[0, e_0]^2$) y que $U_0(e,e) = e$ (en este caso, U_0 viene dada por una t-conorma , S_1 en la región $[e, n]^2$).

Entoces, la RU-implicación I_{U_0} es U condicional si y solo si se verican las siguientes condiciones:

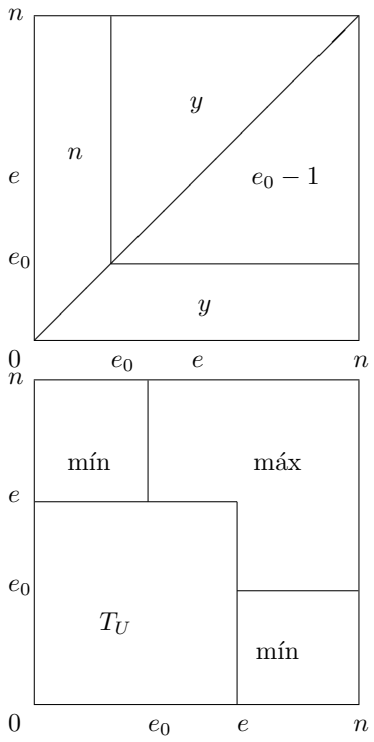


Figura 2. Estructura de la uninorma de U (abajo) y de la función de implicación I_{U_0} (arriba).

i) $U(x,y) = \text{mín}(x,y)$ para todo $x, y \in R_0$ siendo $R_0 = [0, e_0] \times [e, n] \cup [e, n] \times [0, e_0]$.

ii) La implicación residual derivada de la t-norma T_0, I_{T_0}, verifica

$$T_1(x, I_{T_0}(x,y)) \le y \text{ para todo } y < x \le e_0.$$

iii) $S_U(x, R_{S_1}(x,y)) \le y$ para todo $x \le y$.

V. U-Modus Ponens para implicaciones residuales obtenidas de uninormas idempotentes

En esta sección se estudiará el U-Modus Ponens para funciones de implicación residuadas derivadas de uninormas idempotentes discretas. Para ello, recordemos la estructura general de una implicación residual derivada de una uninorma idempotente.

Proposición 21: [2] Sea $U_0 \equiv< g_0, e_0 >_{ide}$ una uninorma idempotente con elemento neutro $e_0 \in (0, n)$ tal que $g_0(0) = n$. Entonces su operador residual I_{U_0} es una función de implicación y vendrá dada por

$$I_{U_0}(x,y) = \begin{cases} \text{máx}(\overline{g}_0(x), y) & \text{si} \quad x \le y \\ \text{mín}(\overline{g}_0(x), y) & \text{si} \quad x > y, \end{cases}$$

o, equivalentemente,

$$I_{U_0}(x,y) = \begin{cases} g_0(x) & \text{si} \quad x < e_0 \quad \text{con } x \le y < g_0(x) \\ \overline{g}_0(x) & \text{si} \quad x > e_0 \quad \text{con } \overline{g}_0(x) \le y < x \\ y & \text{en cualquier otro caso.} \end{cases}$$

Un primer resultado que nos da condiciones necesarias es el siguiente.

Proposición 22: Sea U una uninorma conjuntiva discreta con elemento neutro $e \in L_n \setminus \{0, n\}$, $U_0 \equiv <g_0, e_0>_{ide}$ una uninorma idempotente discreta con elemento neutro e_0, siendo $(0 < e_0 \leq e)$ y sea I_{U_0} la implicación residual derivada de U_0. Si I_{U_0} es U-condicional entonces la t-conorma subyacente de U debe ser $S_U = \text{máx}$.

Imponiendo la condición $\overline{g}_0(e) = 0$, se tiene la siguiente caracterización.

Proposición 23: Sea U una uninorma conjuntiva discreta con elemento neutro $e \in L_n \setminus \{0, n\}$, $U_0 \equiv <g_0, e_0>_{ide}$ una uninorma idempotente con elemento neutro $0 < e_0 < e$ tal que $\overline{g}_0(e) = 0$. Sea I_{U_0} la implicación residual derivada de U_0. Entonces, I_{U_0} es U-condicional si y solo si la t-conorma subyacente de U es $S_U = \text{máx}$.

Ejemplo 24: Si consideramos la única uninorma discreta U_0 de la familia de $U_{\text{mín}}$ que es idempotente con elemento neutro e_0, sabemos que su implicación residual viene dada por

$$I_{U_0}(x, y) = \begin{cases} n & \text{si} \quad 0 \leq x < e_0 \text{ con } x \leq y \\ e_0 - 1 & \text{si} \quad e_0 \leq y < x \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Si consideramos cualquier uninorma discreta U con elemento neutro $e > e_0$ que tenga como estructura la representada en la figura 3, de acuerdo al resultado anterior, I_{U_0} es U-condicional.

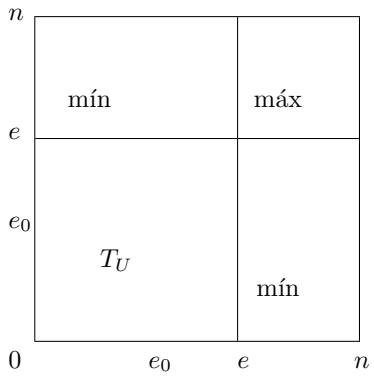


Figura 3. Estructura de la uninorma de U con elemento neutro e.

VI. Conclusión

En este trabajo se ha investigado, una de las reglas de inferencia más importantes, el Modus Ponens, también llamada t-condicionadad dada en la expresión (1). Para ello, se ha sustituido la t-norma discreta T por una uninorma discreta U, dando lugar, a una generalización de la misma, que hemos llamado, U-condicionalidad (de manera semejante al estudio hecho en el invervalo unidad) y que viene dada por la fórmula 2. Se ha demostrado, que dicho operador necesariamente debe de ser conjuntivo y que las funciones de implicación discretas a considerar no pertenenecen a las familias más clásicas, como es el caso de las R, (S,N), D o Q-implicaciones. Por ello, se han estudiado las implicaciones residuadas derivadas de uninormas conjuntivas discretas haciendo especial incapié para el caso de dos familias importantes de las mismas, las obtenidas a partir de uninormas discretas de la familia de $U_{\text{mín}}$ y de la familia de las uninormas idempotentes. Como trabajo futuro, queremos investigar caracterizaciones para el caso de la familia de uninormas idempotentes, así como investigar dicha propiedad, para el caso particular, en que las t-normas y t-conormas subyacentes sean suaves

Agradecimientos

Isabel Aguiló y Juan Vicente Riera están parcialmente subvencionados por el proyecto R+D+i PID2020-113870GB-I00-"Desarrollo de herramientas de Soft Computing para la Ayuda al Diagnóstico Clínico y a la Gestión de Emergencias (HESOCODICE)", financiado por MCIN/AEI/10.13039/501100011033/ P. Fuster-Parra está parcialmente subvencionada por el Ministerio de Ciencia, Innovación y Universidades con el proyecto PID2022-139248NB-I00 subvencionado por MCIN/AEI/10.13039/501100011033 y "ERDF A way of making Europe".

Referencias

[1] M. Baczynski, B. Jayaram, Fuzzy Implications. Studies in Fuzziness and Soft Computing, vol. 231. Springer, Berlin Heidelberg, 2008.

[2] B. De Baets, J. Fodor, D. Ruiz-Aguilera and J. Torrens, "Idempotent uninorms on finite ordinal scales," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 17 pp. 1–14, 2009.

[3] M. Mas, G. Mayor and J. Torrens, "t-Operators and uninorms on a finite totally ordered sets," J. Intell. Systems, vol.14, pp. 909–922 (1999).

[4] M. Mas, G. Mayor, M. Monserrat and J. Torrens, "Residual Implications from Discrete Uninorms. A Characterization," in L. Magdalena et. (Eds.), Enric Trillas: A Passion for Fuzzy Sets, Studies in Fuzzines and Soft Computing, Springer 2015, pp. 27–40.

[5] M. Mas, M. Monserrat and J. Torrens, "S-implications and R-implications on a finite chain," Kybernetika, vol.40, pp. 3–20, (2004).

[6] M. Mas, M. Monserrat and J.Torrens," On two types of discrete implications," International Journal of Approximate Reasoning, vol. 40, pp. 262–279, 2005.

[7] M. Mas, M. Monserrat, J. Torrens, E. Trillas, "A survey on fuzzy implication functions," IEEE Transactions on Fuzzy Systems, 15(6), 1107–1121, 2007.

[8] M. Mas, M. Monserrat and J. Torrens, "Modus ponens i modus tollens in discrete implications," International Journal jof Approximate Reasoning, vol.49, pp. 422–435, 2008.

[9] M. Mas, M. Monserrat, D. Ruiz-Aguilera, J. Torrens "RU and (U,N)-implications satisfying Modus Ponens," International Journal of Approximate Reasoning, 73, 123-137, 2016.

[10] M. Mas, D. Ruiz-Aguilera, J. Torrens, "On a generalization of the Modus Ponens: U-conditionality," in Proceedings of IPMU-2016, Part I, CCIS 610, J.P. Carvalho et al. Eds. 2016, pp. 1–12.

[11] M. Mas, D. Ruiz and J. Torrens, "Uninorm based residual implications satisfying the Modus Ponens property with respect to a uninorm, Fuzzy Sets and Systems, vol. 359, pp. 22–41, 2019.

[12] G. Mayor and J. Torrens, Triangular norms in discrete settings, in: E.P. Klement and R. Mesiar (Eds.), *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*. Elsevier, Amsterdam, 2005, pp. 189–230.

[13] E. Trillas, M. Mas, M. Monserrat and J. Torrens, "On the representation of fuzzy rules, International Journal of Approximate Reasoning," vol. 48, pp. 583–597, 2008.

[14] E. Trillas, L. Valverde, "On Modus Ponens in fuzzy logic," in 15th International Symposium on Multiple-Valued Logic, pp. 294–301. Kingston, Canada, 1985.

Sistema Difuso TSK Escalable

Javier Martín Moreno
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
javier.martin@dti.uhu.es

Francisco Alfredo Márquez Hernández
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
alfredo.marquez@dti.uhu.es

Francisco José Moreno Velo
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
francisco.moreno@dti.uhu.es

Antonio Peregrín Rubio
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
peregrin@dti.uhu.es

***Resumen*—En la actualidad, el avance tecnológico ha derivado en la generación masiva de datos, impulsada por la expansión de Internet y la presencia generalizada de sensores, donde, además, las personas contribuyen de manera activa a esta generación de datos mediante el uso de dispositivos móviles y redes sociales. Ante la dificultad que enfrentan las técnicas convencionales de Inteligencia Artificial al procesar volúmenes de datos tan elevados, se recurre a las tecnologías de Big Data para abordar estos desafíos de manera escalable. Este trabajo se centra específicamente en el desarrollo de un Sistema Basado en Reglas Difusas de tipo TSK, que sea escalable, relativamente sencillo y fácilmente reproducible.**

Palabras clave—Big Data; Spark; Sistema Basado en Reglas Difusas TSK; Algoritmo Evolutivo; RPROP

I. INTRODUCCIÓN

Los Sistemas Basados en Reglas Difusas (SBRD) han sido ampliamente utilizados para resolver problemas que requieren modelar el conocimiento de una manera similar a cómo lo percibe el ser humano. Este enfoque permite procesar correctamente información incompleta y/o imprecisa, particularmente en entornos dinámicos de incertidumbre [1].

El proceso de aprendizaje de reglas a partir de un conjunto de ejemplos conlleva un coste computacional proporcional a la magnitud de dicho conjunto. Los algoritmos utilizados para esta labor encuentran limitaciones en la cardinalidad del problema. En respuesta a esta problemática, surge la necesidad de aplicar técnicas de Big Data con el fin de desarrollar un modelo paralelo y distribuido que sea escalable, abordando así eficazmente las demandas computacionales inherentes en los grandes conjuntos de datos [2].

Diversas propuestas se han llevado a cabo en el ámbito de los SBRD escalables. Inicialmente, surgieron iniciativas en el campo del agrupamiento difuso, destacando las contribuciones de T. C. Havens et al., quienes proponen una implementación del algoritmo C-Means diseñada para procesar grandes volúmenes de datos [3]. Por otro lado, S. del Río et al. proponen un clasificador difuso basado en el algoritmo Chi, enfocado en resolver problemas de clasificación difusa en estos grandes conjuntos de datos [4]. En cuanto al dominio de la regresión difusa, cabe resaltar entre otros el trabajo de I. Rodríguez-Fernández et al. [5], quienes proponen un sofisticado SBRD evolutivo para resolver problemas de regresión de gran escala.

La complejidad computacional asociada al diseño de SBRDs reside en la generación y ajuste de la Base de Conocimiento. En la literatura podremos encontrar diversas metodologías para abordar este desafío. Estos métodos se clasifican en:

- Métodos evolutivos (Genetic Fuzzy Systems - GFS): Estos enfoques, comúnmente conocidos como sistemas difusos genéticos, han sido ampliamente explorados [6]. La utilización de algoritmos bioinspirados para llevar a cabo la búsqueda de parámetros presenta como ventaja principal su capacidad de exploración, si bien es cierto que requieren más tiempo de ejecución.
- Métodos basados en la teoría de aprendizaje de redes neuronales (Neuro Fuzzy Systems – NFS): Estos métodos, habitualmente denominados sistemas neurodifusos, han destacado en la literatura [7]. La aplicación de algoritmos basados en búsquedas de parámetros guiadas por el descenso del gradiente ofrece soluciones efectivas de manera más rápida, aunque ocasionalmente pueden quedar atrapados en mínimos locales.

En este trabajo se propone un SBRD TSK escalable y distribuido llamado TSKEvoRPROP, diseñado específicamente para abordar problemas de regresión donde la precisión del modelo desempeña un papel fundamental. El objetivo principal de este modelo es garantizar su reproducibilidad, siguiendo una filosofía sencilla que facilite tanto su comprensión como su implementación práctica. Además, se asegura la escalabilidad a medida que aumenta el volumen de las instancias procesadas en la etapa de aprendizaje y ajuste del SBRD, apoyándose en técnicas de Big Data para lograrlo, como son el almacenamiento distribuido del conjunto de datos en

particiones de menor tamaño y el procesamiento en paralelo de cada subconjunto de datos.

Este trabajo sigue la siguiente estructura: La Sección II aborda los modelos TSK desde la perspectiva de la escalabilidad. En la Sección III se define el modelo TSK escalable propuesto. La Sección IV se centra en el análisis de los resultados obtenidos. Finalmente, en la Sección V se presentan las conclusiones principales de este trabajo.

II. Modelos TSK y la Escalabilidad

En esta sección, se presentan de forma concisa las características de los SBRD TSK. Posteriormente, se lleva a cabo un breve análisis de la literatura en el ámbito de la escalabilidad de estos modelos.

Como es bien sabido, los SBRD TSK emplean reglas con antecedentes representados por etiquetas lingüísticas y consecuentes expresados como polinomios en función de las variables de entrada. La estructura común de las reglas TSK es la siguiente:

$$R_j\text{: Si } X_1 \text{ es } A_1 \text{ y } \ldots \text{ y } X_n \text{ es } A_n\text{, entonces}$$
$$Z = p_1 \cdot X_1 + \ldots + p_n \cdot X_n + p_0$$

Donde X representa el conjunto de entradas al sistema compuesto por n variables, A denota el conjunto de etiquetas lingüísticas que componen el antecedente de la regla, Z es el resultado del consecuente de la regla y p constituye el conjunto de componentes del polinomio del consecuente, siendo p_0 el término independiente. El subíndice j refiere el número de regla, y el i es el número de variable.

Dada la entrada X, el grado de pertenencia de X_i a su correspondiente etiqueta A_i se calcula usando el método de fuzzificación no puntual, expresado como $h_i = f(A_i, X_i)$, siendo h_i un valor decimal entre 0 y 1.

$$h_i = f(A_i, X_i) \tag{1}$$

El grado de activación de la regla R_j dada una entrada X se representa como $H_j = \mu_A(X)$, siendo habitual el operador de conjunción Mínimo.

$$H_j = min(h_1, \ldots, h_n) \tag{2}$$

A continuación, se calcula la salida de la regla R_j.

$$Z_j = \sum(X_i \cdot p_i) + p_0 \tag{3}$$

Finalmente, el resultado Y' generado por el sistema se calcula mediante la media ponderada de los consecuentes, multiplicados por el grado de activación de cada regla:

$$Y' = \sum Z_j \cdot H_j \,/\, \sum H_j \tag{4}$$

En sistemas de este tipo, el aprendizaje de los consecuentes se realizará necesariamente mediante algoritmos. Asimismo, los antecedentes se generan mediante métodos guiados por cubrimiento de ejemplos y se ajustan posteriormente. Para ajustar estos parámetros, se debe minimizar el error cuadrático medio de las predicciones del modelo sobre el *dataset*.

Debido a la complejidad computacional inherente a los grandes conjuntos de datos, se ha abordado previamente la problemática desde una perspectiva escalable. Inicialmente, antes de la aparición de las tecnologías Big Data, se recurrió a técnicas de preprocesamiento como la reducción del conjunto de datos para hacer frente a este desafío. En la literatura se destacan trabajos como los de M. J. Gacto et al. quienes proponen un método para modelar un sistema difuso preciso capaz de abordar problemas de regresión de gran magnitud denominado METSK-HD^e [8]. La contribución principal de este trabajo radica en la integración el filtro de Kalman con el ajuste del SBRD TSK basado en dispersión, lo que resulta en una mejora significativa de la eficiencia del algoritmo.

No obstante, en épocas más recientes con la aparición de tecnologías Big Data, ha surgido la posibilidad de abordar estos desafíos procesando estas grandes cantidades de datos mediante la escalabilidad horizontal de la capacidad de cómputo. Este nuevo paradigma permite resolver problemas de manera eficiente sin recurrir a técnicas de selección ni dispersión, las cuales pueden resultar en pérdidas de información no deseadas. En este contexto, cabe destacar la contribución de I. Rodríguez-Fernández et al. quienes proponen un sistema difuso genético distribuido y escalable para resolver problemas de regresión, denominado S-FRULER [5]. La principal aportación de este trabajo se encuentra en la subdivisión del problema en particiones más reducidas que se procesan de manera distribuida, ofreciendo así una solución escalable. Sin embargo, es importante señalar que estas soluciones suelen incorporar diversas etapas que integran múltiples técnicas de optimización, haciendo el modelo muy complejo y de difícil reproducibilidad.

En este trabajo se propone una implementación sencilla y más fácilmente reproducible de un SBRD TSK basado en tecnología Big Data para resolver problemas de gran magnitud.

Para abordar el desafío de la optimización paramétrica asociada al ajuste de los antecedentes y los consecuentes de la Base de Reglas, se ha optado por un algoritmo determinista que emplea modelos matemáticos basados en la teoría del descenso por el gradiente. De esta manera, se busca mejorar iterativamente la precisión del SBRD mediante la modificación de los parámetros en la dirección óptima de la derivada del error cuadrático medio.

En cuanto a la escalabilidad del modelo, se ha logrado mediante la adaptación de este al paradigma de procesamiento en paralelo sobre datos distribuidos. El modelo propuesto exhibe la capacidad de ejecutarse en múltiples nodos de cómputo, llevando a cabo cada una de estas tareas en paralelo sobre particiones individuales de los datos. Con el objetivo de preservar la consistencia del modelo distribuido, su diseño se ha concebido de tal manera que el resultado final de la ejecución del algoritmo sea idéntico, ya sea ejecutándolo en paralelo sobre varios nodos del clúster o sobre un solo nodo de forma secuencial. Esta consistencia se asegura al optar por la paralelización de las etapas más costosas, realizando cálculos parciales que se agregarán posteriormente para mantener la integridad de los resultados del algoritmo original.

III. Modelo Propuesto: TSKEvoRPROP

En esta sección se presenta TSKEvoRPROP, un SBRD TSK-1 escalable que permite resolver problemas de regresión de gran magnitud. A grandes rasgos, el modelo se compone de tres etapas secuenciales. En primer lugar, un algoritmo evolutivo aborda la búsqueda de la granularidad de cada una de las variables de entrada, dividiendo el universo de discurso y generando así la Base de Datos inicial. A continuación, se genera la Base de Reglas inicial mediante un algoritmo de cubrimiento de ejemplos. Posteriormente, esta Base de Reglas se ajusta de manera iterativa mediante técnicas de descenso por gradiente, tratando de minimizar el error cuadrático medio de las predicciones del SBRD sobre el conjunto de datos. Finalmente, se obtiene como resultado la Base de Reglas ajustada que mejor rendimiento ha ofrecido. La Figura 1 ilustra el diagrama del modelo propuesto.

A continuación, se profundiza en los detalles de cada una de las fases del modelo.

A. *Algoritmo evolutivo para escoger granularidad*

En particular, se ha seleccionado el algoritmo genético CHC [9] como técnica evolutiva para determinar la granularidad del SBRD. Este algoritmo destaca por su balance entre la exploración y la explotación del espacio de búsqueda. Para lograrlo, combina técnicas de elitismo, que enfocan la búsqueda en las zonas prometedoras mediante presión selectiva, con técnicas de prevención de incesto, que evitan la convergencia prematura, y con técnicas de rearranque, que diversifican las áreas de búsqueda.

Para identificar la mejor granularidad, este algoritmo utiliza cromosomas con tantos genes como variables de entrada tenga el problema. Los valores de los genes están comprendidos en el rango de granularidades de 3 a 7, y se añade el valor 0 para aplicar una selección de variables implícita en la búsqueda de la mejor granularidad para cada variable. Dado que los valores de los genes son enteros, se ha optado por un cruce HUX, asegurando que cada cromosoma hijo contenga un gen con el valor de cada padre. De este cruce se generan dos hijos, cada uno con un gen de cada padre. La prevención de incesto consiste en que los padres que no superen cierto umbral de cruce, no se cruzan, manteniendo así la diversidad genética a lo largo de las generaciones. Nótese que, al tratarse de un espacio de búsqueda reducido, el coste computacional será bajo.

B. *Generación de la Base de Reglas por cubrimiento*

Después de obtener la Base de Datos inicial a partir de la granularidad, se utilizan estas etiquetas para generar reglas cuyos antecedentes cubran cada ejemplo del conjunto de datos. La técnica de cubrimiento seleccionada es la clásica propuesta por Wang y Mendel [10].

Este algoritmo está orientado a generar la base de reglas a partir de una base de datos predefinida. Basándose en criterios de cobertura, genera una regla que cubra cada ejemplo dado. Se calcula a qué término lingüístico pertenece cada parámetro del ejemplo de entrada, generando así el antecedente de la regla que lo cubre. Una vez generada una regla que cubra a cada ejemplo, se eliminan las reglas duplicadas, obteniendo así la base de reglas inicial por cubrimiento de los ejemplos dados.

C. *Aprendizaje de consecuentes*

En el caso de los sistemas TSK-1, el algoritmo de Wang y Mendel no puede calcular el consecuente, ya que fue concebido para generar sistemas tipo Mamdani, donde el consecuente también se define mediante una etiqueta lingüística. Por lo tanto, para la optimización paramétrica se ha optado por un algoritmo heurístico de descenso por gradiente, en lugar de otras técnicas conocidas como los algoritmos evolutivos. Aunque más complejo, se busca potenciar la velocidad y la precisión del modelo aprendido mediante el uso del algoritmo matemático Resilient Propagation (RPROP) [11]. Se trata de una técnica de aprendizaje supervisado, donde a partir de un conjunto de datos etiquetados, se ajustan los valores de los parámetros del sistema para obtener el resultado deseado.

A los sistemas difusos que hacen uso de técnicas de aprendizaje se les denomina sistemas neuro-difusos debido a la relación demostrada entre las redes neuronales y los sistemas difusos. En este caso nos enfocamos en los algoritmos de descenso por gradiente, los cuales utilizan el gradiente de la función de error para encontrar su mínimo local. Para aplicarlos al aprendizaje de reglas consideraremos cada configuración de los parámetros a ajustar, en este caso el conjunto de los parámetros de los polinomios de los consecuentes, como un punto en el espacio de búsqueda. La Figura 2 muestra gráficamente cómo el algoritmo RPROP ajusta los parámetros sobre la función de error en tan sólo 22 iteraciones, mientras el algoritmo de descenso por gradiente original necesita 598 iteraciones.

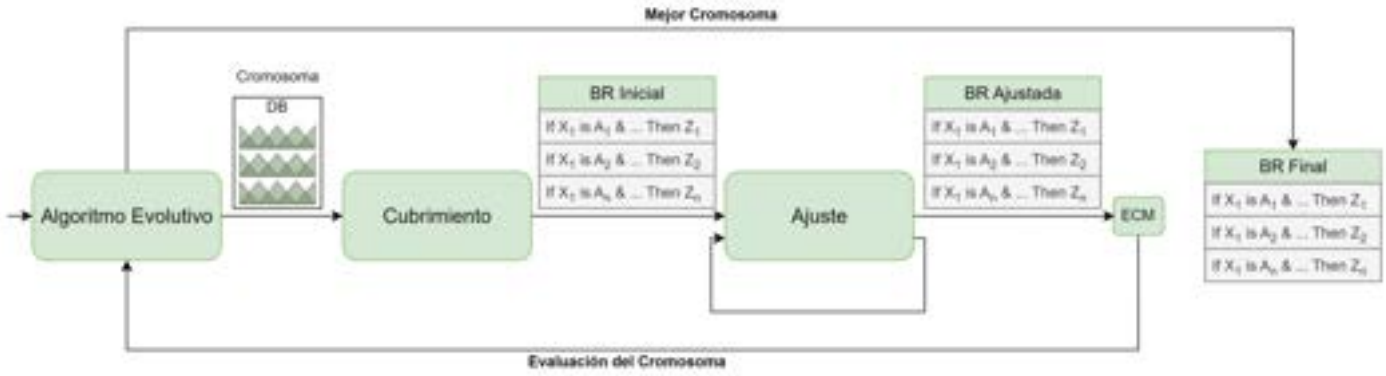


Fig. 1. Diagrama TSKEvoRPROP

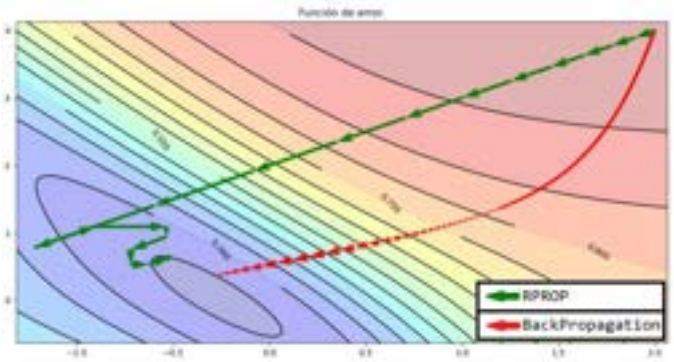


Fig. 2. Función de error: RPROP vs BackPropagation

El algoritmo RPROP soluciona los problemas fundamentales del algoritmo backpropagation clásico, acelerando su convergencia. Por un lado, utiliza un factor de aprendizaje dinámico en lugar de constante, y por otro, emplea un factor de aprendizaje individual para cada parámetro, permitiendo que cada uno se actualice según sus necesidades específicas.

Este algoritmo se fundamenta en el algoritmo Manhattan para realizar la actualización de los parámetros. En lugar de usar el valor de la derivada, se emplea únicamente su signo, mejorando así el comportamiento al acercarse al mínimo. La heurística utilizada para acelerar la convergencia del algoritmo implica considerar el valor de las derivadas de iteraciones sucesivas al realizar la actualización de cada parámetro. De esta manera, cada parámetro mantiene su propio factor de aprendizaje e incorpora información de segundo orden. Por tanto, para calcular los pasos de actualización de cada parámetro se usa la siguiente ecuación, que describe el desplazamiento del parámetro en función de su evolución con respecto al instante anterior:

$$\Delta_p(t) \begin{cases} \Delta_p(t-1)\cdot\sigma^+ \; si \; \partial E/\partial p_{\,t-1} \cdot \partial E/\partial p_{\,t} > 0 \\ \Delta_p(t-1)\cdot\sigma^- \; si \; \partial E/\partial p_{\,t-1} \cdot \partial E/\partial p_{\,t} < 0 \\ \Delta_p(t-1) \quad si \; \partial E/\partial p_{\,t-1} \cdot \partial E/\partial p_{\,t} = 0 \end{cases} \tag{5}$$

A continuación, para actualizar los parámetros se utiliza la siguiente ecuación, que modifica el parámetro utilizando sólo el signo de la derivada del error respecto al parámetro:

$$p_{ij}^{(t+1)} = p_{ij}^{(t)} - signo\,(\,\partial E/\partial p_{\,t}) \cdot \Delta_{ij}^{(t)} \tag{6}$$

Por tanto, el algoritmo RPROP se fundamenta en la derivada de la función de error respecto de cada parámetro para calcular el desplazamiento de cada uno de ellos. Debido a la estructura interna de los SBRD, es necesario usar la regla de la cadena para calcular las derivadas de cada parámetro. Para calcular el gradiente de la función de error, es necesario seguir de manera inversa los cálculos realizados por el sistema difuso en el proceso de inferencia, partiendo de la función del error cuadrático medio. En primer lugar, se define la derivada de la función de error cuadrático medio del SBRD respecto de cada parámetro como:

$$\partial E/\partial p = \sum \partial e_i/\partial p \tag{7}$$

Luego, se calcula la derivada del error del SBRD para cada ejemplo respecto a las funciones de pertenencia y al grado de activación de cada regla:

$$\partial e_i/\partial p = 2 \cdot (f(p,x)-y)\cdot\partial f(p,x)/\partial p \tag{8}$$

A partir del grado de activación se puede calcular la derivada respecto al valor de activación de cada variable de entrada utilizando las derivadas de los operadores difusos. Por último, se calculan las derivadas del grado de activación de cada variable respecto de cada parámetro. Por lo tanto, para calcular las derivadas de la función de error respecto a cada parámetro será necesario derivar los operadores difusos y las funciones de pertenencia que componen el proceso de inferencia.

D. Tunning de antecedentes

Además del aprendizaje de los consecuentes, se ha empleado RPROP para llevar a cabo el ajuste o tunning de los antecedentes usando el modelo 3-tuplas [12]. Este método consiste en ajustar tanto la posición como la anchura de cada una de las etiquetas lingüísticas de los antecedentes de las reglas. Para ello, al igual que los parámetros del consecuente, se ha calculado la derivada de la función del error respecto a la posición y la anchura de cada antecedente mediante la regla de la cadena.

E. Implementación escalable

Una vez explicadas las diferentes etapas del algoritmo propuesto, se analizan las características que lo hacen escalable. Para hacer frente a grandes conjuntos de datos se ha implementado siguiendo la filosofía Big Data que consiste en procesar los datos en paralelo de manera distribuida. Para ello, se ha dividido el conjunto de datos en particiones que se han distribuido en diferentes nodos de cálculo. Cada uno de estos nodos procesa la parte más pesada de los algoritmos descritos anteriormente en paralelo sobre su propia partición de datos. Posteriormente, estos resultados parciales de cada nodo se agregan continuando el flujo de resultados hacia la siguiente etapa. La Figura 3 muestra el diagrama del modelo escalable, indicando qué etapas se ejecutan de manera escalable en paralelo sobre las particiones de datos distribuidas.

La primera etapa consiste en ejecutar el algoritmo genético CHC, cuyo objetivo es evaluar la granularidad que ofrece mayor precisión. En esta fase, los cromosomas se evalúan secuencialmente en el nodo central, siendo la evaluación de cada cromosoma, es decir, el cálculo del error cuadrático medio sobre el conjunto de datos, la parte pesada que se realiza en paralelo sobre las diferentes particiones del conjunto de datos.

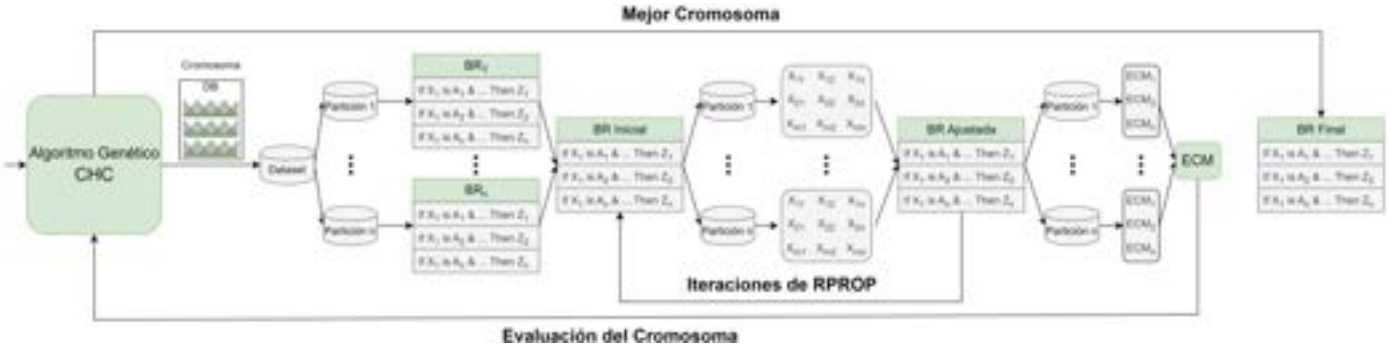


Fig. 3. Arquitectura del Modelo TSKEvoRPROP paralelo y distribuido

La evaluación del cromosoma se puede desglosar en las siguientes etapas:

- Generar la Base de Datos a partir de la granularidad representada por el cromosoma en el nodo central.
- Generar la Base de Reglas inicial por cubrimiento usando el algoritmo de Wang y Mendel en paralelo sobre cada partición y agregar las reglas posteriormente.
- Ajustar la Base de Reglas haciendo uso el algoritmo RPROP en paralelo sobre cada partición.
- Calcular el error cuadrático medio del SBRD en paralelo y agregar los errores posteriormente.
- Evaluar el siguiente cromosoma del algoritmo CHC.
- Obtener la Base de Reglas con mayor precisión.

En resumen, este trabajo propone un modelo de regresión integral que utiliza una implementación paralela y distribuida basada en Big Data, brindando escalabilidad horizontal. Este modelo genera un SBRD TSK-1, garantizando una alta precisión en los resultados. Se ha priorizado la sencillez y la reproducibilidad del modelo, utilizando un algoritmo genético CHC para la búsqueda eficiente de la granularidad y el algoritmo de aprendizaje supervisado RPROP para acelerar el ajuste de los parámetros.

IV. Estudio Experimental

En esta sección se presentan los resultados derivados de la experimentación llevada a cabo con el propósito de someter a evaluación la implementación del modelo TSKEvoRPROP. El enfoque central es valorar los resultados en términos de precisión y escalabilidad, proporcionando así un análisis detallado de la efectividad y la capacidad de adaptación del modelo propuesto en el contexto de problemas de regresión de grandes dimensiones.

Se ha evaluado la capacidad predictiva del modelo sobre un conjunto de *datasets* representativos. Se ha llevado a cabo una comparación exhaustiva con otros algoritmos de vanguardia en el ámbito de los SBRDs escalables para regresión. La evaluación de la precisión del modelo se ha medido mediante el error cuadrático medio sobre el conjunto de prueba.

Primero, se muestran los *datasets* utilizados en este estudio experimental. Posteriormente, se detalla la configuración del entorno de experimentación. Finalmente, se muestran los resultados obtenidos y se realiza la comparativa con otros modelos.

A. *Datasets*

Se han seleccionado diversos *datasets* de gran magnitud, los cuales se han obtenido del repositorio KEEL. La Tabla I muestra las principales características de los problemas de regresión elegidos para realizar el estudio experimental.

TABLA I. DATASETS UTILIZADOS EN EL ESTUDIO EXPERIMENTAL

Problema	Abr.	Variables	Instancias
Delta Ailerons	DELAIL	5	7129
Delta Elevators	DELELV	6	9517
California Housing	CAL	8	20640
Mv Artificial Domain	MV	10	40768
House	HOU	16	22768
Elevators	ELV	18	16599
Computer Activity	CA	21	8192
Pole Telecommunications	POL	26	14998
Pumadyn	PUM	32	8192
Airelons	AIL	40	13750
Tic	TIC	85	9822

B. *Entorno de experimentación*

Para la ejecución de las pruebas reales, se ha configurado un clúster Hadoop/Spark formado por 6 servidores. Este clúster HPC permite validar la escalabilidad del modelo con una capacidad de cómputo de 200 núcleos a 2.2 GHz y 1.25 TB de memoria RAM en total.

La configuración de los hiperparámetros del modelo ha sido ajustada en base a la experiencia en ensayos controlados. El algoritmo CHC se ha configurado con un tamaño de población de 6 cromosomas y un límite de 60 cruces. Por otro lado, el algoritmo RPROP se ha configurado con un límite de 150 iteraciones y un desplazamiento inicial (Δ_0) del 0,05% del universo de discurso de cada variable.

Sería necesario realizar un estudio de mayor profundidad para afinar los hiperparámetros, pero hemos considerado tanto el algoritmo genético CHC como el algoritmo RPROP como herramientas con configuraciones fijas. Sin embargo,

probablemente se deban considerar configuraciones adaptativas para cada *dataset* o, al menos, diversas configuraciones prediseñadas para *datasets* similares en número de variables y complejidad. Respecto a la configuración de RPROP, se ha detectado que, para diferentes *datasets*, funcionan mejor diferentes configuraciones. Se han usado estos hiperparámetros debido a que son los empleados en el trabajo original y aunque son efectivos, cabe destacar que el hiperparámetro Δ_0 tiene un impacto significativo en el resultado. A pesar de que teóricamente no debe influir demasiado, ya que Δ se ajustará progresivamente, se observan casos en los que el algoritmo RPROP puede llegar a converger prematuramente en mínimos locales debido al valor de Δ_0.

C. Resultados

En comparación con algoritmos del estado del arte en el ámbito de los SBRDs escalables para regresión, la Tabla II recoge los resultados obtenidos sobre los conjuntos de prueba tanto de TSKEvoRPROP, como del algoritmo S-FRULER. Los mejores resultados para el conjunto de prueba se encuentran resaltados.

TABLA II. COMPARATIVA DE RESULTADOS

Dataset	**TSKEvoRPROP**		**S-FRULER**
	ECM Ent	***ECM Pru***	***ECM Pru***
DELAIL	$0{,}80\cdot10^{-8}$	$1{,}65\cdot10^{-8}$	$\mathbf{1{,}44\cdot10^{-8}}$
DELELV	$0{,}34\cdot10^{-6}$	$2{,}50\cdot10^{-6}$	$\mathbf{1{,}12\cdot10^{-6}}$
CAL	$3{,}30\cdot10^{8}$	$\mathbf{3{,}45\cdot10^{8}}$	$2{,}18\cdot10^{9}$
MV	$0{,}18\cdot10^{-4}$	2,13	**0,05**
HOU	$2{,}57\cdot10^{8}$	$\mathbf{2{,}65\cdot10^{8}}$	$8{,}20\cdot10^{8}$
ELV	$0{,}06\cdot10^{-5}$	$1{,}64\cdot10^{-5}$	$\mathbf{3{,}20\cdot10^{-6}}$
CA	1,29	2,27· 10	**4,60**
POL	3,21	$3{,}24\cdot10^{4}$	$\mathbf{1{,}24\cdot10^{2}}$
PUM	$0{,}10\cdot10^{-9}$	$4{,}71\cdot10^{-4}$	$\mathbf{0{,}34\cdot10^{-4}}$
AIL	$0{,}38\cdot10^{-8}$	$1{,}99\cdot10^{-8}$	$\mathbf{1{,}40\cdot10^{-8}}$
TIC	$2{,}61\cdot10^{-2}$	$\mathbf{2{,}73\cdot10^{-2}}$	-

En comparación con S-FRULER, TSKEvoRPROP muestra una mayor precisión sobre los *datasets* CAL, y HOU. Es probable que aún no se alcance la precisión deseada en este estudio preliminar debido a la necesidad afinar la configuración de los hiperparámetros del modelo. Se anticipa que el modelo propuesto proporcionará mejores resultados al controlar el evidente sobreaprendizaje en este momento. Para abordar este desafío, se propone reducir la intensidad de RPROP en general, regulándola a lo largo de la búsqueda para que el modelo tenga una mayor capacidad de exploración al principio y una mejor capacidad de explotación en la fase final.

En definitiva, hemos logrado resultados prometedores. Es crucial destacar que nuestro objetivo es mantener la reproducibilidad de nuestro modelo. Por lo tanto, como primera aproximación consideramos que el resultado ha sido un éxito.

V. CONCLUSIONES

En este trabajo se propone un SBRD TSK-1 denominado TSKEvoRPROP, escalable en instancias, que destaca por su relativa sencillez y reproducibilidad frente a otras alternativas que también hacen uso de las tecnologías de Big Data para ofrecer una alta escalabilidad. Nuestra propuesta combina un algoritmo evolutivo con un algoritmo de descenso por gradiente optimizado para acelerar la convergencia de las soluciones dentro del espacio de búsqueda.

Si bien los resultados actuales aún no son óptimos, éstos se sitúan bastante próximos a los de modelos más complejos, pues se ha priorizado la escalabilidad y la simplicidad del modelo para facilitar su reproducibilidad y empleabilidad práctica.

AGRADECIMIENTOS

Este trabajo ha sido soportado por el proyecto PID2020-119478GB-I00 del Ministerio de Ciencia e Innovación de España.

REFERENCIAS

[1] Fernandez, A., Herrera, F., Cordon, O., del Jesus, M. J., & Marcelloni, F. (2019). Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?. *IEEE Computational intelligence magazine*, *14*(1), 69-81.

[2] Abdalla, H. B., & Abuhaija, B. (2023). Comprehensive analysis of various big data classification techniques: A challenging overview. *Journal of Information & Knowledge Management*, *22*(01), 2250083.

[3] Havens, T. C., Bezdek, J. C., Leckie, C., Hall, L. O., & Palaniswami, M. (2012). Fuzzy c-means algorithms for very large data. *IEEE Transactions on Fuzzy Systems*, *20*(6), 1130-1146.

[4] del Rio, S., López, V., Manuel Benitez, J., & Herrera, F. (2015). A mapreduce approach to address big data classification problems based on the fusion of linguistic fuzzy rules. *International journal of computational intelligence systems*, *8*(3), 422-437.

[5] Rodríguez-Fdez, I., Mucientes, M., & Bugarín, A. (2016). S-FRULER: Scalable fuzzy rule learning through evolution for regression. *Knowledge-Based Systems*, *110*, 255-266.

[6] Fernandez, A., Lopez, V., del Jesus, M. J., & Herrera, F. (2015). Revisiting evolutionary fuzzy systems: Taxonomy, applications, new trends and challenges. Knowledge-Based Systems, 80, 109-121.

[7] Nauck, D., & Kruse, R. (2020). Neuro–Fuzzy Systems. In *Handbook of Fuzzy Computation* (pp. 319-D2). CRC Press.

[8] Gacto, M. J., Galende, M., Alcalá, R., & Herrera, F. (2014). METSK-HDe: A multiobjective evolutionary algorithm to learn accurate TSK-fuzzy systems in high-dimensional and large-scale regression problems. *Information Sciences*, *276*, 63-79.

[9] Eshelman, L. J. (1991). The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In *Foundations of genetic algorithms* (Vol. 1, pp. 265-283). Elsevier.

[10] Wang, L. X., & Mendel, J. M. (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on systems, man, and cybernetics*, *22*(6), 1414-1427.

[11] Riedmiller, H. B. M., & Braun, H. (1993, February). Rprop: a fast and robust backpropagation learning strategy. In *Proc. of the ACNN*.

[12] Alcalá, R., Alcalá-Fdez, J., Gacto, M. J., & Herrera, F. (2007). Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation. *Soft Computing*, *11*(5), 401-419.

Sistemas Difusos Federados en el Borde con Autoajuste en Línea

Javier Martín Moreno
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
javier.martin@dti.uhu.es

Francisco Alfredo Márquez Hernández
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
alfredo.marquez@dti.uhu.es

Antonio Peregrín Rubio
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Centro de Estudios Avanzados en Física, Matemáticas y Computación (CEAFMC)
Universidad de Huelva
Huelva, España
peregrin@dti.uhu.es

***Resumen*—El aprendizaje federado facilita la colaboración entre dispositivos ubicuos para desempeñar tareas de aprendizaje automático de manera distribuida. Este enfoque aprovecha el aumento en la capacidad de cómputo de estos dispositivos, combinado con la optimización los algoritmos de aprendizaje automático, lo que permite realizar la adquisición de conocimiento de manera local. Además, el uso de Sistemas Basados en Reglas Difusas permite la modelización del conocimiento extraído de forma interpretable para los seres humanos. Este trabajo propone el ajuste paramétrico de modelos difusos en el propio dispositivo, seguido de su federación con el objetivo de mejorar su precisión de forma colaborativa.**

Palabras clave—Aprendizaje Federado; Sistema Basado en Reglas Difusas; Internet of Things; TinyML; Regresión

I. INTRODUCCIÓN

La generación constante de datos por parte de dispositivos con recursos limitados es una realidad común en el ámbito del Internet de las Cosas (IoT) [1]. Estos dispositivos, al estar interconectados, forman una red de datos que requiere procesamiento. Tradicionalmente, debido a las restricciones de recursos de los dispositivos IoT, estos datos se han enviado a la nube para su posterior análisis. Sin embargo, con la creciente preocupación por la privacidad de los datos, surge el concepto de Edge Computing [2], que propone llevar a cabo el procesamiento de los datos en el propio dispositivo. A medida que mejora la capacidad de cómputo de los dispositivos IoT, aumenta la posibilidad de implementar diversos algoritmos de procesamiento de datos en estos dispositivos.

Siguiendo esta evolución, surge el TinyML, que emerge de la optimización de los modelos clásicos de aprendizaje automático (Machine Learning – ML) con el fin de adaptarlos a los requisitos específicos de los dispositivos IoT [3]. Esta adaptación permite implementar algoritmos de Inteligencia Artificial (IA) en el borde de manera local, sin necesidad de comunicar los datos con el exterior. El uso de esta metodología deriva en varias ventajas en comparación con el procesamiento de datos en la nube [4]:

- Privacidad de los datos: Cuando la privacidad de los datos es primordial, procesar los datos en el propio dispositivo reduce el riesgo de ataques.
- Tiempo de respuesta real: En aplicaciones de control y monitorización, la reducción de la latencia al procesar los datos localmente evita los retrasos asociados con el envío de estos a través de la red y la espera de una respuesta.
- Eficiencia energética: A medida que aumenta la potencia de los centros de datos, también lo hace la demanda energética. La optimización de algoritmos para su ejecución en dispositivos IoT contribuye a disminuir el consumo energético asociado al procesamiento de datos.

Es importante tener en cuenta que, en un principio, la filosofía de TinyML se propone considerando que la etapa más costosa del proceso, esto es, el aprendizaje o el ajuste de los modelos, se realiza fuera del dispositivo. De esta manera, sólo es necesario adaptar el modelo obtenido para su ejecución en el dispositivo. Esto implica que el dispositivo puede ejecutar el modelo entrenado para realizar inferencia localmente en un tiempo de respuesta mínimo, manteniendo los datos en el dispositivo. Este enfoque consigue preservar la privacidad de los datos y reduce la necesidad de transmitir información a través de la red para su procesamiento en el exterior.

Posteriormente, surgió la metodología On Device Learning (ODL), que busca adaptar, además, la propia etapa de aprendizaje o ajuste satisfaciendo las restricciones de los dispositivos de recursos limitados [5], utilizando los datos locales. En ciertas situaciones esta capacidad es esencial, ya que permite actualizar los modelos en tiempo real en entornos dinámicos donde los datos y las condiciones cambian con frecuencia, garantizando que los modelos estén siempre actualizados y sean más efectivos en su desempeño.

Por otro lado, se denomina Tiny Online Learning (TinyOL) a una subárea del ODL, que se caracteriza por la adaptación de la etapa de ajuste a dispositivos de recursos limitados, inspirada

en la filosofía de aprendizaje incremental [6]. Esta técnica se basa en el análisis de flujos de datos en tiempo real, de manera que los datos no se almacenan, y por tanto se satisface la escasa capacidad de almacenamiento. Considerando un modelo inicial, se realiza la actualización de manera iterativa en cada dispositivo mediante el uso de ventanas de datos adaptadas a los requisitos específicos del dispositivo en cuestión.

Por último, en este sentido, cabe destacar que la capacidad de ajuste en el propio dispositivo está condicionada por la localidad de los datos. Para superar esta limitación es necesario implementar técnicas de colaboración entre múltiples dispositivos aislados. El Aprendizaje Federado (Federated Learning – FL) surge como respuesta a esta necesidad, combinando los resultados del ajuste local de diferentes nodos para generar un modelo global colaborativo [7]. Esta metodología mantiene las ventajas de ambas filosofías, permitiendo el ajuste del modelo sobre diversas fuentes de datos mientras se preserva su privacidad. Además, se aumenta la escalabilidad de la capacidad de cómputo global, ya que cada dispositivo puede procesar sus datos de manera local, lo que facilita el procesamiento de grandes volúmenes de datos de manera distribuida.

Por otro lado, como es bien sabido, los Sistemas Basados en Reglas Difusas (SBRD) han sido ampliamente utilizados para resolver problemas que requieren modelar el conocimiento de manera similar a la percepción humana. Este enfoque permite procesar correctamente información incompleta y/o imprecisa, particularmente en entornos dinámicos de incertidumbre. Los SBRDs permiten resolver problemas de aprendizaje supervisado, tanto de clasificación como de regresión, y problemas de aprendizaje no supervisados como clustering.

En la actualidad, la importancia de generar modelos de IA transparentes, interpretables y explicables, ha propiciado que se mantenga la atención y el interés en los SBRDs [8], pues son relevantes en un en el marco de la Inteligencia Artificial Explicable (Explainable Artificial Intelligence – XAI).

El desafío en el diseño de SBRDs consiste en obtener una Base de Conocimiento (BC) que pueda realizar predicciones precisas. Tradicionalmente se han utilizado algoritmos de optimización paramétrica para abordar el problema. En la literatura se proponen diversas metodologías para resolverlo, como los algoritmos difusos evolutivos (Genetic Fuzzy Systems – GFS) [9], o los algoritmos neuro-difusos (Neuro Fuzzy Systems – NFS) [10]. En ambos casos, el objetivo es optimizar la capacidad predictiva del modelo, ya sea para tareas de clasificación, regresión o clustering, mediante la modificación de los parámetros que componen la BC.

El objetivo de este trabajo es proponer la implementación de un nuevo modelo federado cuyos nodos realizan el ajuste de SBRDs locales de forma también novedosa mediante TinyOL, denominado FedFuzzy-OL.

La organización de este trabajo es la siguiente: La Sección II aborda el ajuste federado de modelos difusos. En la Sección III se define el modelo difuso TinyOL federado propuesto. La Sección IV ilustra la utilidad del modelo propuesto mediante la resolución de un problema real. Finalmente, en la Sección V se presentan las conclusiones principales de este trabajo.

II. Ajuste Federado de Modelos Difusos

Debido a su naturaleza liviana, los SBRDs han estado presentes en los dispositivos de IoT históricamente [11]. Sin embargo, en estos casos el modelo siempre se ha mantenido estático, con la BC aprendida/ajustada fuera del dispositivo en unidades de cómputo más potentes. Esta solución no es la idónea en entornos dinámicos, pues los modelos pueden quedar obsoletos ante cambios en el entorno. Para superar este desafío, es esencial ajustar el modelo en el propio dispositivo, procesando los datos en tiempo real.

La filosofía ODL propone la adaptación de los algoritmos de ajuste paramétricos tradicionales, como GFS o NFS, a los recursos disponibles en los dispositivos IoT. Esto implica acercar el procesamiento a la fuente de datos, evitando la necesidad de exportar los datos para su procesamiento en el exterior. No obstante, esta tarea no es trivial, ya que la adaptación del algoritmo puede variar significativamente según el tipo de algoritmo y el entorno de aplicación en cuestión.

La Figura 1 simboliza el ajuste paramétrico de la BC de un modelo difuso en el propio dispositivo mediante ODL.

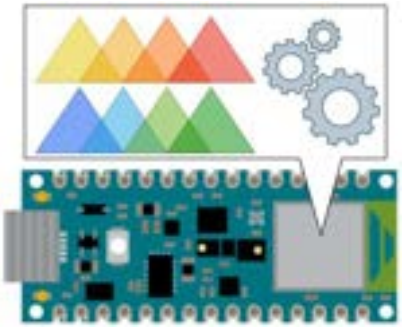


Fig. 1. Representación del ajuste paramétrico de la BC en el dispositivo.

Una forma de implementar soluciones ODL es a través del aprendizaje incremental. La metodología TinyOL, se basa en el procesamiento de datos en tiempo real, sin tener que almacenarlos. Utilizando una ventana de datos, el modelo se ajusta de manera iterativa, ejecutando el algoritmo sobre los datos generados. De esta manera, se elimina la necesidad de almacenar un gran conjunto de datos, ya que estos se procesan gradualmente a medida que se generan, lo que reduce el coste computacional del algoritmo al operar sobre particiones de datos menores.

Este enfoque, combinado con otras técnicas de optimización como la reducción del espacio de búsqueda mediante la discretización o la delimitación del espacio de búsqueda, permite la implementación de algoritmos de ajuste paramétrico directamente en el propio dispositivo.

En entornos donde los datos generados en el dispositivo no sean suficiente para alcanzar buena precisión adecuada, se recurre a técnicas de aprendizaje colaborativo como el FL. Esta estrategia permite mejorar la precisión del modelo mientras se mantiene el bajo coste computacional en dispositivos con

recursos limitados. Por otro lado, desde un punto de vista ético, se incrementa la privacidad del usuario compartiendo el modelo difuso en lugar de los datos.

La Figura 2 representa la colaboración mediante FL de múltiples dispositivos difusos que han realizado el ajuste paramétrico del modelo localmente.

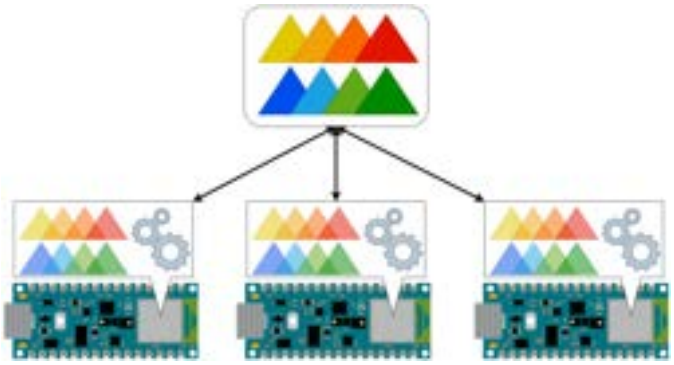


Fig. 2. Federación de múltiples dispositivos para alcanzar un modelo global de BC ajustada.

La filosofía de FL nos lleva a plantear una arquitectura compuesta por múltiples nodos, donde cada nodo ajusta su propio modelo local de manera independiente. Posteriormente, estos modelos locales se comparten para crear un modelo global que integre el conocimiento de todos los nodos. Finalmente, este modelo global se distribuiría nuevamente a los nodos, quienes actualizarían sus modelos locales con las aportaciones del modelo global. Este ciclo de actualización se repite de manera iterativa a lo largo del tiempo, permitiendo que el modelo global se beneficie continuamente de la evolución alcanzada por todos los nodos.

La implementación de esta metodología supone ciertas decisiones de diseño:

- Arquitectura del sistema: En función de la presencia o ausencia de un nodo central encargado de la agregación de las aportaciones de los demás nodos, se pueden clasificar en modelos centralizados [7] o descentralizados [12].
- Coordinación del sistema: Si la agregación de los modelos locales no se realiza hasta que todos hayan participado se denomina sistema síncrono [7]; por el contrario, si la agregación se realiza cuando se cumple cierta condición, como, por ejemplo, transcurrido cierto tiempo, se considera un sistema asíncrono [13].
- Distribución de los datos en el sistema: Si los datos de los nodos comparten las mismas variables, se considera un sistema horizontal [7], mientras que, si los nodos tienen diferentes variables, pero corresponden a las mismas instancias, se trata de un sistema vertical [14].
- Método de agregación del sistema: Dependiendo de la técnica empleada para combinar el conocimiento de los modelos se pueden distinguir diferentes enfoques, siendo el más habitual el promedio (FedAvg) [7], comúnmente utilizado en arquitecturas de redes neuronales, donde los pesos del modelo global son calculados como el promedio de los pesos de los modelos locales. Se puede considerar también el uso de la agregación ponderada (PW) [15], en la que los pesos del modelo global son el promedio de los modelos locales ponderado a su respectiva precisión.
- Tipo de modelos aprendidos: Si todos los nodos aprenden el mismo tipo de modelo o no, se distingue entre sistemas homogéneos [7] y heterogéneos [16].
- Propósito: En determinadas ocasiones, el enfoque se centra en optimizar los resultados del modelo local, centrados en el consumidor, donde la prioridad es ajustar el modelo para adaptarse mejor a las necesidades específicas de cada dispositivo [17]. En otros casos, el objetivo principal es mejorar el modelo global, centrados en el proveedor [7]. Además, existe el concepto de FL personalizado, donde el objetivo es adaptar el modelo global para abordar la heterogeneidad entre los datos de los dispositivos, de manera similar a los sistemas centrados en el consumidor [18].
- Tipo de nodo: El FL puede realizarse en nodos con una gran capacidad de cómputo, denominados sistemas *cross-silo*, donde cada nodo puede estar compuesto por una o varias computadoras que posteriormente compartirán su modelo local entrenado sobre sus propios datos [19]. Por otro lado, también puede implementarse en dispositivos IoT de bajos recursos, denominados sistemas *cross-device*, donde estos tienen una capacidad de cómputo más reducida, pero un mayor número de nodos [7].
- Variedad de los nodos: Si todos los nodos tienen las mismas características la arquitectura es homogénea [7]; en caso contrario, se trata de una arquitectura heterogénea [20].
- Ubicación del aprendizaje: Normalmente el entrenamiento del modelo se lleva a cabo íntegramente en los dispositivos, y luego se agregan todos los modelos en el servidor central [7], pero es posible, en determinadas situaciones, utilizar la técnica divide y vencerás para realizar en cada nodo sólo una parte del entrenamiento, de modo que el servidor central se encargue de agregar todos estos cálculos parciales y completar las demás etapas del entrenamiento [21].

Otros factores adicionales que se deben considerar son: la cantidad de pasos de comunicación en cada ronda de aprendizaje, la posible aplicación de un proceso de selección de nodos participantes, la frecuencia con la que se actualiza el modelo, la implementación de medidas para proteger la privacidad de los datos, así como el refuerzo de la seguridad con técnicas de defensa ante ataques.

En la siguiente sección se desarrolla el modelo propuesto teniendo en cuenta los aspectos anteriores esenciales para garantizar la eficacia y la integridad del sistema federado.

III. Modelo Propuesto: FedFuzzy-OL

En esta sección se presenta FedFuzzy-OL, un SBRD TinyOL Federado. La representación del conocimiento se

realizará mediante el diseño de un SBRD, utilizado tanto para inferir la salida del sistema frente a problemas de regresión, como para predecir la clase frente a problemas de clasificación. Este SBRD se ajustará en tiempo real en el propio dispositivo de manera local mediante una metaheurística liviana que optimiza los resultados de las predicciones. Posteriormente, se combinarán los SBRDs ajustados localmente en los dispositivos dando lugar a un SBRD global federado que integra el conocimiento adquirido en todos los dispositivos. Este modelo global se comparte con los dispositivos, reemplazando los modelos locales previos.

A. Ajuste del Modelo Local

En cada dispositivo, de manera local, se utiliza el modelo en tiempo real para dar respuesta a los datos de entrada. Estos datos de entrada se almacenarán temporalmente en un buffer. Cuando se considera que el modelo está desfasado debido a la evolución del entorno, se ajustará en el propio dispositivo para mantener la precisión de las predicciones. El ajuste se basará en la metodología de aprendizaje incremental TinyOL. Se utiliza el pequeño conjunto de datos almacenado en el buffer para ajustar el modelo local previo. Mediante una metaheurística liviana que se ajusta a los requerimientos del dispositivo, se realiza la actualización del modelo local. Dependiendo de la implementación del modelo, se puede configurar un umbral que defina cuándo el modelo local requiere ser actualizado, o definir un período de tiempo de actualización.

El sistema partirá de un SBRD inicial proporcionado por un experto, el cuál compartirán todos los nodos al inicio. Cuando el SBRD se considera obsoleto, se ejecuta el método de ajuste en el propio dispositivo sobre sus propios datos locales, evitándose la latencia provocada por el tráfico de datos, preservando su privacidad. Para el ajuste del modelo se propone el empleo de una técnica evolutiva configurada acorde a las restricciones del dispositivo, aunque se hace hincapié en que esta propuesta no se refiere a un mecanismo de ajuste en concreto.

Siguiendo el enfoque del aprendizaje incremental de la filosofía TinyOL, no será necesario ejecutar esta técnica de manera exhaustiva, ya que se efectuará varias veces a lo largo del tiempo, sobre ventanas de datos relativamente pequeñas. De esta forma, el algoritmo evolutivo no es necesario que complete la búsqueda de la solución en una sola pasada, sino que se dividirá en varias ejecuciones a lo largo del tiempo, lo que permitirá ajustar gradualmente el SBRD y adaptarlo a los cambios en las tendencias de los datos.

El ajuste de la BC del SBRD utiliza etiquetas lingüísticas codificadas mediante 3-tuplas lingüísticas: los parámetros para optimizar son la posición y la anchura de cada etiqueta.

Un reducido número de parámetros supone una longitud de población reducida en el algoritmo evolutivo, en el que cada cromosoma representa un SBRD, cuyos genes son los parámetros del modelo.

B. Agregación del Modelo Global

Una vez que cada nodo ha ajustado su modelo localmente evaluándolo sobre sus propios datos, este se comparte con los demás nodos del conjunto mediante comunicación directa. Para evitar la centralización, se propone la implementación de un sistema donde la agregación del modelo sea descentralizada. De manera asíncrona, cuando cada nodo cumpla las condiciones establecidas, agregará las aportaciones de los modelos locales disponibles.

Concretamente, un nodo recibirá los modelos locales de varios nodos vecinos, y cuando el modelo local se considere obsoleto, se actualiza teniendo en cuenta las aportaciones de los demás nodos que hayan compartido su SBRD hasta ese momento. Por tanto, la agregación descentralizada y asíncrona en cada nodo consistirá en combinar varios SBRD. La agregación propuesta consiste en promediar los valores de posición y anchura de las etiquetas de los diferentes modelos, de tal modo que cada nodo genera su propio modelo global.

Para finalizar la etapa de ajuste federado del modelo, cada nodo evaluará el nuevo modelo global sobre su ventana de datos temporal y escogerá entre el modelo local previo o el modelo global actualizado, en función de cuál obtenga una mejor evaluación. Al implementar esta técnica de elitismo, en lugar de sustituir siempre el modelo local por el global independientemente de la precisión, se fomenta la filosofía centrada en el consumidor, priorizando la precisión del modelo local, personalizando el modelo a cada dispositivo.

El modelo actualizado será utilizado durante la siguiente etapa de inferencia, tomando decisiones cada vez que se disponga de datos de entrada para generar una salida, hasta que vuelva a quedar obsoleto. En ese momento, se actualizará localmente de nuevo y se volverá a agregar con las aportaciones de los demás nodos.

En definitiva, se ha implementado un SBRD ajustado en el propio dispositivo mediante TinyOL, apoyado en el aprendizaje federado, mediante un sistema *cross-device* homogéneo descentralizado y asíncrono, con una arquitectura de datos horizontal. Además, se ha propuesto un método de agregación promedio. El sistema se ha enfocado en la precisión de los nodos utilizando una política de reemplazo elitista.

Con objeto de mostrar la utilidad del modelo propuesto, a continuación, se ilustra con un problema real.

IV. Estudio Experimental

Un modelo difuso generado de manera federada en dispositivos IoT puede ser útil en gran variedad de aplicaciones. En este trabajo, proponemos como ejemplo de aplicación un sistema federado de gestión de fuentes de energía eléctrica para instalaciones de energía renovable en viviendas dotadas de sistemas fotovoltaicos (paneles solares y almacenamiento en baterías) no aislados, es decir, con conexión a la red eléctrica. Este modelo es de gran interés, ya que en lugar de depender de grandes baterías como haría un sistema autónomo, que suelen ser el elemento más costoso de estas instalaciones, utiliza una cantidad de almacenamiento menor y más económica, para reducir la factura energética.

El objetivo de este modelo, implementado en dispositivos IoT, es tomar decisiones sobre el consumo energía renovable o convencional de la red eléctrica, de manera que se optimice el

coste para el usuario de la vivienda. Las decisiones se basarán en factores relevantes para nuestro objetivo, utilizando como entradas el consumo esperado de energía en la vivienda, la energía renovable disponible almacenada, la radiación solar esperada y el coste de la energía convencional; por ejemplo, si se dispone de suficiente energía renovable, el consumo esperado es bajo y el coste de la energía convencional es elevado, se preferirá usar energía renovable. La Figura 3 ilustra el entorno de experimentación, formado por tres dispositivos, cada uno ubicado en cada vivienda.

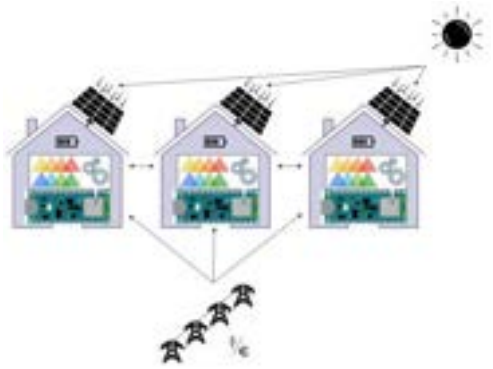


Fig. 3. Entorno de experimentación.

Adicionalmente, de manera periódica, las viviendas compartirán su modelo local para ser agregados en un modelo global, de manera que cada vivienda pueda beneficiarse de lo aprendido por los dispositivos de las demás.

El conjunto de datos distribuido se ha generado especialmente para ilustrar esta aplicación. Por un lado, se han obtenido los datos de coste de la electricidad usando la API de ESIOS, expresados en euros por kilovatios por hora. Por otro lado, para simular el rendimiento de la instalación fotovoltaica, se han empleado los datos de la Agencia Andaluza de la Energía sobre la radiación solar de la zona, medida en vatios por hora por metro cuadrado. Estos datos se compartirán por todos los nodos, ya que serán comunes para la zona a la que pertenecen. Los datos privados de cada dispositivo serán los relacionados con el consumo de cada vivienda. Para ello, hemos extraído los consumos medidos en kilovatios hora usando los recursos que cada distribuidora eléctrica pone a disposición de sus clientes. La Figura 4 ilustra la representación gráfica de los datos pertenecientes a un día en un solo nodo.

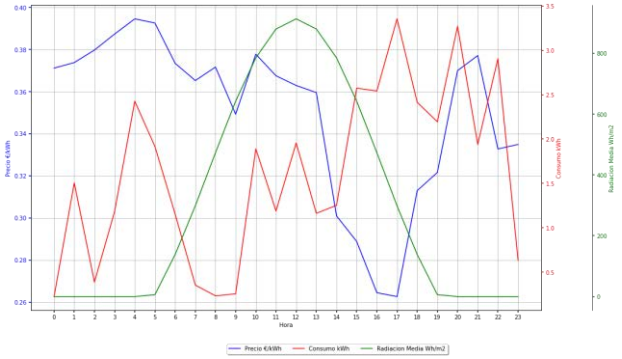


Fig. 4. Representación gráfica de los datos usados en la experimentación.

Para facilitar la labor del ajuste del SBRD se ha optado por una agregación de variables, lo que permite transformar el conjunto de datos de entrada inicial formado por cuatro variables, en un nuevo conjunto con solo dos variables: el precio de la electricidad y el equilibrio energético. Para ello, se han combinado las variables relacionadas con el consumo eléctrico, la radiación solar y la energía almacenada en una sola variable que indica si el equilibrio energético en conjunto tiende a ser renovable o no. Así, un equilibrio energético positivo indica que se dispone de suficiente energía renovable para afrontar el consumo energético esperado, mientras que uno negativo indica que es necesario recurrir además al suministro de energía tradicional. Adicionalmente, se han aplicado otras decisiones para favorecer el procesamiento en el dispositivo, como la normalización, discretización y limitación del universo de discurso de ambas variables, usando valores enteros entre 0 y 100.

Se parte de un SBRD con una Base de Datos (BD) compuesta por dos variables de entrada y una de salida, cuya Base de Reglas (BR) será definida por un experto. Se emplean tres etiquetas lingüísticas por variable, dando lugar a una BR de cubrimiento completo con 9 reglas.

La reducción deliberada de parámetros tiene como objetivo reducir el espacio de búsqueda del algoritmo evolutivo encargado de ajustar la BD, para el que se ha optado por modelo de tipo CHC. La configuración de este algoritmo tiene en cuenta que se ejecutará varias veces a lo largo del tiempo siguiendo la filosofía TinyOL, por lo que usa una población reducida de 10 cromosomas y sólo 100 evaluaciones.

El citado ajuste local del SBRD se lleva a cabo diariamente, y dado que los datos se generan cada hora, se realiza el ajuste con una ventana de 24 instancias. Una vez ajustado, se envía a los demás nodos, los cuales agregan todos los SBRD recibidos de los nodos del sistema cuando consideren que su modelo local está obsoleto.

Finalmente, se evalúa la precisión del nuevo SBRD global sobre los datos almacenados en el buffer, y en el caso de que mejore las predicciones del SBRD local previo, se sustituye, manteniendo actualizado así el SBRD para la nueva llegada de datos del día siguiente. Si bien es cierto que lo más habitual es que el modelo local presente mejores resultados que el modelo global, en determinadas ocasiones, si el entorno ha variado lo suficiente y los dispositivos cercanos han anticipado mejor esa circunstancia, el sistema se beneficia del aprendizaje colaborativo.

La medida de precisión usada para evaluar los SBRDs es el porcentaje de ahorro derivado de las decisiones de usar energía renovable o no. Con objeto de comparar los resultados se ha considerado como modelo base la gestión del sistema fotovoltaico mediante el SBRD inicial definido por el experto de manera estática, sin ajuste local ni federado. La Figura 5 muestra el porcentaje de ahorro diario de un nodo en el mes de mayo de 2022 obtenido por FedFuzzy-OL, mientras que la Figura 6 muestra el porcentaje de ahorro mensual de un nodo obtenido por FedFuzzy-OL, en ambos casos frente al modelo base. En líneas generales, el modelo propuesto logra mejorar los resultados obtenidos por el modelo base, por tanto, se optimiza el consumo, lo que se traduce en un incremento del

porcentaje de ahorro en la mayoría de los casos, proporcionando así un mayor beneficio para los clientes que colaboran en el sistema federado. En general, todos los nodos logran mejorar su porcentaje de ahorro acorde a sus perfiles de consumo, que pueden ser diferentes en cuanto a cantidad o tramos horarios.

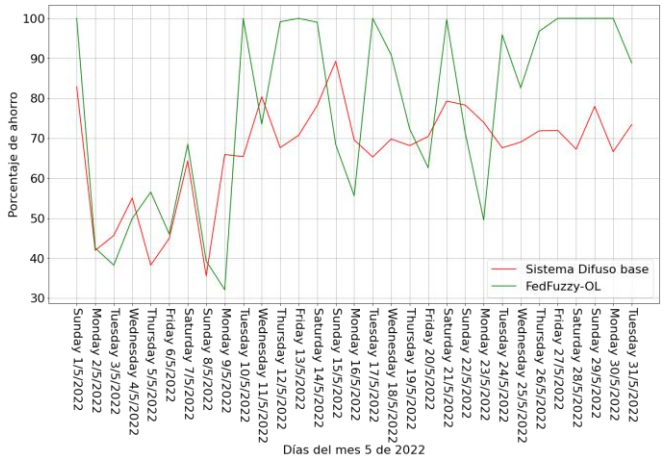


Fig. 5. Porcentaje de ahorro diario durante el mes de mayo de 2022.

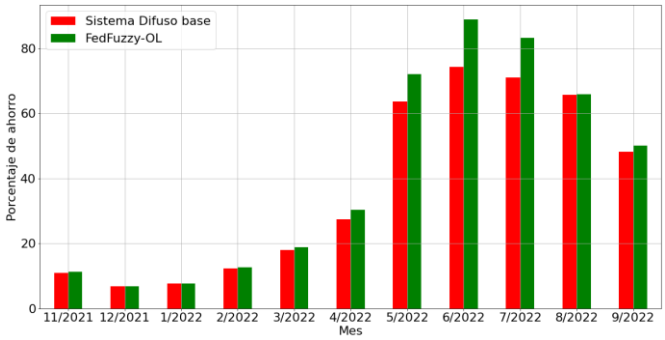


Fig. 6. Porcentaje de ahorro mensual.

V. Conclusiones

En entornos distribuidos, donde la preservación de la privacidad de los datos pueda ser fundamental, es necesario que el procesamiento de datos se realice en cada dispositivo. Estos dispositivos pueden beneficiarse del conocimiento adquirido por otros dispositivos al comunicar el modelo ajustado en lugar de sus datos. Los sistemas difusos son un elemento clásico para dotar de inteligencia a los dispositivos de bajos recursos que destaca por su explicabilidad inherente.

Este trabajo aborda preliminarmente una forma sencilla de potenciar los SBRDs adaptándolos a la filosofía de TinyOL, utilizando mecanismos de ajuste en línea que se puedan integrar en lugar de ser estáticos, combinado con técnicas de aprendizaje colaborativo entre dispositivos a través del FL.

Agradecimientos

Este trabajo ha sido soportado por el proyecto PID2020-119478GB-I00 del Ministerio de Ciencia e Innovación de España.

Referencias

[1] Madakam, S., Lake, V., Lake, V., & Lake, V. (2015). Internet of Things (IoT): A literature review. *Journal of Computer and Communications*, *3*(05), 164.

[2] Khan, W. Z., Ahmed, E., Hakak, S., Yaqoob, I., & Ahmed, A. (2019). Edge computing: A survey. *Future Generation Computer Systems*, *97*, 219-235.

[3] Ray, P. P. (2022). A review on TinyML: State-of-the-art and prospects. *Journal of King Saud University-Computer and Information Sciences*, *34*(4), 1595-1623.

[4] Abadade, Y., Temouden, A., Bamoumen, H., Benamar, N., Chtouki, Y., & Hafid, A. S. (2023). A comprehensive survey on tinyml. *IEEE Access*.

[5] Dhar, S., Guo, J., Liu, J., Tripathi, S., Kurup, U., & Shah, M. (2021). A survey of on-device machine learning: An algorithms and learning theory perspective. *ACM Transactions on Internet of Things*, *2*(3), 1-49.

[6] Ren, H., Anicic, D., & Runkler, T. A. (2021, July). Tinyol: Tinyml with online-learning on microcontrollers. In *2021 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

[7] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

[8] Fernandez, A., Herrera, F., Cordon, O., del Jesus, M. J., & Marcelloni, F. (2019). Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?. *IEEE Computational intelligence magazine*, *14*(1), 69-81.

[9] Fernandez, A., Lopez, V., del Jesus, M. J., & Herrera, F. (2015). Revisiting evolutionary fuzzy systems: Taxonomy, applications, new trends and challenges. *Knowledge-Based Systems*, *80*, 109-121.

[10] Nauck, D., & Kruse, R. (2020). Neuro–Fuzzy Systems. In *Handbook of Fuzzy Computation* (pp. 319-D2). CRC Press.

[11] Krishnan, R. S., Julie, E. G., Robinson, Y. H., Raja, S., Kumar, R., & Thong, P. H. (2020). Fuzzy logic based smart irrigation system using internet of things. *Journal of Cleaner Production*, *252*, 119902.

[12] Roy, A. G., Siddiqui, S., Pölsterl, S., Navab, N., & Wachinger, C. (2019). Braintorrent: A peer-to-peer environment for decentralized federated learning. *arXiv preprint arXiv:1905.06731*.

[13] Chen, Y., Ning, Y., Slawski, M., & Rangwala, H. (2020, December). Asynchronous online federated learning for edge devices with non-iid data. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 15-24). IEEE.

[14] Liu, Y., Kang, Y., Zou, T., Pu, Y., He, Y., Ye, X., ... & Yang, Q. (2024). Vertical Federated Learning: Concepts, Advances, and Challenges. *IEEE Transactions on Knowledge and Data Engineering*.

[15] Reyes, J., Di Jorio, L., Low-Kam, C., & Kersten-Oertel, M. (2021). Precision-weighted federated learning. *arXiv preprint arXiv:2107.09627*.

[16] Hu, L., Yan, H., Li, L., Pan, Z., Liu, X., & Zhang, Z. (2021). MHAT: An efficient model-heterogenous aggregation training scheme for federated learning. *Information Sciences*, *560*, 493-503.

[17] Mestoukirdi, M., Zecchin, M., Gesbert, D., Li, Q., & Gresset, N. (2021, December). User-centric federated learning. In *2021 IEEE Globecom Workshops (GC Wkshps)* (pp. 1-6). IEEE.

[18] Tan, A. Z., Yu, H., Cui, L., & Yang, Q. (2022). Towards personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems*.

[19] Huang, C., Huang, J., & Liu, X. (2022). Cross-silo federated learning: Challenges and opportunities. *arXiv preprint arXiv:2206.12949*.

[20] Nishio, T., & Yonetani, R. (2019, May). Client selection for federated learning with heterogeneous resources in mobile edge. In *ICC 2019-2019 IEEE international conference on communications (ICC)* (pp. 1-7). IEEE.

[21] Szegedi, G., Kiss, P., & Horváth, T. (2019, September). Evolutionary Federated Learning on EEG-data. In ITAT (pp. 71-78).

Dotando de nuevas funcionalidades a los portales de comercio electrónico actuales

María Garrido, José Jesús Castro-Schez, David Vallejo, Rubén Grande, Santiago Schez-Sobrino y Javier Albusac
Escuela Superior de Informática (Departamento de Tecnologías y Sistemas de Información)
Universidad de Castilla-La Mancha
Paseo de la Universidad 4, 13071 Ciudad Real, España
{maria.garrido, josejesus.castro, david.vallejo, ruben.grande, santiago.sanchez, javieralonso.albusac}@uclm.es

***Abstract*—Este artículo presenta una innovadora propuesta para los portales de comercio electrónico, centrada en el apoyo a las pequeñas y medianas empresas (PYMEs), integrando el concepto de "negocios satélite" en los portales. El objetivo es promover una relación simbiótica entre compras en línea, compras en negocios físicos y actividades de ocio. El portal animará a la compra o recogida en tiendas o en puntos locales específicos, con los beneficios que esto tiene para el medio ambiente y el fortalecimiento de las economías locales. Además, generarán y propondrán "planes de ocio" personalizados basados en el perfil del usuario a realizar en los negocios satélite. Esto último enriquecería la experiencia del usuario aportando un avance significativo en la evolución de los portales de comercio electrónico.**

***Index Terms*—Perfilado de usuario, Generador planes de ocio, Portales de Comercio electrónico.**

I. Introducción

En el contexto empresarial actual, el comercio electrónico desempeña un papel crucial al actuar como un impulsor clave para las empresas que buscan expandir su presencia en el mercado y mejorar su eficiencia operativa [1]. Específicamente, permite a las pequeñas y medianas empresas (PYMEs) competir con empresas más grandes [2]. Aunque las ventajas del comercio electrónico son evidentes, especialmente en la era posterior a COVID-19 [3], también presenta desafíos notables, especialmente para las PYMEs.

Uno de los desafíos principales que enfrentan las PYMEs es la brecha en habilidades digitales. Esta deficiencia dificulta su crecimiento y competitividad al obstaculizar la eficiencia, la innovación y la adaptación a las tendencias del mercado.

Además, la financiación insuficiente para la transformación digital es otro problema crítico que enfrentan las PYMEs. La falta de recursos financieros necesarios para llevar a cabo dicha transformación y desarrollar su presencia en línea, obstaculiza su capacidad para adaptarse a la era digital, donde la presencia en línea es cada vez más crucial para llegar a los clientes y competir en el mercado.

Así pues todas empresas, independientemente de su tamaño –grandes, medianas y pequeñas– deberían no sólo tener presencia en la web, sino también vender a través de ella. Para lograrlo, las empresas tienen dos alternativas principales: una es desarrollar su propia solución de comercio electrónico y la otra es vender a través de portales de comercio electrónico que dependen de terceros. Estos les permiten vender a través de su plataforma a cambio de alguna forma de compensación económica, p.e. comisión por venta (e.g. Amazon, AliExpress, Alibaba). Cada una de estas alternativas tiene sus pros y sus contras. Sin embargo, la segunda alternativa permite tener una presencia web más inmediata al usar una solución ya desarrollada. Además, estas soluciones cuentan con un alto nivel de madurez e implementación de las tecnologías actuales. Estar al tanto de las novedades tecnológicas e incorporarlas al portal es fundamental para los propietarios del portal, ya que es un elemento diferenciador que les ayudará a atraer negocios a su portal y ser competitivos en el mercado. Por ello, no sólo deben estar al tanto de las nuevas tecnologías que se están proponiendo, sino también proponer nuevas funcionalidades.

En cuanto a la incorporación de nuevas tecnologías, habría que tener en cuenta la Inteligencia Artificial (IA) y la Realidad Virtual (RV). La IA permite la personalización [4], optimización [5] y automatización [6] de varios aspectos del comercio. La RV permite mejorar la experiencia de compra en línea del cliente, acercándola a la experiencia de una tienda física, resolviendo así las limitaciones de espacio y de tiempo disponibles [7]. Al adoptar e integrar estas tecnologías se consigue una ventaja competitiva con respecto a otros portales competidores.

Tras esta introducción, el resto del artículo se organiza de la siguiente manera. En la Sección II se presenta la idea de portal sobre el que se podría implementar la idea propuesta en este trabajo. En la Sección III se define cómo se realiza el perfilado del usuario, fundamental a la hora de elaborar el plan de ocio. El algoritmo sugerido para obtener los planes de ocio se presenta en la Sección IV. Un ejemplo de uso del algoritmo será incluido en la Sección V. Finalmente, en la Sección VI se presentan nuestras conclusiones y el trabajo futuro.

II. Portal de Comercio Electrónico

En este trabajo se propone una nueva funcionalidad que puede ser de interés para los portales de comercio electrónico del futuro, que harán uso de IA y RV. Esta funcionalidad innovadora pretende dar respuesta a la nueva generación de consumidores, más responsables y críticos con el proceso de compra, intentando reducir el impacto del comercio electrónico en el medio ambiente y en los negocios tradicionales y prosperidad de las economías locales [8]. En el portal de comercio electrónico propuesto convivirán tiendas online y

físicas con negocios y lugares de ocio, entretenimiento y cultura, tales como bares, restaurantes, teatros, cines, museos, clubs deportivos, gimnasios, salas de exposiciones, entre otros. A este tipo de negocios se les denominan como "negocios satélite" y se definirán formalmente como:

$$NS = (id, nombre, dirección, latitud, longitud, email, pwd) \quad (1)$$

Los campos de datos cruciales incluyen un identificador (id) que distingue a cada negocio de manera única, el nombre del negocio ($nombre$) para fines de identificación, la dirección física para una ubicación precisa ($dirección$) e información geográfica como latitud y longitud para facilitar una representación espacial precisa ($latitud$ y $longitud$). Además, las credenciales del negocio, que incluyen correo electrónico ($email$) y contraseña (pwd), se almacenan para garantizar seguridad y autenticación en las interacciones con el portal.

Todos los negocios satélite en la ciudad que deseen ofrecer actividades de ocio podrán registrarse en el portal. De esta manera, existirá un número de negocios satélite en el portal, a los que nos referiremos como LNS ($LNS = \{NS_1, NS_2, \ldots, NS_n\}$).

Cada negocio satélite (NS_i) podrá proponer cuantas actividades de ocio desee, éstas irán desde la asistencia a un evento deportivo hasta la asistencia a una obra de teatro, una exposición o un concierto, pasando por actividades tan cotidianas como disfrutar de una cena, una comida o un aperitivo. En general, actividades de ocio destinadas a cautivar y mejorar la experiencia de un usuario. El portal tendrá una lista de actividades de ocio, nos referiremos a ella como LAC, donde se podrá almacenar y proporcionar la información necesaria para cada actividad de ocio (AC_i).

Cada actividad de ocio (AC_i) tendrá un identificador único (id) y una descripción detallada de la actividad ($descripcion$), estando asociada al negocio satélite que la ofrece (id_{ns}) y relacionada con una o varias de las categorías de productos existentes en el portal ($categoria$). Esta categorización permite a los usuarios explorar fácilmente una amplia gama de actividades, y será de utilidad a la hora de elaborar planes de ocio para usuarios basados en sus intereses. La inclusión de la fecha en la que tendrá lugar el evento ($fecha$), junto con las horas de inicio y fin precisas ($inicio$ y fin), es también fundamental a nivel informativo y para elaborar los planes. También se tiene en cuenta el precio de la actividad de ocio ($precio$). Formalmente, una actividad de ocio (AC) se puede definir respectivamente como sigue:

$$AC = (id, descripcion, id_{ns}, categoria, fecha, inicio, fin, precio) \quad (2)$$

De esta manera, existirá un número de actividades de ocio en el portal, al que nos referiremos como LAC (es decir, $LAC = \{AC_1, AC_2, \ldots, AC_m\}$).

La idea del portal es promover, siempre y cuando sea posible, que el comprador vaya caminando a comprar (o recoger) los productos que ha adquirido a través del portal a un negocio local o punto de recogida y a la vez pueda disfrutar de un plan de ocio recomendado por el portal en función del perfil del comprador, de la compra realizada y del punto de compra o recogida y momento de recogida seleccionado.

Los puntos de recogida son lugares en los que los usuarios podrán recoger sus productos comprados a través del portal. Estos puntos pueden ser seleccionados por el usuario o recomendados por el portal. Formalmente, los puntos de recogida (PR) pueden definirse como:

$$PR = (id, nombre, latitud, longitud, email, pwd) \quad (3)$$

almacenando su identificador único (id), el nombre del punto ($nombre$), sus coordenadas geográficas ($latitud$ y $longitud$) y su información de credenciales ($email$ y pwd).

Lo primero que hará el portal será comprobar si el usuario que ha realizado la compra pertenece a la nueva generación de consumidores: responsables y críticos con el proceso de compra. Esto se sabe en función de su historial de compras. Si es así, le mostrará primero en la lista de resultados en sus búsquedas, los productos que se venden en locales cercanos a él y/o con recogida en puntos de la localidad.

Si el usuario decide comprar o recoger en local, el portal elaborará para él un plan de ocio en función de la oferta de actividades que existen en la localidad y el perfilado del usuario.

III. Perfilado de usuario

El perfilado de usuarios se obtendrá mediante un algoritmo de *clustering* difuso, que obtendrá un conjunto de agrupaciones de usuarios similares. Esta agrupación se realizará en función de diferentes aspectos de los usuarios, aspectos clave para la elaboración de planes, como por ejemplo las preferencias del usuario en las diferentes categorías de productos que se venden en el portal, la edad y el gasto que realizan en el portal. Hacer el perfilado del usuario, es decir comprender qué compran o qué le interesa del portal, cómo se comportan los usuarios en función de su edad o del gasto que realizan en el portal es importante para el análisis de la similitud entre usuarios y para el funcionamiento de los sistemas de recomendación (tratándose de un filtrado colaborativo en este caso) y para la generación y recomendación de planes en particular. Veamos cómo se hace ese perfilado de usuarios.

Primero, calculamos las diferentes matrices de similitud para las características relevante del usuario (es decir, sus preferencias, su edad y su gasto previo en el portal). En este caso, calculamos las siguientes matrices de similitud:

- **Preferencias del usuario en categorías de productos del portal** (MP). Esta matriz captura la similitud entre usuarios basada en sus preferencias en las diferentes categorías de productos de C (es decir, $C = \{c_1, c_2, \ldots, c_r\}$). Previamente, se obtendrá para cada usuario su matriz de preferencias en las diferentes categorías de productos del portal. Estas pueden ser explícitas o implícitas, ambas con valores normalizados. Las preferencias explícitas se obtienen de la retroalimentación directa del usuario

sobre las categorías de interés, obtenidas a través de un cuestionario que se le realiza cuando se da de alta en el portal, y generando un vector EP de dimensión $|C|$. Las preferencias del usuario u_x para cada categoría c_i de C se almacenan en $EP_{u_x}[i]$. Se utilizará un valor de 0 en las categorías donde el usuario no tenga interés y un valor mayor que 0 en aquellas en las que el usuario sí tenga algún interés (es decir, $EP_{u_x}[i] \in [0, 1]$).

Por otro lado, las preferencias implícitas se obtendrán a partir de su comportamiento en el portal, teniendo en cuenta aspectos como el historial de compras (HC), manipulación de productos en 3D ($M3D$), visualización de productos (VP) y teleportación ($TE3D$) del usuario en cada categoría $c_i \in C$. Las preferencias implícitas en cada categoría se almacenan en $IP_{u_x}[i]$ y se calculan considerando los comportamientos del usuario en el portal sobre la categoría c_i de C ($BH_{u_x}[i]$) (ver ecuación 4), ponderado con coeficientes de importancia para cada aspecto (α para HC, β para M3D, γ para TE3D y θ para VP), como se puede observar en la Ecuación 5. De esta forma se puede dar mayor o menor importancia a cada uno de los aspectos considerados.

$$BH_{u_x}[i] = HC_{u_x}(c_i) \times \alpha + M3D_{u_x}(c_i) \times \beta + VP_{u_x}(c_i) \times \gamma + TE3D_{u_x}(c_i) \times \theta \quad (4)$$

$$IP_{u_x}[i] = \frac{BHu_x[i]}{\sum_{j=1}^{C} BH_{u_x}[j]} \quad (5)$$

Para obtener la matriz de preferencias del usuario u_x en las categorías del portal C, denominada UP_{u_x}, se considerarán las preferencias explícitas EP_{u_x} y las implícitas IP_{u_x} del usuario ponderadas conforme a la siguiente ecuación 6.

$$UP_{u_x}[i] = \alpha_{IP} \times IP_{u_x}[i] + \alpha_{EP} \times EP_{u_x}[i] \quad (6)$$

donde α_{IP} y α_{EP} son dos parámetros que nos permiten determinar la importancia de cada una de las dos preferencias implícitas y explícitas, respectivamente.

Una vez calculadas las matrices de preferencias de cada usuario en las categorías del portal UP_{u_x} ya se puede calcular la matriz que captura la similitud entre usuarios del portal, se denomina MP y se calcula de acuerdo con la ecuación 7.

$$MP[u_x, u_y] = \sum_{i=1}^{|C|} |UP_{u_x}[i] - UP_{u_y}[i]| \quad (7)$$

- **Edad del Usuario** (ME). Esta matriz captura la similitud entre usuarios según su edad y se calcula empleando la similitud coseno, tal y como se indica en la ecuación 8. Se emplea lógica difusa para definir el grupo de edad al que pertenece cada usuario, empleando los valores lingüísticos: CE={'niño', 'adolescente', 'joven', 'adulto', 'mediana edad' ,'mayor'} (ver Figura 1).

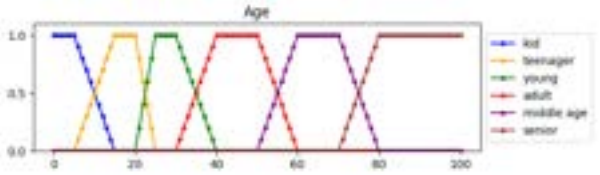


Fig. 1. Función de membresía de la variable *edad* (μ).

$$ME(u_x, u_y) = \frac{\sum_{i=1}^{|CE|} \mu_i(E(u_x)) \times \mu_i(E(u_y))}{\sqrt{\sum_{i=1}^{|CE|} \mu_i(E(u_x))^2} \sqrt{\sum_{i=1}^{|CE|} \mu_i(E(u_y))^2}} \quad (8)$$

donde E(u_x) es una función que devuelve la edad numérica del usuario u_x.

- **Gasto Monetario del Usuario** (MG). Esta matriz almacena la similitud entre los usuarios en función de sus patrones de gasto en el portal dentro de cada categoría de producto del portal utilizando la similitud del coseno (ver ecuación 9). Se vuelven a emplear valores lingüísticos para categorizar a los usuarios en niveles de gasto en cada categoría $c_i \in C$: CG_i= {'bajo', 'medio', 'alto'} (ver Figura 2, donde G(u_x,i) es una función que devuelve el gasto monetario realizado por el usuario u_x en productos de la categoría c_i). La definición de estos conjuntos se realizará de manera automática mensualmente en función del gasto medio que se realice en el portal y en cada categoría.

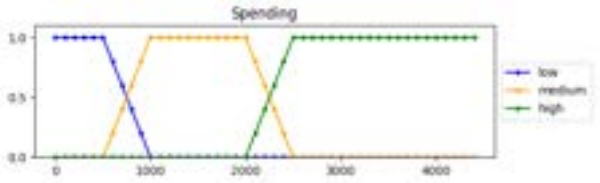


Fig. 2. Función de membresía de la variable *gasto* (ρ_i), siendo i una categoría.

A partir de estas matrices se calcula una matriz compuesta (M) utilizando la ecuación 10. Como se puede comprobar esta matriz se calcula a partir de una ponderación de las matrices previamente calculadas, donde la matriz de preferencias (MP) tendrá más importancia, y las matrices de edad (ME) y gasto (MG) serán empleadas en caso de haber poca información sobre las preferencias en categorías.

$$M[u_x, u_y] = \alpha_{MP} \times MP[u_x, u_y] + \alpha_{ME} \times ME[u_x, u_y] + \alpha_{MG} \times MG[u_x, u_y] \quad (10)$$

Una vez que se obtiene la matriz ponderada M de distancias entre usuarios, se aplica el algoritmo de *clustering* difuso c-means. Elegimos $|C|$ como el número de grupos que deseamos obtener (que coincidirá con el número de categorías disponibles en el portal). Cada grupo estará representado por su centroide.

Por otra parte para cada usuario tendremos un vector (GP_{u_x}) con dimensión $|C|$ que muestra los grados de pertenencia del usuario u_x a cada uno de los grupos (es decir, a cada categoría). En $GP_{u_x}[i]$ se almacena la pertenencia de u_x al grupo i e informa del interés del usuario en la categoría c_i.

IV. Generación de Planes de Ocio

Los planes de ocio se generan dinámicamente en función del perfilado del usuario, se comprueba a qué grupo pertenece para ver qué categorías les interesan y en función de esto construir un plan de ocio para él.

Cada plan de ocio (PO) comprenderá sugerencias personalizadas de actividades de ocio (ACs). PO se define como una colección de estas actividades (es decir, $PO = AC_x, AC_y, \ldots, AC_k$), cada una elegida para alinearse con los intereses únicos del usuario y el cronograma de las actividades (es decir, $ACs \in PO$ están ordenadas en función de sus tiempos ($inicio$ y fin)).

En esta sección se presenta un algoritmo diseñado específicamente para la generación de planes de ocio. Se sugerirá un plan de ocio ajustado al usuario, en función de su perfilado (es decir, del clúster al que pertenece) y las actividades de ocio disponibles.

Para iniciar el proceso, el primer paso implica determinar el clúster o clústers a los que pertenece el usuario u_x. Para ello se usará el vector GP_{u_x}. Se comprobará si el usuario está muy interesado en una única categoría c_i (es decir $GP_{u_x}[i]$ supera el umbral α). Si no es así, se verifica si la diferencia entre varios valores de GP_{u_x} es menor que otro umbral β. Si es así, el usuario está asociado con varios clústers; de lo contrario, el usuario se asigna al clúster con el grado de pertenencia más alto. Todo esto se realiza en el algoritmo 1.

Datos: GP_{u_x}, α, β
Resultado: Categoría/s relevantes para el usuario (CAT).
$max\_val = \max_{i=1\ldots|C|}(GP_{u_x}[i])$
$CAT = \emptyset$;
si $max\_val > \alpha$ **entonces**
  $CAT = CAT \cup \{i \mid \max_i\{GP_{u_x}[i]\}\}$
en otro caso
  para $i = 1$ *en* $|C|$ **hacer**
    si $|GP_{u_x}[i] - max\_value| < \beta$ **entonces**
      $CAT = CAT \cup \{i\}$
    fin
  fin
fin
devolver CAT

Algoritmo 1: Algoritmo de clúster dominante.

Con la información de las categorías dominantes del usuario u_x que ha realizado la compra (es decir, CAT), se ejecutará el algoritmo que genera el plan de ocio, teniendo en cuenta el punto de recogida, la fecha y hora en la que el usuario lo recogerá (ver algoritmo 2).

El algoritmo 2 emplea un proceso de filtrado multifacético. En primer lugar, asegura que las actividades de ocio se alineen con la fecha de recogida especificada por el usuario. Luego, incluye selectivamente actividades que caen bajo las categorías representadas por las preferencias dominantes del usuario (CAT). A continuación, el algoritmo calcula las distancias entre cada actividad de ocio y el punto de recogida designado. El enfoque incluye selectivamente solo actividades de ocio que estén dentro de una proximidad razonable, asegurando que las distancias calculadas no excedan el umbral especificado α. Esto mitiga el potencial malestar del usuario asociado con ubicaciones distantes. Finalmente, se excluyen cualquier actividad de ocio que ocurra antes de la hora de recogida especificada por el usuario.

Datos: CAT, α, β y punto de recogida PR, fecha y hora de la recogida, LAC
Resultado: Plan de Ocio PO.
$PO = LAC$;
para *cada* AC_i *en* PO **hacer**
  si $AC_i.fecha \neq fecha$ **entonces**
    $PO = PO - \{AC_i\}$
  fin
fin
para *cada* AC_i *en* PO **hacer**
  si $AC_i.categoría \cap CAT = \emptyset$ **entonces**
    $PO = PO - \{AC_i\}$
  fin
fin
para *cada* AC_i *en* PO **hacer**
  si $distancia(AC_i, PR) ==' lejos'$ **entonces**
    $PO = PO - \{AC_i\}$
  fin
fin
para *cada* AC_i *en* PO **hacer**
  si $AC_i.inicio < hora$ **entonces**
    $PO = PO - \{AC_i\}$
  fin
fin
$LP = organizar(LP, \beta)$
devolver LP;

Algoritmo 2: Algoritmo Generador de Planes de Ocio

Dentro del algoritmo 2, *distancia* es una función difusa que calcula la distancia entre el lugar donde se lleva a cabo la actividad de ocio y el punto de recogida distinguiendo

$$ME(u_x, u_y) = \frac{\sum_{i=1}^{|C|}\sum_{j=1}^{|CG|}\rho_{ij}(G(u_x,i)) \times \rho_{ij}(G(u_y,i))}{\sqrt{\sum_{i=1}^{|C|}\sum_{j=1}^{|CG|}\rho_{ij}(G(u_x,i))^2}\sqrt{\sum_{i=1}^{|C|}\sum_{j=1}^{|CG|}\rho_{ij}(G(u_y,i))^2}} \quad (9)$$

TABLE I
NEGOCIOS SATÉLITE Y PROPUESTAS DE ACTIVIDADES.

Nombre del negocio	**Categoría**	**Actividad de ocio**
Centro Ayurveda Fusionatur	Belleza	Sesión de masaje
Primor	Belleza	Curso de maquillaje
Biblioteca de Ciudad Real	Cultura	Sábado de videojuegos
Auditorio de la Granja	Cultura	Concierto: Pablo López
Antiguo Casino	Hogar	Casa y Jardín Expo 2023
Pabellón Ferial	Hogar	ExpoHogar
Plaza Mayor	Moda	Fashion Trend Showcase 2023
Moda re- Ciudad Real	Moda	Recogida de ropa para la caridad
Polideportivo Rey Juan Carlos	Deportes	Torneo provincial de natación
Quijote Arena	Deportes	Partido de balonmano Caserío CR vs Sinfin
Espacio Serendipia	Tecnología	Machines learning, humans on alert
Living Room	Tecnología	Math Street Fighter: Maths vs Humans
A'Pares	*	-
Entretapas	*	-

en 3 conjuntos: 'cerca', 'medio' y 'lejos'. Y *organizar* es una función que selecciona aleatoriamente un número de actividades de ocio diferentes de la lista y las organiza según sus tiempos de inicio y fin, asegurando que no se superpongan y que haya un margen de tiempo entre ellas para permitir el desplazamiento (β).

La implementación de estos pasos conduce a la creación de un plan de ocio personalizado basado en las preferencias del usuario, asegurando una experiencia atractiva y personalizada.

V. EJEMPLO ILUSTRATIVO DEL ALGORITMO PROPUESTO

Para este caso de uso específico dentro del portal, nuestro enfoque se reduce a 14 negocios satélites (NS) seleccionados en Ciudad Real, descritos en la tabla I. Cada ubicación actúa como un centro representativo, ofreciendo actividades de ocio únicas en seis categorías diferentes C - belleza, cultura, hogar, moda, deporte y tecnología. También hemos incluido dos establecimientos que se consideran restaurantes o cafés que pertenecen simultáneamente a múltiples categorías.

Presentamos al usuario X como un usuario representativo, ofreciendo una visión de sus compras e interacciones dentro del portal. El usuario X, llamado Ana, una joven de 25 años interesada en deportes y tecnología, nos ofrece una perspectiva detallada para evaluar qué tan efectivo es nuestro sistema.

Después de recopilar los datos de Ana, procedemos a construir su perfil de usuario utilizando el algoritmo de *clustering*. Los resultados obtenidos para cada grupo se presentan en la Tabla II.

TABLE II
RESULTADOS INICIALES DEL GRADO DE PERTENENCIA (GP).

Belleza	Cultura	Moda	Hogar	Deportes	Tecnología
0.112596	0.076077	0.081246	0.075429	0.340955	0.313697

Como era de esperar, se observan valores más altos en los grupos asociados con deportes y tecnología, reflejando con precisión las preferencias iniciales de Ana. Después de varias interacciones con productos de interés, como vistas de artículos relacionados con deportes y tecnología, así como manipulaciones en 3D para cada uno, Ana finalmente decide comprar un producto tecnológico. Utilizando el algoritmo 1, Ana tiene 2 grupos dominantes: deportes y tecnología (ver Tabla II). La compra del producto desencadena diferentes actividades de ocio, teniendo en cuenta las restricciones temporales, la distancia desde el punto de recogida hasta el negocio satélite y el perfil del usuario, los resultados se organizan en la Tabla III. El resultado para el primer plan de ocio es "Espacio Serendipia" de la categoría de tecnología.

TABLE III
PROPUESTAS PARA LA PRIMERA ACTIVIDAD DE OCIO.

Negocio satélite	**Actividad de ocio**	**Categoría**
Espacio Serendipia	Machines learning, humans on alert	Tecnología
Living Room	Math Street Fighter: Maths vs Humans	Tecnología
Polideportivo Rey Juan Carlos	Torneo provincial de natación	Deportes
Bar Entretapas	-	*
Quijote Arena	Partido de balonmano Caserío CR vs Sinfin	Deportes
A Pares	-	*

Posteriormente, tiene lugar la generación del segundo plan de ocio. Los resultados se representan en la Tabla IV. En este caso, se tiene en cuenta la distancia desde la primera actividad de ocio generada, el perfil del usuario y las restricciones temporales. Como resultado, "Bar Entretapas" se presenta como la opción recomendada, ofreciendo una secuencia de actividades personalizada que no solo se ajusta a las preferencias de Ana, sino que también considera aspectos prácticos para una experiencia fluida y agradable.

TABLE IV
PROPUESTAS PARA LA SEGUNDA ACTIVIDAD DE OCIO.

Negocio satélite	**Actividad de ocio**	**Categoría**
Bar Entretapas	-	*
Quijote Arena	Partido de balonmano Caserío CR vs Sinfin	Deportes
A Pares	-	*

Finalmente, el plan de ocio generado para Ana consiste en recoger el producto comprado en el punto de recogida "PCBox", seguido de una visita a "Espacio Serendipia" para una charla sobre tecnología. Posteriormente, hay oportunidad de relajarse y disfrutar de refrescos en "Bar Entretapas". Este itinerario cuidadosamente diseñado se visualiza en la Figura 3, presentando una secuencia de actividades fluida

Fig. 3. Representación de las actividades de ocio.

adaptada a las preferencias de Ana. Esta serie de actividades no solo garantiza la satisfacción de Ana al coincidir con sus intereses, sino que también impulsa la economía local al fomentar la participación en eventos cercanos. Este enfoque de recomendación personalizada, adaptado a las preferencias y comportamientos de Ana, resalta la capacidad de nuestro modelo para ofrecer experiencias únicas para cada usuario.

VI. Conclusiones y trabajo futuro

Este artículo presenta una idea innovadora para la evolución de los portales de comercio electrónico, con el objetivo de apoyar a las pequeñas y medianas empresas (PYMEs), fomentar las economías locales y reducir el impacto medioambiental del comercio electrónico. La integración de "negocios satélite" diversifica el enfoque tradicional centrado en el producto del comercio electrónico, fomentando una relación sinérgica entre las compras en línea y las actividades de ocio físicas. Los usuarios pueden recoger compras en puntos locales específicos, simplificando la logística y ayudando al medio ambiente. Esto también promueve la generación de "planes de ocio" personalizados que tienen en cuenta las ofertas de actividades de ocio que realizan estos negocios satélite.

Como un esfuerzo de investigación futuro, se propone mejorar la función de perfilado, teniendo en cuenta también las experiencias o planes aceptados previamente por el usuario y el feedback proporcionado sobre ellas por él u otros usuarios similares y la organización de actividades dentro del algoritmo 2 (*organizar*). El objetivo es desarrollar un método más sofisticado que optimice la organización de actividades de ocio, en función de la proximidad, las temporalidad, el histórico y el feedback. Esta mejora busca enriquecer aún más la experiencia de los usuarios proporcionando recomendaciones de ocio más personalizadas alineadas con sus preferencias específicas.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio español de Ciencia e Innovación MICIN/AEI/10.13039/501100000033, y la Unión Europea (NextGenerationEU/PRTR), en el marco del Proyecto de Investigación TED2021-131082B-I00.

Referencias

[1] V. Jain, B. Malviya, and S. Arya, "An overview of electronic commerce (e-commerce)," *Journal of Contemporary Issues in Business and Government*, vol. 27, no. 3, pp. 665–670, 2021.

[2] Ministerio de Asuntos Económicos y Transformación Digital, Gobierno de España, "Sme digitalisation plan 2021-2025," https://portal.mineco.gob.es/RecursosArticulo/mineco/\ministerio/ficheros/210902-digitalisation-smes-plan.pdf, 2021, accessed: 15/12/2023.

[3] V. Pavlova, T. Murovana, N. Sarai, V. Velychko, K. Illyashenko, and H. Hrysshyna, "Crisis phenomena as an incentive to intensify e-commerce of the enterprise," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 19, pp. 4646–4657, 2021.

[4] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273, 2015.

[5] A. Goli, H. Khademi-Zare, R. Tavakkoli-Moghaddam, A. Sadeghieh, M. Sasanian, and R. M. Kordestanizadeh, "An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: a case study," *Network: Computation in Neural Systems*, vol. 32, no. 1, pp. 1–35, 2021.

[6] L. Lundström, "Dynamic pricing services in e-commerce ecosystems: An exploratory study of context, technologies, and practices," 2021.

[7] N. Xi and J. Hamari, "Shopping in virtual reality: A literature review and future agenda," *Journal of Business Research*, vol. 134, pp. 37–58, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0148296321003209

[8] R. Grande, S. Sánchez-Sobrino, D. Vallejo, J. J. Castro-Schez, and J. Albusac, "Supporting small businesses and local economies through virtual reality shopping and artificial intelligence: A position paper," in *Proceedings of the 25th International Conference on Enterprise Information Systems - Volume 2: ICEIS*, INSTICC. SciTePress, 2023, pp. 312–319.

Una aplicación de conjuntos difusos a la asignación óptima de tareas

G. Jaume-Martin[1,2]
[1]*Depto. de Matemáticas e Informática*
Universidad de las Islas Baleares
[2]*Instituto de Investigación Sanitaria*
Illes Balears (IdISBa)
Hospital Universitari Son Espases
Palma de Mallorca, España
gabriel.jaume@uib.cat

J. Antich[1,2]
[1]*Depto. de Matemáticas e Informática*
Universidad de las Islas Baleares
[2]*Instituto de Investigación Sanitaria*
Illes Balears (IdISBa)
Hospital Universitari Son Espases
Palma de Mallorca, España
javier.antich@uib.es

O. Valero[1,2]
[1]*Depto. de Matemáticas e Informática*
Universidad de las Islas Baleares
[2]*Instituto de Investigación Sanitaria*
Illes Balears (IdISBa)
Hospital Universitari Son Espases
Palma de Mallorca, España
o.valero@uib.es

***Abstract*—Uno de los principales retos al trabajar con sistemas multi-robot consiste en asignar, en cada instante de tiempo, la tarea a realizar más adecuada para cada robot. En muchas situaciones reales, estas tareas tienen ligados unos plazos de tiempo límite, que se tienen que cumplir. Dada su probada efectividad a la hora de lidiar con el problema que nos concierne, en este artículo, nos inspiramos en los métodos de umbral de respuesta ("*Response-Threshold Methods*"), los cuales son un caso particular de los métodos de enjambre ("*Swarm-like Methods*"). Siguiendo esta metodología, los robots deciden cuál es la tarea que deben realizar basándose en los estímulos procedentes tanto de la tarea que están llevando a cabo en ese momento, como de la tarea a la cual consideran desplazarse. De esta forma, planteamos una alternativa innovadora a las estrategias de asignación de tareas existentes actualmente en la literatura, que tiene la capacidad de manejar tareas con plazos de tiempo límite asociados. Más concretamente, los estímulos percibidos por los robots son modelados mediante conjuntos difusos, de manera que el robot es capaz de determinar cuál es la tarea más apropiada en cada momento a través de la célebre técnica de optimización difusa de Bellman y Zadeh.**

Con el objetivo de poder evaluar el nuevo enfoque matemático, hemos llevado a cabo un extenso número de simulaciones. Así, hemos podido comprobar que la alternativa propuesta es capaz de modelar la evolución de un sistema de enjambre cuando existen tareas con plazos límite de ejecución.

***Index Terms*—sistemas multi-robot, problema de asignación de tareas, conjuntos difusos, técnica de optimización difusa de Bellman-Zadeh.**

I. Introducción

En los últimos años, los sistemas multi-robot han ganado popularidad y los continuos avances en este campo han permitido una amplia serie de aplicaciones. No obstante, los problemas ligados a los sistemas multi-robot tienden a ser desafiantes y matemáticamente complejos. En este artículo, nos centraremos en uno de los problemas más importantes que pueden surgir al trabajar en esta temática: la asignación de tareas en sistemas multi-robot ("*Multi-Robot Task Allocation*" o MRTA en su forma abreviada).

Esta investigación forma parte del proyecto PID2022-139248NB-I00 financiado por MICIU/AEI/10.13039/501100011033 y "FEDER Una manera de hacer Europa".

La MRTA consiste en asignar a cada grupo de robots las tareas a realizar teniendo en cuenta, por una parte, las características de los robots y, por otra, las características de las tareas. Dado que este problema es típicamente NP-complejo, la mayoría de los algoritmos MRTA están diseñados para tener una gran eficiencia computacional, a la vez que proporcionan resultados tan próximos como sea posible a los óptimos. Entre las distintas metodologías desarrolladas por los investigadores en los últimos años, destacan las metodologías MRTA descentralizadas. En particular, el trabajo presentado en este artículo está inspirado en los enfoques tipo enjambre, más concretamente en los métodos de umbral de respuesta ("*Response-Threshols Methods*" o RTMs). Los RTMs asumen que cada tarea está asociada a un estímulo, del cual los agentes son capaces de captar información. Posteriormente, se sigue un proceso de toma de decisiones probabilístico, dando una mayor probabilidad de transición a aquellas tareas que generan un estímulo de mayor intensidad. Típicamente, estos procesos de toma de decisión han sido modelados matemáticamente mediante cadenas de Markov. Es importante destacar que estos métodos son bastante simples, lo cual restringe el rango de tareas que pueden realizar.

En este artículo, el modelo matemático que proponemos ha sido desarrollado aprovechando el marco teórico basado en conjuntos difusos introducido en [1]–[3]. En las referencias mencionadas, se utilizan conjuntos difusos para modelar las funciones de respuesta y la asignación de tareas se lleva a cabo mediante cadenas de Markov posibilísticas. Este enfoque presenta diversas ventajas frente a la aproximación clásica probabilística. Entre ellas, cabe destacar las siguientes. Por un lado, la convergencia del sistema a una distribución estacionaria se realiza en un número finito de pasos, y no de modo asintótico. Por otro lado, la predicción del comportamiento del sistema mejora enormemente en presencia de vaguedad en las mediciones tomadas por parte de los robots de las diferentes características del sistema como, por ejemplo, las distancias entre tareas. Inspirados por estos hechos, en la estrategia de asignación de tareas para RTMs que presentamos, los estímulos percibidos por los robots son modelados a

través de conjuntos difusos como en las referencias [1]–[3]. Sin embargo, siguiendo con el planteamiento dado en [4], nuestro enfoque trata con tareas que tienen plazos de tiempo límite y el proceso de toma de decisiones es implementado como un problema de optimización. Así, permitimos que la decisión en cada momento del tiempo sea obtenida mediante la técnica de optimización de Bellman y Zadeh [5]. Este enfoque, utilizando conjuntos difusos, tiene numerosas ventajas, de entre las cuales destacan las siguientes: la sencillez del sistema de comunicación requerido (compatible con el de cualquier sistema de enjambre); el bajo coste computacional del proceso de toma de decisión; y la capacidad de poder modelar objetivos y restricciones de nuestro problema que no estén claramente definidos en términos precisos. Los resultados obtenidos de los experimentos de simulación realizados demuestran que el enfoque propuesto es capaz de modelar eficazmente la evolución de un sistema de enjambre, cuando se tienen en cuenta plazos límite de tiempo para las tareas. Cabe decir que nuestro planteamiento presenta ciertas diferencias con el establecido en [4], las cuales serán clarificadas en la Sección III.

Con respecto al resto del artículo, en la Sección II ofrecemos una explicación más detallada del proceso de toma de decisiones. Posteriormente, en la Sección III presentaremos la formulación matemática del problema que intentamos abordar y el modelo propuesto para la estrategia MRTA. En la Sección IV, se evalúa el rendimiento del modelo mediante un extenso número de simulaciones, analizando los resultados obtenidos. Finalmente, en la Sección V, resumimos las conclusiones que hemos podido sacar e indicamos algunas propuestas de trabajo futuro.

II. Proceso de toma de decisiones

El objetivo de esta sección es introducir la metodología para la toma de decisiones en entornos difusos que hemos incorporado en nuestra estrategia MRTA. En particular, hablamos de la célebre técnica de optimización de Bellman y Zadeh, la cual hace uso de conjuntos difusos y funciones de agregación. Asumimos que el lector está familiarizado con las nociones básicas de esta técnica (véase [6]). Aún así, resumiremos su funcionamiento para facilitar el entendimiento de las secciones posteriores.

De acuerdo con [5], consideremos un objetivo G a alcanzar y una restricción C que limite su consecución. Denotemos por $\widehat{G}$ el objetivo difuso (versión difusa del objetivo), por $\widehat{C}$ la restricción difusa (versión difusa de la restricción) y por $\widehat{D}$ la decisión difusa (versión difusa del conjunto de soluciones del problema bajo estudio). Sea X el conjunto de todas las posibles alternativas, conteniendo la decisión óptima. Sean $\mu_{\widehat{G}} : X \to [0,1]$ y $\mu_{\widehat{C}} : X \to [0,1]$ las funciones de pertenencia de $\widehat{G}$ y $\widehat{C}$, respectivamente. Entonces, los valores de $\mu_{\widehat{G}}(x)$ y $\mu_{\widehat{C}}(x)$ con $x \in X$ pueden ser interpretados como el nivel con el que la alternativa x satisface el objetivo G y la restricción C.

Por supuesto, la solución al proceso de toma de decisiones tiene que satisfacer simultáneamente el objetivo y la restricción. Esto se modela en el entorno difuso a través del conjunto decisión difusa $\widehat{D}$. Con este objetivo, dicho conjunto se describe mediante una función de pertenencia obtenida mediante la agregación de las funciones de pertenencia del objetivo y de la restricción. Así pues, dada una función de agregación F (véase [7] para una exposición detallada sobre la teoría de funciones de agregación), la función de pertenencia de una decisión difusa $\widehat{D}$ sobre X, $\mu_{\widehat{D}} : X \to [0,1]$, se obtiene mediante $\mu_{\widehat{D}}(x) = F(\mu_{\widehat{G}}(x), \mu_{\widehat{C}}(x))$ para todo $x \in X$. Notemos que la función de agregación es simétrica, de manera que el papel que juegan el objetivo y la restricción son intercambiables. Este hecho resulta fundamental en aquellos casos en los que la frontera entre ellos no es clara y, por ese motivo, ambos pueden ser tratados simplemente como aspectos del problema sin establecer una distinción precisa entre ellos.

Recordemos que, a pesar que la solución difusa proporciona el nivel de adecuación de cada una de las diferentes alternativas, es necesario seleccionar la solución más apropiada para el problema bajo consideración. Dado que $\mu_{\widehat{D}}(x)$ proporciona para cada $x \in X$ el nivel de satisfacción combinado de $\mu_{\widehat{G}}(x)$ y $\mu_{\widehat{C}}(x)$, en [5] la metodología de Bellman y Zadeh establece que la alternativa óptima $x^* \in X$ es la que maximiza la función de pertenencia de $\widehat{D}$, en otras palabras, $x^* = \arg\max(\mu_{\widehat{D}})$.

III. Formulación matemática del problema

En lo que sigue, vamos a describir el problema de asignación de tareas que nos proponemos abordar, así como el enfoque matemático que usaremos para modelar el comportamiento individual de los robots, que conducirá a la colaboración necesaria de la que emergerá la inteligencia enjambre que permitirá solucionar dicho problema.

A. Descripción del problema

Sea $R = \{a_1, ..., a_n\}$ el conjunto de todos los robots y $T = \{t_1, ..., t_m\}$ el conjunto de todas las tareas que se tienen que llevar a cabo, siendo $n, m \in \mathbb{N}$ y $n \leq m$. En cada instante de tiempo t, considerando el tiempo discreto, necesitamos asignar a cada tarea $t_j \in T$ un subconjunto de robots $RA(t) \subseteq R$. El objetivo del sistema de enjambre es completar cada tarea cumpliendo con el plazo de tiempo límite que tiene asociado, que denotaremos por $t_{j,dl} \in \mathbb{N}$. Para alcanzarlo, se recompensa al sistema con una cantidad variable de utilidad por cada tarea completada. Si la tarea se consigue completar en un tiempo menor o igual a $t_{j,dl}$, entonces se proporciona la máxima utilidad posible para esa tarea, $t_{j,u} \in [0,1]$. Por contra, si se completa en un tiempo superior a $t_{j,dl}$, se logra una proporción (menor) de $t_{j,u}$.

Adicionalmente, denotaremos por $a_{k,wc} \in \mathbb{N}$ a la cantidad de trabajo que es capaz de realizar el robot a_k por unidad de tiempo. Cabe destacar que $a_{k,wc}$ es independiente de la tarea a ejecutar. Así, supondremos conocido el tiempo, $t_{j,et} \in \mathbb{N}$, que necesitaría un único robot con capacidad de trabajo igual a uno, $a_{k,wc} = 1$, para completar la tarea t_j. En consecuencia, el tiempo necesario para ejecutar la tarea t_j dependerá tanto

de $t_{j,et}$ como del conjunto de robots que estén trabajando en ella.

Al inicio, cada robot tendrá una ubicación específica, la cual podrá ser la misma para todos los robots, o bien cada robot podrá tener una posición diferente dentro del entorno. En cada instante de tiempo posterior, se utilizará el proceso de optimización difuso descrito en la Sección II para determinar la tarea más adecuada para cada robot, en el instante de tiempo t. Para poder llegar a las tareas, cada robot es capaz de desplazarse a una velocidad constante $a_{k,v} \in \mathbb{N}$.

La idoneidad con la que percibe el robot cada tarea en un instante de tiempo t se obtiene a partir de tres estímulos distintos: la capacidad de terminar la tarea antes de su plazo límite —*urgencia*— sdl_{a_k,t_j} (interpretado como un objetivo del problema); la cantidad máxima de robots que pueden trabajar simultáneamente en una misma tarea —*interferencia física*— I_{L,t_j}; y la voluntad de un robot de quedarse en la tarea actual hasta ejecutarla por completo —*inercia*— IN_{cl,t_j}, siendo cl la ubicación actual del robot (I_{L,t_j} e IN_{cl,t_j} interpretados como restricciones del problema). Esta información es percibida por cada robot a_k, en cualquier instante de tiempo t, como un único estímulo global $s_{k,cl,j}(t)$ calculado usando una función de agregación. Una vez el robot a_k dispone del estímulo procedente de cada tarea, este aplica la técnica de optimización de Bellman y Zadeh para seleccionar la tarea a realizar. Esto significa que la tarea seleccionada t_{j^*} será la que satisfaga $s_{k,cl,j^*}(t) = \max_{j\in\{1,\dots,m\}} s_{k,cl,j}(t)$.

Notemos que, aunque los robots sean heterogéneos en términos de velocidad y capacidad de trabajo por unidad de tiempo, todos ellos reciben y procesan los estímulos de la misma manera, tomando la decisión óptima del mismo modo, es decir, a partir de la misma función de agregación.

B. Modelo Matemático

Ahora, ya estamos en condiciones de introducir las expresiones matemáticas para todos los estímulos previamente mencionados y la metodología exacta para la toma de decisiones en entornos difusos. No obstante, en primer lugar, definiremos el conjunto difuso $U_{t_j} : \mathbb{N} \to [0, t_{j,u}]$ que nos proporciona la utilidad conseguida por el sistema al completar la tarea t_j en el instante de tiempo t del modo siguiente:

$$U_{t_j}(t) = \begin{cases} t_{j,u} & \text{si } t \leq t_{j,dl} \\ t_{j,u} \cdot \frac{0.07 \cdot t_{j,dl}}{(t - t_{j,dl}) + 0.07 \cdot t_{j,dl}} & \text{si } t > t_{j,dl} \end{cases} \quad (1)$$

Seguidamente, presentamos las expresiones matemáticas propuestas para cada uno de los tres estímulos considerados: *urgencia*, *interferencia física* e *inercia*. Por un lado, la *urgencia*, al igual que sucede con la utilidad, tiene que decrecer a medida que el tiempo de ejecución requerido para completar una tarea se aleja más allá del plazo límite. Por otra parte, en el caso de que una tarea no haya sido completada, el estímulo tiene que ser cada vez mayor a medida que se acerca el plazo límite, puesto que el tiempo del que se dispone para terminarla es cada vez menor. Siguiendo esta idea, el conjunto difuso para la urgencia, $sdl_{a_k,t_j} : \mathbb{N} \to [0, t_{j,u}]$, viene dado por:

$$sdl_{a_k,t_j}(t) = \begin{cases} t_{j,u} \cdot \frac{t_{j,dl}}{(t_{j,dl} - (t + ft_{a_k,t_j}(t))) + t_{j,dl}} & \text{si } t + ft_{ak,t_j}(t) \leq t_{j,dl} \\ & \text{y } t_{j,ret}(t) > 0 \\ t_{j,u} \cdot \frac{0.07 \cdot t_{j,dl}}{(t + ft_{a_k,t_j}(t) - t_{j,dl}) + 0.07 \cdot t_{j,dl}} & \text{si } t + ft_{ak,t_j}(t) > t_{j,dl} \\ & \text{y } t_{j,ret}(t) > 0 \\ 0 & \text{si } t_{j,ret}(t) = 0 \end{cases} \quad (2)$$

donde:

- $ft_{a_k,t_j}(t) = \frac{d_{cl,j}}{a_{k,v}} + \frac{t_{j,ret}(t)}{a_{k,wc}}$ es el tiempo que le llevaría al robot ir desde su ubicación actual hasta la tarea t_j y, también, terminar individualmente esa tarea.
- $d_{cl,j}$ es la distancia entre la ubicación actual del robot y la tarea t_j.
- $t_{j,ret}(t)$ es el tiempo restante para finalizar la tarea t_j, en el instante de tiempo t. Se puede calcular restando a $t_{j,et}$ el total de trabajo realizado por los robots sobre la tarea t_j hasta el instante de tiempo t.

Notemos que, en las expresiones (1) y (2), se ha utilizando un caso particular de la llamada casi-métrica difusa estándar (véase, por ejemplo, [8]). Observemos también que la expresión (2) supone una diferencia importante entre nuestro enfoque y el llevado a cabo en [4], puesto que, en lugar de utilizar conjuntos difusos distintos para modelar los estímulos de distancia y urgencia para posteriormente agregarlos, utilizamos un único conjunto difuso, el cual codifica información de ambos estímulos, pero no se deriva a priori de ningún proceso de agregación.

El siguiente estímulo que detallaremos es la *interferencia física*, I_{L,t_j}. Supondremos que existe una cantidad máxima de robots que pueden trabajar simultáneamente en una tarea t_j, que denotaremos por $t_{j,pt} \in \mathbb{N}$. Entonces, el conjunto difuso considerado para $I_{L,t_j} : N_{t_j} \to [0, 1]$, siendo $N_{t_j} = \{0, 1, ..., t_{j,pt}\}$ y $x_{t_j}(t) \in N_{t_j}$ el número de robots trabajando en la tarea t_j en el instante de tiempo t, viene dada por:

$$I_{L,t_j}(x_{t_j}(t)) = 1 - \frac{x_{t_j}(t)}{t_{j,pt}} \quad (3)$$

Solo falta por detallar el estímulo de la *inercia*, $IN_{cl,t_j}(t)$, entendido como la necesidad que siente un robot de quedarse en su posición actual en el instante de tiempo t. El conjunto difuso propuesto para $IN_{cl,t_j} : \mathbb{N} \to \{0, p\}$, donde $p \in (0, 1]$, es:

$$IN_{cl,t_j}(t) = \begin{cases} p & \text{si } cl = t_j \text{ y } t_{j,ret}(t) > 0 \\ 0 & \text{de lo contrario} \end{cases} \quad (4)$$

Ahora, ya estamos en condiciones de explicar cómo se obtiene el estímulo global único que percibe un robot a partir de la agregación de los estímulos de *urgencia*, *interferencia física* e *inercia*. Sea $A : [0, 1] \times [0, 1] \to [0, 1]$ una función de agregación con la particularidad que $A(x, 0) = A(0, x) = 0 \ \forall x \in [0, 1]$; es decir, 0 es un elemento absorbente para A (véase [6] o [7]). Nótese que imponemos dicha condición para prevenir que un robot decida ir a una tarea completada o a una tarea en la que ya está trabajando la cantidad máxima de

robots permitida. Entonces, $\forall t \in \mathbb{N}$, el estímulo global viene dado por:

$$s_{k,cl,j}(t) = \max\{A\left(sdl_{a_k,t_j}(t), I_{L,t_j}(x_{t_j}(t))\right), IN_{cl,t_j}(t)\} \quad (5)$$

En este sentido, el robot a_k decidirá cuál es la tarea óptima, t_{j^*}, a ser ejecutada en el siguiente instante de tiempo siguiendo un caso particular de la técnica de optimización de Bellman y Zadeh, presentada en la Sección II.

IV. Evaluación experimental

En esta sección, nuestro objetivo será evaluar el rendimiento de un RTM basado en el modelo matemático formulado en la sección anterior que aborde la estrategia MRTA. Para poder hacerlo, empezaremos especificando cinco funciones de agregación, A, distintas. Estas serán versiones modificadas del mínimo, el producto, la media armónica (H_{mean}), la media ponderada ordenada ("*Ordered Weighted Averaging*" u *OWA*) y del máximo (*OWA* con peso $w_1 = 1$). Estas funciones toman su valor original siempre que $min\{x, y\} \neq 0$, y en otro caso forzamos a que valgan cero, es decir, consideramos versiones de las funciones de agregación con 0 como elemento absorbente debido a la razón ya indicada en la Sección III. Por último, cabe recalcar que para el cálculo de $s_{k,cl,j}(t)$ y $s_{k,cl,j^*}(t)$ en la sección anterior, *max* representa la versión ordinaria (sin modificar) de la función de agregación máximo.

Se ha llevado a cabo un extenso número de simulaciones, considerando una gran variedad de escenarios posibles, con el fin de poner a prueba el RTM propuesto. Estos escenarios consisten en conjuntos de robots y tareas heterogéneos, en el sentido que los robots poseen diferentes habilidades ($a_{k,wc}$ y $a_{k,v}$) y las tareas pueden variar en términos de $t_{j,u}$, $t_{j,et}$, $t_{j,dl}$ y $t_{j,pt}$. Además, consideraremos las siguientes métricas para poder medir la eficacia del RTM:

- TTC es el tiempo total requerido para terminar todas las tareas.
- U es la proporción media de la utilidad máxima conseguida. Para calcularla, utilizaremos la expresión: $U = \sum_{j=1}^{m} \frac{U_{t_j}}{t_{j,u} \cdot m}$, donde U_{t_j} indica la cantidad de utilidad lograda para la tarea t_j.
- D es la distancia total media recorrida por cada robot. Esta se puede obtener como $D = \sum_{k=1}^{n} \frac{a_{k,d}}{n}$, siendo $a_{k,d}$ la distancia total de viaje del robot a_k. Obsérvese que valores más bajos de D suponen un menor consumo de recursos.
- $TbDL$ determina, en promedio, el tiempo por debajo del plazo de tiempo límite. Este se puede calcular mediante $TbDL = \sum_{j=1}^{m} \frac{t_{j,dl} - t_{j,c}}{m}$, donde $t_{j,c}$ es el tiempo en el que la tarea t_j fue completada.

Conocer las limitaciones de un RTM basado en un sistema de enjambre es indispensable para poder entender el motivo de la selección de las métricas presentadas para evaluar su rendimiento. En situaciones reales, la duración de las baterías de los robots suele ser bastante limitada, de manera que D y TTC son dos de las métricas más importantes a considerar. Tenerlas en cuenta es la mejor forma de prevenir la recarga de baterías de los robots durante la ejecución del problema. Por otra parte, dado que tenemos plazos de tiempo límite para completar las tareas, es importante comparar cómo de cercano ha sido el rendimiento obtenido al ideal. Esta información puede ser estimada gracias a la combinación de las métricas U y $TbDL$.

TABLE I
Resultados de las medias de las métricas para las configuraciones mencionadas, específicamente en el caso FC.

	A	FC			
		$\overline{TTC}$	$\overline{U}$	$\overline{TbDL}$	$\overline{D}$
$p = 1$	min	48.22	98.99	32.21	1655.91
	$product$	59.90	92.96	28.26	22485.67
	H_{mean}	59.23	94.03	28.26	2485.67
	OWA: $w_1 = 1$	60.67	88.64	27.61	2546.10
	OWA: $w_1 = 0.75$	63.67	90.45	26.82	2766.10
	OWA: $w_1 = 0.50$	60.33	92.39	28.00	2515.55
$p = 0$	min	151.09	69.40	-24.33	9736.75
	$product$	531.96	20.99	-260.00	38323.15
	H_{mean}	368.43	21.55	-172.87	26158.67
	OWA: $w_1 = 1$	824.55	4.12	-523.49	59863.09
	OWA: $w_1 = 0.75$	805.77	18.25	-439.09	58687.02
	OWA: $w_1 = 0.50$	684.51	25.43	-328.87	49646.75

A. Resultados de las simulaciones

El código desarrollado para ejecutar todas las simulaciones se ha escrito utilizando la versión 3 del lenguaje Python[1]. Estas simulaciones se han llevado a cabo fijando $n = 10$ robots y $m = 20$ tareas, distribuyendo estos robots y estas tareas aleatoriamente en un entorno de 640×480 píxeles. Los valores concretos de los parámetros que configuran a los robots y a las tareas han sido seleccionados aleatoriamente dentro de los siguientes intervalos: $t_{j,u} \in [0.75, 1.00]$, $t_{j,et} \in [4, 30]$, $t_{j,dl} \in [2.5 \cdot t_{j,et}, 4.0 \cdot t_{j,et}]$, $t_{j,pt} = 3$, $a_{k,v} \in [65, 85]$ (las unidades utilizadas para la velocidad son píxeles por unidad de tiempo) y $a_{k,wc} \in \{1, 2\}$.

También se han planteado diferentes configuraciones con respecto al proceso de toma de decisiones, calculando mil simulaciones para cada configuración. Para empezar, hemos considerado la posibilidad de tener inercia completa o falta total de inercia, es decir, hemos configurado el estímulo IN_{cl,t_j} con $p = 1$ o $p = 0$ respectivamente. En relación al operador *OWA*, los diferentes conjuntos de pesos estudiados son: $w = \{1, 0\}$, $w = \{0.75, 0.25\}$ y $w = \{0.50, 0.50\}$, donde $w = \{w_1, w_2\}$. Finalmente, se han explorado dos situaciones distintas en términos de las capacidades en la toma de decisiones de los robots: FC (siendo los robots capaces de cambiar la decisión con respecto a la mejor tarea a ejecutar mientras se están moviendo de una tarea a otra) y RC (no permitiendo que los robots cambien de decisión con respecto a su destino mientras viajan). Los resultados obtenidos en las condiciones mencionadas pueden encontrarse en las Tablas I y II.

Empecemos notando que cuando $TbDL$ es negativo nos indica que, en promedio, el tiempo en el que se ha comple-

[1]El pseudocódigo, el código y otra información relevante puede ser encontrada en el siguiente repositorio de GitHub: https://github.com/BielJM1/MRTAOptima

TABLE II
RESULTADOS DE LAS MEDIAS DE LAS MÉTRICAS PARA LAS CONFIGURACIONES MENCIONADAS, ESPECÍFICAMENTE EN EL CASO RC.

	A	RC			
		$\overline{TTC}$	$\overline{U}$	$\overline{TbDL}$	$\overline{D}$
$p=1$	min	36.57	99.61	37.03	817.04
	$product$	33.89	99.65	37.47	671.15
	H_{mean}	34.33	99.65	37.37	692.42
	OWA: $w_1 = 1$	32.68	97.17	37.51	602.83
	OWA: $w_1 = 0.75$	33.61	99.67	37.56	655.92
	OWA: $w_1 = 0.50$	33.62	99.65	37.52	661.04
$p=0$	min	84.67	86.95	2.84	4916.95
	$product$	105.05	58.09	-20.16	6702.61
	H_{mean}	106.95	58.68	-19.75	6803.40
	OWA: $w_1 = 1$	102.82	29.97	-19.17	6558.48
	OWA: $w_1 = 0.75$	105.39	50.12	-21.01	6758.35
	OWA: $w_1 = 0.50$	103.77	57.13	-20.16	6623.49

TABLE III
RESULTADOS DE LAS MÉTRICAS SEGÚN EL MÉTODO DE EVALUACIÓN POR RANKING EN EL CASO FC.

	A	FC				
		$\overline{TTC}$	$\overline{U}$	$\overline{TbDL}$	$\overline{D}$	Sum
$p=1$	min	7	6	7	7	27
	$product$	9	9	9	9	36
	H_{mean}	8	8	8	8	32
	max	11	12	11	11	45
	OWA: $w_1 = 0.75$	12	11	12	12	47
	OWA: $w_1 = 0.50$	10	10	10	10	40
$p=0$	min	19	14	19	19	71
	$product$	21	22	21	21	85
	H_{mean}	20	21	20	20	81
	max	24	24	24	24	96
	OWA: $w_1 = 0.75$	23	23	23	23	92
	OWA: $w_1 = 0.50$	22	20	22	22	86

TABLE IV
RESULTADOS DE LAS MÉTRICAS SEGÚN EL MÉTODO DE EVALUACIÓN POR RANKING EN EL CASO RC.

	A	RC				
		$\overline{TTC}$	$\overline{U}$	$\overline{TbDL}$	$\overline{D}$	Sum
$p=1$	min	6	5	6	6	23
	$product$	4	4	4	4	16
	H_{mean}	5	3	5	5	18
	$OWA: w_1 = 1$	1	7	3	1	12
	$OWA: w_1 = 0.75$	2	1	1	2	6
	$OWA: w_1 = 0.50$	3	2	2	3	10
$p=0$	min	13	13	13	13	52
	$product$	16	16	16	16	64
	H_{mean}	18	15	15	18	66
	$OWA: w_1 = 1$	14	19	14	14	61
	$OWA: w_1 = 0.75$	17	18	18	17	70
	$OWA: w_1 = 0.50$	15	17	17	15	64

tado una tarea t_j ha sido mayor que $t_{j,dl}$. En consecuencia, podemos ver de las Tablas I y II que, en promedio, cuando no se considera el estímulo de inercia ($p = 0$), las tareas no se terminan dentro del plazo deseado, al ser prácticamente todos los valores de $TbDL$ negativos. Además, si este estímulo es considerado en su máxima expresión ($p = 1$), el mínimo es la función de agregación que mejores resultados ha proporcionado en la situación FC, mientras que en el caso RC no hay un claro vencedor. Más específicamente, cuando nos encontramos en el caso RC con $p = 1$, el *OWA* modificado con pesos $w = \{1, 0\}$ (máximo modificado) ha sido la función de agregación que mejores resultados ha proporcionado para las métricas TTC y D, pero para las métricas U y $TbDL$ la mejor función de agregación ha sido el *OWA* modificado con pesos $w = \{0.75, 0.25\}$. Por este motivo, en el caso de tener recursos limitados, sería más adecuado usar la versión modificada de la función de agregación máximo, pero si fuera indispensable terminar todas las tareas a tiempo, sin importar los recursos consumidos, sería mejor usar el *OWA* modificado con pesos $w = \{0.75, 0.25\}$.

Como podemos observar, con los resultados presentados en las Tablas I y II es muy difícil comparar directamente el rendimiento de las diferentes configuraciones, porque en muchos casos este está fuertemente relacionado con la métrica bajo consideración. Por este motivo, en las Tablas III y IV ofrecemos la posición obtenida por cada una de las configuraciones en un ranking elaborado para cada una de las métricas propuestas. Además, calculamos la suma de dichas posiciones para cada configuración. De esta manera, uno podría afirmar que la mejor configuración será aquella cuya suma sea lo menor posible, puesto que esa será la que tenga globalmente la mejor posición entre todas las métricas. Evidentemente, otras metodologías podrían ser consideradas para poder escoger la configuración óptima para el problema planteado, dando más importancia a las métricas que ayuden más a las necesidades específicas del usuario. El motivo por el cual hemos decidido usar el método del ranking es porque este nos proporciona la configuración más versátil.

Comparando los resultados obtenidos en las tablas III y IV, lo primero que llama la atención es el gran salto de eficiencia que se obtiene cuando $IN_{cl,t_j} \neq 0$. De hecho, tanto es así que la peor configuración considerando $p = 1$ es claramente mejor que la mejor configuración considerando $p = 0$. De la misma manera, restringir las capacidades de toma de decisiones de los robots mientras están viajando —situación RC— hace que el sistema sea mucho más eficiente. De estos hechos, deducimos que no permitir que los robots cambien su destino mientras viajan, y obligarles a quedarse en la tarea hasta terminarla, se traduce en mucho menos tiempo invertido en desplazarse y en que más tareas se terminen a tiempo, permitiendo así que las puntuaciones en las distintas métricas sean mejores. Para terminar, notemos que, en base a nuestro método de evaluación por ranking, el mejor rendimiento se obtiene con el operador *OWA* modificado bajo la situación RC y con $p = 1$. Más concretamente, los mejores resultados se han conseguido con los pesos $w = \{0.75, 0.25\}$, los segundos mejores con los pesos $w = \{0.50, 0.50\}$, y los terceros mejores con los pesos $w = \{1, 0\}$ (máximo modificado).

V. CONCLUSIONES Y TRABAJO FUTURO

En este artículo, hemos presentado una estrategia para métodos de umbral de respuesta (RTM) que permite afrontar tareas con plazos de tiempo límite asociados en problemas de tipo MRTA. Con esta estrategia, se modelan los estímulos que reciben los robots mediante conjuntos difusos, lo cual permite que se decida la mejor tarea a realizar solucionando un problema de optimización difuso mediante la técnica de Bellman y Zadeh.

Para poder evaluar la mencionada metodología, se han llevado a cabo una extensa cantidad de simulaciones, considerando distintos parámetros, así como cinco posibles funciones de agregación diferentes. Los parámetros que se han ajustado afectan tanto a la detección del estímulo procedente de la *inercia*, como a la capacidad de cambiar la tarea objetivo mientras se está viajando. Los resultados obtenidos demuestran que el enfoque matemático propuesto es capaz de modelar el tipo de sistema deseado manejando tareas con plazos de tiempo límite. Concretamente, el mejor rendimiento se ha obtenido cuando los robots sienten la necesidad de quedarse en la tarea en la que se encuentran en ese momento hasta terminarla y sin tener la capacidad de cambiar su destino mientras se están desplazando. Además, la mejor asignación de tareas se ha conseguido utilizando una versión modificada del operador *OWA*.

Recordemos que, aunque todos los robots son heterogéneos en términos de su conjunto de habilidades, el proceso de toma de decisiones es homogéneo para todos los robots. Como propuesta de trabajo futuro, planeamos introducir un mayor nivel de heterogeneidad, personalizando para cada robot la función de agregación y los conjuntos difusos usados para modelar los distintos estímulos.

REFERENCES

[1] J. Guerrero, O. Valero, and G. Oliver, "Toward a possibilistic swarm multi-robot task allocation: theoretical and experimental results," *Neural Processing Letters*, vol. 46, pp. 881–897, 2017.

[2] J. Guerrero, J.-J. Miñana, and O. Valero, "On the use of fuzzy preorders in multi-robot task allocation problem," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations. IPMU 2018.* Springer, 2018, pp. 195–206.

[3] M.-d.-M. Bibiloni-Femenias, J. Guerrero, J.-J. Miñana, and O. Valero, "Indistinguishability operators via yager t-norms and their applications to swarm multi-agent task allocation," *Mathematics*, vol. 2, no. 9, p. 190, 2021.

[4] O. Valero, J. Antich, A. Tauler-Rosselló, J. Guerrero, J.-J. Miñana, and A. Ortiz, "Multi-robot task allocation methods: A fuzzy optimization approach," *Information Sciences*, vol. 648, p. 119508, 2023.

[5] R. Bellman and L. Zadeh, "Decision making in a fuzzy environment," *Management Science*, vol. 17, no. 4, pp. 141–164, 1970.

[6] E. Trillas and L. Eciolaza, *Fuzzy Logic. An Introductory Course for Engineering Students*. Springer, 2015.

[7] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation functions*. Academic Press, 2009.

[8] T. Pedraza, J. Rodríguez-López, and O. Valero, "Aggregation of fuzzy quasi-metrics," *Information Sciences*, vol. 581, pp. 362–389, 2021.

Funciones de agregación inspiradas en la integral Choquet

Humberto Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
31006, Pamplona
Email: bustince@unavarra.es

Julio Lafuente
Departamento de Estaística, Informática y Matemáticas
Universidad Pública de Navarra
31006, Pamplona
Email: lafuente@unavarra.es

Xabier Gonzalez-Garcia
Departamento de Estaística, Informática y Matemáticas
Universidad Pública de Navarra
31006, Pamplona
Email: xabier.gonzalez@unavarra.es

Graçaliz P. Dimuro
Computer Science Center,
Federal University of Rio Grande,
Rio Grande, Brazil
Email: gracalizdimuro@furg.br

Radko Mesiar
Faculty of Civil Engineering, Slovak University of Technology,
Radlinského 11, 810 05
Bratislava, Slovakia
Email: mesiar@math.sk

***Resumen*—En este trabajo presentamos una nueva clase de funciones de agregación. Para la definición de estas nuevas funciones nos inspiramos en el método de construcción de las integrales Choquet, reemplazando las medidas por funciones adecuadas. Tras discutir la definición de las nuevas funciones, estudiamos algunas de su propiedades básicas y consideramos su relación con otras funciones de agregación utilizadas en la literatura, como los estadísticos de orden o las funciones de overlap y grouping.**

I. Introducción

La fusión de información es un paso crucial en la resolución de la mayor parte de los problemas científicos. Por este motivo, se han dedicado grandes esfuerzos a desarrollar técnicas apropiadas para abordar este problema. Entre los métodos más utilizados se encuentran los que hacen uso de las funciones de agregación [2], [1], [10]. Recordemos que las funciones de agregación son funciones monótonas crecientes que están habitualmente definidas en el intervalo unidad $[0, 1]$ y que proporcionan un valor único, también en el intervalo unidad, como representante de un conjunto de datos numéricos de entrada.

Las funciones de agregación se han usado con éxito en problemas muy diversos, incluyendo el aprendizaje automático [13], la toma de decisiones [9] o la neurociencia computacional [8], entre muchos otros. Una familia particular de funciones de agregación, las llamadas integrales difusas, ha experimentado un amplio desarrollo en los últimos años, ya que, por medio de medidas difusas, pueden tener en cuenta relaciones entre diferentes datos para elegir un valor que los represente de una forma adecuada.

Posiblemente el ejemplo más representativo de las inregrales difusas lo constituyen las integrales de Choquet [5] y sus generalizaciones [16], [17], [7], [15]. Muchas de estas generalizaciones se han obtenido considerando operaciones más generales que la suma, el producto, el máximo o el mínimo utilizados en la definición original de integrales difusas [6], o reemplazando la resta en esas mismas definiciones por otras medidas de información, como, por ejemplo, disimilitudes [3], lo que, a su vez, ha permitido la extensión de estos operadores de agregación a marcos más generales que el numérico.

El principal objetivo de este trabajo es proponer una nueva forma de construir funciones de agregación que se inspira en la definición de integral Choquet y reemplaza las medidas por funciones más generales. Estas funciones, al igual que las medidas, son dependientes de los datos a fusionar en cada caso. En concreto, consideramos una combinación convexa de los datos de entrada, haciendo que los pesos de dicha combinación sean también dependientes de las entradas.

La estructura del trabajo es la siguiente. Comenzamos presentado algunos conceptos y resultados preliminares necesarios para el desarrollo del trabajo. En la Sección III, introducimos la definición de la nueva familia de funciones inspiradas en la integral Choquet. En la Sección IV estudiamos algunas primeras propiedades de estas funciones y, en la Sección V, consideramos su relación con algunas funciones de agregación bien conocidas. Terminamos con unas breves conclusiones y referencias.

II. Preliminares

Dado $n \geq 2$, denotamos $N = \{1, \ldots, n\}$. 2^N denota el conjunto de todos los subconjuntos de N.

Si $n \geq 3$, dados $i, j \in \{1, \ldots, n\}$ con $i \neq j$ y dado $\mathbf{x} = (x_1, \ldots, x_n) \in [0,1]^n$, denotamos por $\hat{\mathbf{x}}_{i,j}$ el resultado de la proyección $P_{(i,j)} : [0,1]^n \to [0,1]^{n-2}$ en la que se omiten las componentes i y j .

Dado $(x_1, \ldots, x_n) \in [0,1]^n$, denotamos por (i) la posición i-ésima de una permutación de N tal que:

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)} \ .$$

Para $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, escribimos:

$$\|\mathbf{x}\| = \sum_{i=1}^{n} |x_i| \quad ,$$

esto es, la norma L^1 usual en $\mathbb{R}^n$.

Recordamos a continuación algunos de los conceptos básicos para este trabajo.

Definición 2.1: Una función de agregación n-dimensional es una función $A : [0,1]^n \to [0,1]$ monótona creciente y tal que $A(0, \ldots, 0) = 0$ y $A(1, \ldots, 1) = 1$.

Recordemos que una función de agregación es simétrica si se verifica que $A(x_1, \ldots, x_n) = A(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ para cualquier permutación σ de N.

Definición 2.2: Una función de agregación $G_O : [0,1]^n \to [0,1]$ es una función de overlap si es simétrica, continua y tal que:

1. $G_O(x_1, \ldots, x_n) = 0$ si y solo si $x_1 \cdots x_n = 0$, y
2. $G_O(x_1, \ldots, x_n) = 1$ si y solo si $x_1 \cdots x_n = 1$.

Definición 2.3: Una función de agregación $G_G : [0,1]^n \to [0,1]$ es una función de grouping si es simétrica, continua y:

1. $G_G(x_1, \ldots, x_n) = 0$ si y solo si $x_1 = \cdots = x_n = 0$, y
2. $G_G(x_1, \ldots, x_n) = 1$ si y solo si existe $i \in \{1, \ldots, n\}$ tal que $x_i = 1$.

A continuación, recordamos la definición de los operadores OWA [18].

Definición 2.4: Sea $n \geq 2$. Un vector de pesos es una n-tupla $(w_1, \ldots, w_n) \in [0,1]^n$. A lo largo de este trabajo, impondremos además, que $w_1 + \cdots + w_n = 1$.

Definición 2.5: Sea $n \geq 2$ y fijemos un vector de pesos $\mathbf{w} = (w_1, \ldots, w_n) \in [0,1]^n$. El operador OWA definido por $\mathbf{w}$ es la función $F_{\mathbf{w}} : [0,1]^n \to [0,1]$ dada por:

$$F_{\mathbf{w}}(x_1, \ldots, x_n) = w_1 x_{(n)} + w_2 x_{(n-1)} + \cdots + w_n x_{(1)}$$

para todo $(x_1, \ldots, x_n) \in [0,1]^n$.

Finalmente, recordamos la noción de integral Choquet [5].

Definición 2.6: Sea $n \geq 2$ and $N = \{1, \ldots, n\}$. Una medida difusa en N es una función $m : 2^N \to [0,1]$ tal que:

1. $m(\emptyset) = 0$ y $m(N) = 1$;
2. si $A \subseteq B \subseteq N$, entonces $m(A) \leq m(B)$

Definición 2.7: Sea $m : 2^N \to [0,1]$ una medida difusa. La integral de Choquet definida por m es la función $C_m : [0,1]^n \to [0,1]$ dada, para cada $\mathbf{x} = (x_1, \ldots, x_n) \in [0,1]^n$, por:

$$C_m(\mathbf{x}) = \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \cdot m(A_{(i)})$$

donde $A_i = \{(i), (i+1), \ldots, (n)\}$ y donde se define $x_{(0)} = 0$.

III. Funciones de agregación inspiradas en la integral Choquet

La definición fundamental de este trabajo es la siguiente.

Definición 3.1: Sea $n \geq 3$ y sea $F : [0,1]^{n-2} \to [0,1]$ una función simétrica y creciente. Definimos la función (inspirada en Choquet) $CI_F : [0,1]^n \to [0,1]$ como:

$$CI_F(x_1, \ldots, x_n) = x_{(1)} + \sum_{i=1}^{n-1} (x_{(i+1)} - x_{(i)}) F(\hat{\mathbf{x}}_{(i+1),(i)}) \tag{1}$$

donde, como se ha indicado anteriormente, $(\cdot)$ una permutación de $\{1, \ldots, n\}$ tal que:

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$$

Nótese que la función CI_F está bien definida incluso si existe más de una posible permutación creciente $(\cdot)$ para $\mathbf{x}$, debido a la simetría de F.

Al igual que sucede en el caso de la integral Choquet usual, la expresión de CI_F dada por Eq. (1) puede también reescribirse de la forma siguiente:

$$\begin{aligned} CI_F(x_1, \ldots, x_n) &= x_{(1)}(1 - F(\mathbf{x}_{(1),(2)})) \\ &+ \sum_{i=2}^{n-1} x_{(i)}(F(\mathbf{x}_{(i-1),(i)}) - F(\mathbf{x}_{(i),(i+1)})) \\ &+ x_{(n)}(F(\mathbf{x}_{(n-1),(n)})) \end{aligned} \tag{2}$$

Es decir, CI_F puede verse como una combinación convexa de las entradas con coeficientes variables. Dichos coeficientes son dependientes de los propios datos de entrada.

Ejemplo 1: Sea $k \in [0,1]$. Consideremos la función constante $F \equiv k$ (que es trivialmente simétrica y creciente). Entonces, para cualquier $n \geq 3$ y para cualquier $x_1, \ldots, x_n \in [0,1]$, es fácil calcular que:

$$CI_F(\mathbf{x}) = (1-k)\,\text{mín}(x_1, \ldots, x_n) + k\,\text{máx}(x_1, \ldots, x_n) \ .$$

Esto es, recuperamos una combinación convexa del mínimo y el máximo con coeficiente k. En particular:

- Si $k = 0$, recuperamos el mínimo.
- Si $k = 1$, recuperamos el máximo.

Ejemplo 2: Sea $n = 4$. Para cualquier función $F : [0,1]^2 \to [0,1]$ simétrica y creciente:

$$\begin{aligned} CI_F(x_1, x_2, x_3, x_4) &= x_{(1)} + (x_{(2)} - x_{(1)}) F(x_{(3)}, x_{(4)}) \\ &+ (x_{(3)} - x_{(2)}) F(x_{(1)}, x_{(4)}) + (x_{(4)} - x_{(3)}) F(x_{(1)}, x_{(2)}) \end{aligned}$$

En particular, si $F(x, y) = \frac{x+y}{2}$, entonces tenemos que:

$$CI_F(x_1, x_2, x_3, x_4) = x_{(1)}(1 - \frac{x_{(2)} + x_{(3)}}{2}) + x_{(4)}(\frac{x_{(2)} + x_{(3)}}{2})$$

Es decir, de nuevo CI_F es una combinación convexa (o, equivalentemente, una media aritmética ponderada) de los valores extremos, pero en este caso los pesos vienen dados por los valores intermedios (centrales) de la entrada.

IV. Algunas propiedades de las funciones de agregación inspiradas en la integral Choquet

En esta sección, estudiamos algunas primeras propiedades de las funciones inspiradas en la integral Choquet que acabamos de definir. Empezamos probando que, de hecho, las funciones resultantes son funciones de agregación y que además son idempotentes.

Proposición 4.1: Sea $F : [0,1]^{n-2} \to [0,1]$ simétrica y creciente. Entonces, CI_F es una función de agregación idempotente.

Demostración.

La idempotencia es directa, y de e3lla se deducen de forma directa las condiciones de contorno. Probemos, pues, la monotonía. Sean $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in [0,1]^n$ tales que $x_i \leq y_i$ para todo $i \in \{1,\ldots,n\}$. Podemos asumir, sin pérdida de generalidad, que tenemos dos n-tuplas $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in [0,1]^n$ que:

$$x_1 \leq x_2 \leq \ldots x_n \text{ y } y_1 \leq y_2 \leq \ldots y_n \ .$$

y existe $i_0 \in \{1,\ldots,n\}$ con $y_{i_0} = x_{i_0} + \delta$ para algún $\delta > 0$ y $x_i = y_i$ para cualquier $i \neq i_0$.

Consideremos los tres casos siguientes:

1.- Si $i_0 = 1$, entonces:

$$\begin{aligned} & CI_F(y_1,\ldots,y_n) - CI_F(x_1,\ldots,x_n) \\ = \ & (x_1+\delta) + (x_2 - (x_1+\delta)F(\hat{\mathbf{y}}_{(1),(2)}) \\ + \ & \sum_{i=2}^{n-1}(x_{i+1}-x_i)F(\hat{\mathbf{y}}_{(i+1),(i)}) \\ - \ & x_1 - (x_2-x_1)F(\hat{\mathbf{x}}_{(1),(2)}) \\ - \ & \sum_{i=2}^{n-1}(x_{i+1}-x_i)F(\hat{\mathbf{x}}_{(i+1),(i)})) \end{aligned}$$

Como $F(\hat{\mathbf{y}}_{(1),(2)}) = F(\hat{\mathbf{x}}_{(1),(2)})$ (porque solo $x_1 \neq y_1$), tenemos que:

$$\begin{aligned} & (x_1+\delta) + (x_2 - (x_1+\delta)F(\hat{\mathbf{y}}_{(1),(2)}) \\ - \ & x_1 - (x_2-x_1)F(\hat{\mathbf{x}}_{(1),(2)}) \\ = \ & \delta(1 - F(\hat{\mathbf{x}}_{(1),(2)})) > 0 \end{aligned}$$

Como:

$$\sum_{i=2}^{n-1}(x_{i+1}-x_i)F(\hat{\mathbf{y}}_{(i+1),(i)}) \geq \sum_{i=2}^{n-1}(x_{i+1}-x_i)F(\hat{\mathbf{x}}_{(i+1),(i)})$$

por la monotonía de F, vemos que $CI_F(y_1,\ldots,y_n) - CI_F(x_1,\ldots,x_n) \geq 0$.

2.- Si $i_0 = n$, podemos aplicar un argumento similar, teniendo en cuenta que:

$$(x_n + \delta - x_{n-1})F(\hat{\mathbf{y}}_{(n),(n-1)}) \geq (x_n - x_{n-1})F(\hat{\mathbf{x}}_{(n),(n-1)}))$$

3.- Finalmente, si $1 < i_0 < n$, tenemos que:

$$\begin{aligned} & (x_{i_0} + \delta - x_{i_0-1})F(\hat{\mathbf{y}}_{(i_0),(i_0-1)}) \\ + \ & (x_{i_0+1} - x_{i_0} - \delta)F(\hat{\mathbf{y}}_{(i_0+1),(i_0)}) \\ - \ & (x_{i_0} - x_{i_0-1})F(\hat{\mathbf{x}}_{(i_0),(i_0-1)}) \\ - \ & (x_{i_0+1} - x_{i_0})F(\hat{\mathbf{x}}_{(i_0+1),(i_0)}) \end{aligned}$$

que es mayor o igual que cero porque

$$\delta F(\hat{\mathbf{y}}_{(i_0),(i_0-1)}) \geq \delta F(\hat{\mathbf{y}}_{(i_0+1),(i_0)})$$

por la monotonía de F. □

El siguiente resultado es directo

Proposición 4.2: Sea $F : [0,1]^{n-2} \to [0,1]$ una función simétrica creciente y continua y sea CI_F definida por la Ec. (1). Entonces, CI_F es continua.

Ejemplo 3: En general, la función inspirada en Choquet CI_F no tiene por qué ser continua para una F arbitraria, no continua. Por ejemplo, sea $n = 3$, y tomemos la función $F : [0,1] \to [0,1]$ dada por:

$$F(x) = \begin{cases} 0 & \text{si } x < \frac{1}{2} \\ 1 & \text{en otro caso.} \end{cases}$$

Entonces $CI_F(0,0{,}4,0{,}5) = 0{,}4$, mientras que, si tomamos $\epsilon \in]0,0{,}1]$, $CI_F(0,0{,}4,0{,}5-\epsilon) = 0$, luego la función resultante no es continua en este caso.

A continuación consideramos el comportamiento de las funciones CI_F cuando F viene dada por una combinación convexa de dos funciones F_1 and F_2. En particular, tenemos el siguiente resultado:

Proposición 4.3: Sean $F_1, F_2 : [0,1]^{n-2} \to [0,1]$ dos funciones simétricas crecientes, y $0 < k < 1$. Entonces, para todo $(x_1,\ldots,x_n) \in [0,1]^n$, se verifica que:

$$\begin{aligned} & CI_{kF_1+(1-k)F_2}(x_1,\ldots,x_n) \\ = \ & kCI_{F_1}(x_1,\ldots,x_n) + (1-k)CI_{F_2}(x_1,\ldots,x_n) \end{aligned}$$

Demostración. Basta observar que::

$$\begin{aligned} & CI_{kF_1+(1-k)F_2}(x_1,\ldots,x_n) = x_{(1)} \\ + \ & \sum_{i=1}^{n-1}(x_{(i+1)} - x_{(i)})(kF_1(\hat{\mathbf{x}}_{(i+1),(i)}) \\ + \ & (1-k)F_2(\hat{\mathbf{x}}_{(i+1),(i)})) \\ = \ & kx_{(1)} \\ + \ & \sum_{i=1}^{n-1}(x_{(i+1)} - x_{(i)})(kF_1(\hat{\mathbf{x}}_{(i+1),(i)}) \\ + \ & (1-k)x_{(1)} + \sum_{i=1}^{n-1}(x_{(i+1)} - x_{(i)})((1-k)kF_2(\hat{\mathbf{x}}_{(i+1),(i)}) \\ = \ & kCI_{F_1}(x_1,\ldots,x_n) + +(1-k)CI_{F_2}(x_1,\ldots,x_n) \quad \square \end{aligned}$$

Corolario 4.4: Sea $(w_1,\ldots,w_r) \in [0,1]^r$ tal que $\sum_{i=1}^r w_i \leq 1$. Entonces, para cualquier familia de funciones

simétricas crecientes $F_1, \ldots, F_r : [0,1]^{n-2} \to [0,1]$, y para cualquier $(x_1, \ldots, x_n) \in [0,1]^n$, se verifica que:

$$CI_{\sum_{j=1}^r w_j F_j}(x_1, \ldots, x_n) = (1 - \sum_{j=1}^r w_j)x_{(1)} + \sum_{j=1}^r w_j CI_{F_j} \ .$$

Nota 1: Nótese que, si $F_1 + F_2 \leq 1$, entonces $CI_{F_1+F_2}(\mathbf{x}) = CI_{F_1}(\mathbf{x}) + CI_{F_2}(\mathbf{x}) - x_{(1)}$.

Pasamos ahora a considerar la propiedad de homogeneidad. Recordemos que una función $M : [0,1]^n \to [0,1]$ es homogénea positiva si , para cada $\mathbf{x} \in [0,1]^n$ y $\lambda > 0$ tal que $\lambda\mathbf{x} \in [0,1]^n$, se verifica que:

$$M(\lambda\mathbf{x}) = \lambda M(\mathbf{x}) \ .$$

Para una función simétrica creciente F y para $\mathbf{x} \in [0,1]^n$ y $\lambda > 0$ tales que $\lambda\mathbf{x} \in [0,1]^n$, vemos que:

$$CI_F(\lambda\mathbf{x}) = \lambda x_{(1)} + \sum_{i=1}^{n-1} (\lambda x_{(i+1)} - \lambda x_{(i)}) F(\lambda\hat{\mathbf{x}}_{(i+1),(i)}) \quad (3)$$

Por tanto, tenemos el siguiente resultado:

Proposición 4.5: Sea $F : [0,1]^{n-2} \to [0,1]$ una función simétrica creciente. Entonces, CI_F es homogénea positiva si y solo si F es constante.

Demostración.

Si F es constante, el hecho de que CI_F es positiva homogénea se sigue directamente de la Ec.(3).

Para ver el recíproco, supongamos que $CI_F(\lambda\mathbf{x}) = \lambda CI_F(\mathbf{x})$. Entonces:

$$\begin{aligned} & \sum_{i=1}^{n-1} (\lambda x_{(i+1)} - \lambda x_{(i)}) F(\lambda\hat{\mathbf{x}}_{(i+1),(i)}) \\ = \ & \lambda \sum_{i=1}^{n-1} (x_{(i+1)} - x_{(i)}) F(\hat{\mathbf{x}}_{(i+1),(i)}) \end{aligned}$$

De la monotonía de F, si F no es constante, existe $\lambda > 0$ y $(y_1, \ldots, y_{n-2}) \in [0,1]^{n-2}$ tales que :

$$F(\lambda y_1, \ldots, \lambda y_{n-2}) \neq F(y_1, \ldots, y_{n-2})$$

Supongamos que $\lambda < 1$. Entonces, de nuevo a partir de la monotonía de F, tenemos que:

$$F(\lambda y_1, \ldots, \lambda y_{n-2}) < F(y_1, \ldots, y_{n-2}) \ .$$

Por la simetría de F podemos asumir, sin pérdida de generalidad, que $y_1 \leq y_2 \leq \cdots \leq y_{n-2}$.

Consideremos la n-tupla $\mathbf{y} = (0, y_1, y_1, \ldots, y_{n-2}) \in [0,1]^n$. Entonces:

$$\begin{aligned} CI_F(\lambda\mathbf{y}) &= \lambda \sum_{i=1}^{n-2} (y_{(i+1)} - y_{(i)}) F(\lambda\hat{\mathbf{y}}_{(i+1),(i)}) \\ < \ & y_1 F(\ y_1, \ldots, y_{n_2}) + \sum_{i=1}^{n-2} (y_{(i+1)} - y_{(i)}) F(\hat{\mathbf{y}}_{(i+1),(i)}) \end{aligned}$$

de nuevo por la monotonía de F.

Como podemos repetir un argumento similar para $\lambda > 1$, simplemente dando la vuelta a la desigualdad, tenemos el resultado. □

V. Algunos casos específicos

En primer lugar, vamos a considerar el caso de los estadísticos de orden. Es decir, dado $i \in \{1, \ldots, n-2\}$, consideramos:

$$F(y_1, \ldots, y_{n-2}) = Q_i(y_1, \ldots, y_{n-2}) = y_{(i)}$$

Un cálculo directo muestra que:

$$\begin{aligned} CI_F(\mathbf{x}) &= x_{(1)} + \sum_{j=1}^{i} (x_{(j+1)} - x_{(j)}) x_{(i+2)} \\ + & \sum_{j=i+1}^{n-1} (x_{(j+1)} - x_{(j)}) x_{(i)} \end{aligned}$$

donde alguno de los sumatorios podría ser vacío. Esta expresión es equivalente a:

$$\begin{aligned} & CI_F(\mathbf{x}) = x_{(1)} + (x_{(i+1)} - x_{(1)}).x_{(i+2)} \\ + \ & (x_{(n)} - x_{(i+1)}).x_{(i)} = x_{(1)}.(1 - x_{(i+2)}) \\ + \ & x_{(i+1)}.(x_{(i+2)} - x_{(i)}) + x_{(n)}.x_{(i)} \end{aligned}$$

Para $i = 1$ obtenemos:

$$CI_F(\mathbf{x}) = x_{(1)} + (x_{(2)} - x_{(1)}) x_{(3)} + \sum_{j=2}^{n-1} (x_{(j+1)} - x_{(j)}) x_{(1)})$$

En particular,

- si $x_{(1)} = 0$, obtenemos que

$$CI_F(\mathbf{x}) = x_{(2)} x_{(3)} \ .$$

- Más generalmente, si $x_{(i)} = 0$, tenemos que $x_{(j)} = 0$ para $1 \leq j \leq i$, y si tomamos $F = Q_i$

$$CI_F(\mathbf{x}) = x_{(i+1)} x_{(i+2)} \ .$$

Por otra parte, el operador OWA es simplemente una combinación convexa de operadores estadísticos de orden. Por tanto, tenemos el siguiente resultado, cuya prueba es directa

Proposición 5.1: Sea $\mathbf{w} = (w_1, \ldots, w_{n-2}) \in [0,1]^{n-2}$ un vector de pesos y sea $F_{\mathbf{w}} : [0,1]^{n-2} \to [0,1]$ el operador OWA definido por $\mathbf{w}$. Entonces:

$$CI_{F_{\mathbf{w}}}(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i CI_{Q_{n-i+1}}(x_1, \ldots, x_n)$$

En cuanto a la relación con las funciones de overlap y grouping, tenemos, en primer lugar, el siguiente resultado.

Proposición 5.2: Para cualquier fumnción simétrica creciente $F : [0,1]^{n-2} \to [0,1]$, la función CI_F no tiene divisores de cero

Demostración. Basta observar que, para cualquier n-tupla $\mathbf{x}$ con todas sus componentes estrictamente positivas, se verifica que $x_{(1)} > 0$, y, por tanto, $CI_F(\mathbf{x}) \geq x_{(1)} > 0$. □

Claramente, en general CI_F no es una función de overlap (porque ni siquiera tiene por qué ser continua). Más aún, tenemos el siguiente resultado:

Proposición 5.3: Sea $n > 2$ y $F : [0,1]^{n-2} \to [0,1]$ una función de overlap. Entonces, $CI_F(\mathbf{x}) = 0$ si y solo si $x_{(1)} = x_{(2)} = 0$.

Demostración.

Si $x_{(1)} = x_{(2)} = 0$, es directo que $CI_F(x_1, \ldots, x_n) = 0$.

Para ver el converso, si F es una función de overlap, entonces $F(y_1, \ldots, y_{n-2}) = 0$ siempre que $y_1 \cdot \cdots \cdot y_{n-2} = 0$. Como todos los términos en la Ec. (1) son no negativos, $CI_F(\mathbf{x}) = 0$ implica que:

1. $x_{(1)} = 0$, y
2. $(x_{(2)} - x_{(1)})F(\hat{\mathbf{x}}_{(2),(1)}) = 0$

Si $x_{(2)} = x_{(1)}$, tenemos el resultado. En otro caso, para que 2. se verifique, ha de cumplirse que $F(\hat{\mathbf{x}}_{(2),(1)}) = 0$. Pero esto no es posible porque todas las componentes de $\hat{\mathbf{x}}_{(2),(1)}$ son mayores y iguales que $x_{(2)}$ (y, por tanto, positivas) y F es una función de overlap. □

De hecho, podemos caracterizar cuándo CI_F es una función de overlap.

Teorema 5.4: La función $CI_F : [0,1]^n \to [0,1]$ es una función de overlap si y solo si la función F es idénticamente cero.

Demostración.

Tomemos $(0, x_2, x_3, \ldots, x_n) \in [0,1]^n$ con $0 < x_2 < x_3 < \ldots < x_n$. Si CI_F es una función de overlap, entonces $CI_F(0, x_2, \ldots, x_n) = 0$. Pero:

$$\begin{aligned} & CI_F(0, x_2, \ldots, x_n) = x_{(2)} F(\hat{\mathbf{x}}_{(2),(1)}) \\ + & \sum_{i=2}^{n-1} (x_{(i+1)} - x_{(i)}) F(\hat{\mathbf{x}}_{(i+1),(i)}) \end{aligned}$$

que, por la definición de F y la elección de $(0, x_2, x_3, \ldots, x_n)$,es siempre cero si y solo si F es idénticamente cero.

El recíproco es directo. □

Corolario 5.5: CI_F es una función de overlap si y solo si $CI_F(x_1, \ldots, x_n) = \min(x_1, \ldots, x_n)$.

Con respecto a las funciones de grouping, tenemos el siguiente resultado.

Teorema 5.6: Sea $F : [0,1]^{n-2} \to [0,1]$ una función simétrica creciente y continua. La función $CI_F : [0,1]^n \to [0,1]$ es una función de grouping si y solo si la función F es tal que:

1. $F(x_1, \ldots, x_{n-2}) = 1$ siempre que $x_{(n-2)} = 1$, y
2. si $F(x_1, \ldots, x_{n-2}) = 0$, entonces $x_1 = \cdots = x_{n-2} = 0$.

Demostración.

La continuidad se sigue de las hipótesis. Sin pérdida de generalidad, supongamos que existe $(x_1, x_2, x_3, \ldots, x_{n-2}) \in [0,1]^{n-2}$ tal que $x_1 \leq x_2 \leq \cdots \leq x_{n-2}$ con $x_{n-2} = 1$ y tal que $F(x_1, \ldots, x_{n-2}) < 1$. Entonces, de la monotonía de F, se sigue que $F(0, \ldots, 0, 1) < 1$. Pero, en este caso:

$$CI_F(0, 0, \ldots, 0, 1) = x_{(n)} F(\hat{\mathbf{x}}_{(n),(n-1)}) < 1$$

lo que contradice la definición de función de grouping. Por tanto, debe tenerse que $F(x_1, \ldots, x_{n-2}) = 1$ siempre que $x_{(n-2)} = 1$.

Además, si existe $(x_1, \ldots, x_{n-2}) \in [0,1]^n$ con $x_{(n-2)} > 0$ y tal que $F(x_1, \ldots, x_{n-2}) = 0$, por la monotonía y la simetría de F tenemos que $F(0, \ldots, 0, x_{(n-2)}) = 0$. De donde:

$$\begin{aligned} & CI_F(0, 0, \ldots, 0, x_{(n-2)}, x_{(n-2)}, x_{(n-2)}) \\ = \; & x_{(n-2)} F(0, \ldots, 0, x_{(n-2)}) = 0 \end{aligned}$$

que, de nuevo, es una contradicción.

El recíproco es directo. □

VI. Conclusiones

En este artículo hemos propuesto una familia de funciones de agregación inspiradas en la integral Choquet. Estas funciones están definidas en términos de una combinación convexa donde los pesos dependen de las entradas consideradas.

Este enfoque puede verse como un primer paso en el desarrollo de un marco general que englobe integrales difusas, por una parte, y otras funciones de agregación, por otra. Cabe esperar, por tanto, que sean aplicables en un amplio abanico de problemas.

Agradecimientos

H. Bustince, X. Gonzalez-Garcia and G.P. Dimuro han sido financiados por el proyecto de investigación PID2022-136627NB-I00 (MCIN/AEI/10.13039/501100011033/FEDER, UE)

Referencias

[1] G. Beliakov, H. Bustince, T. Calvo, A practical guide to averaging functions, Springer, Berlin, 2016.

[2] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Berlin, 2007.

[3] H. Bustince, R. Mesiar, J. Fernandez, M. Galar, D. Paternain, A. Altalhi, G. P. Dimuro, B. Bedregal, Z. Takáč, d-choquet integrals: Choquet integrals based on dissimilarities, Fuzzy Sets and Systems 414 (2021), 1-27

[4] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Directional monotonicity of fusion functions, European Journal of Operational Research 244 (1) (2015) 300–308.

[5] G. Choquet, Theory of capacities, Annales de l'Institut Fourier 5 (1953–1954) 131–295.

[6] G. P. Dimuro, G. Lucca, B. Bedregal, R. Mesiar, J. A. Sanz, C.-T. Lin, H. Bustince, Generalized $c_{F_1 F_2}$-integrals: From Choquet-like aggregation to ordered directionally monotone functions, Fuzzy Sets and Systems 378 (2020) 44–67.

[7] G. P. Dimuro, J. Fernandez, B. Bedregal, R. Mesiar, J. Sanz, G. Lucca, H. Bustince, The state-of-art of the generalizations of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions, Information Fusion, 57 (2020) 27–43.

[8] J. Fumanal-Idocin, Y.-K. Wang , C.-T. Lin, J. Fernández, J. A. Sanz, Humberto Bustince, Motor-Imagery-Based Brain–Computer Interface Using Signal Derivation and Aggregation Functions, IEEE Transactions on Cybernetics 52 (8) (2021) 7944–7955.

[9] M. Grabisch, C. Labreuche, A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, Annals of Operations Research 175 (1) (2010) 247–286.

[10] M. Grabisch, J. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, Cambridge, 2009.

[11] S. Greco, B.Matarazzo, S. Giove, The Choquet integral with respect to a level dependent capacity. Fuzzy Sets and Systems 175 (1) (2011) 1–35.

[12] E. Lehrer, A new integral for capacities, Economic Theory 39 (1) (2009) 157–176.

[13] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, M. J. Asiain, M. Elkano, H. Bustince, CC-integrals: Choquet-like copula-based aggregation functions and its application in fuzzy rule-based classification systems, Knowledge-Based Systems 119 (2017) 32 – 43.

[14] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, H. Bustince, R. Mesiar, CF-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems, Information Sciences 435 (2018) 94 – 110.
[15] G. Lucca, J. Sanz, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince Sola, Pre-aggregation functions: construction and an application, IEEE Transactions on Fuzzy Systems 24 (2) (2016) 260–272.
[16] R. Mesiar, Choquet-like integrals, Journal of Mathematical Analysis and Applications 194 (2) (1995) 477 – 488.
[17] T. Murofushi, M. Sugeno, Fuzzy t-conorm integral with respect to fuzzy measures: Generalization of Sugeno integral and Choquet integral, Fuzzy Sets and Systems 42 (1) (1991) 57 – 71.
[18] R.R.Yager, On ordered weighted averaging aggregation operators in multi-criteria decision making, IEEE Transactions on Systems, Man, and Cybernetics 18, (1988) 183–190.
[19] R. R. Yager, D. Filev, Induced ordered weighted averaging operators, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 29 (2) (1999), 141 – 150.

Towards the Interactive Natural Language Explanation of Processes

Yago Fontenla-Seco, Manuel Lama, Alberto Bugarín-Diz
Centro Singular de Investigación en Tecnologías Intelixentes (CiTIUS), Universidade de Santiago de Compostela
Santiago de Compostela, Spain
{yago.fontenla.seco, manuel.lama, alberto.bugarin.diz}@usc.es

***Abstract*—Despite the efforts of the Process Mining community, the understanding of process mining results and reports by non-technical users still remains an open challenge. For this reason, an increasing need for endowing, integrating and enriching process mining tools and pipelines with mechanisms able to convey the most salient aspects of a process in an understandable manner, arises. Within this context, natural modalities of communication, such as Natural Language, have proven as effective ways to convey information to users in an understandable manner in multiple domains. In this paper, we present a human-in-the-loop approach for the interactive generation of Natural Language Explanations of Processes. The system allows non-technical users to interact, query and retrieve relevant and understandable natural language explanations of a process. Based in process mining algorithms and techniques, the proposed architecture employs fuzzy set theory for capturing the relevant semantics of a process within an ontology-driven system, capable of handling and reasoning with imprecise knowledge. We show the validity of our approach through the generation of a series of natural language explanations for a healthcare process, which have already been proven correct, relevant and understandable in the literature.**

***Index Terms*—Process mining, Natural Language Generation, Fuzzy Quantification, Process Querying, Fuzzy Querying**

I. Introduction

Process mining is a process-centric data-oriented discipline aimed at exploiting event data recorded during the execution of a process, allowing to discover, monitor and ultimately understand and enhance processes [2]. Process discovery is one of the core tasks of process mining, whose aim is to discover understandable and human-interpretable process models [13]. However, in real life scenarios, processes usually turn to be extremely complex. This difficulty, calls for technical knowledge (process analysis, data science, statistics and related) to be able to understand and utilize process mining tools, hindering the accessibility of process mining for non-technical users, whom are not specialized in data analytics and which are in turn the usual stakeholders in organizations [13]. This raises the need for novel ways of conveying the results of process mining to users in a clear, direct and understandable way. In this sense, being the intrinsic way of communication for humans, natural language does not rely on users capabilities to extract relevant information from data or graphs.

Natural Language Generation (NLG) systems, in particular Data-to-Text (D2T) are primarily aimed at generating natural language descriptions from non linguistic input data [16]. These systems have been used to facilitate the understanding of information in multiple domains and application areas. In the process mining domain, natural language descriptions have been proved to help on understanding process mining results [7].

In this paper we propose a new reference ontology-based architecture to support the interactive generation of natural language explanations of processes aiming to improve the understandability and transparency of process mining for non-technical users. Based on process mining and through the use of fuzzy sets in fuzzy quantified sentences, the system is capable of modeling process semantics in a linguistic manner and reason with them, thus effectively generating natural language explanations of processes. The main contributions of this paper are:

- A novel architecture for the interactive generation of natural language explanations of processes, which is based on process mining and ontology-driven. This architecture facilitates a structured approach to extracting and conveying process-related information in a comprehensible natural language format.
- A specialized ontology that integrates fuzzy set theory and process mining. This ontology is designed to model and semantically reason about processes in a linguistically coherent manner. Additionally, the ontology is designed to be user-extendable, which supports the dynamic generation of tailored natural language explanations based on user interactions and extensions.
- Based on the proposed architecture, we develop an algorithm that interprets fuzzy queries and maps them to the developed ontology, generating intermediate logical representations. This allows for the interactive generation of accurate and contextually relevant natural language explanations of processes.
- A proof-of-concept of the generation of natural language explanations of processes which have been already identified as relevant in the healthcare domain.

This research is part of the R+D+i project TED2021-130295B-C33, funded by MCIN/AEI/10.13039/501100011033/ and the "European Union NextGenerationEU/PRTR". It also contributes to the PID2020-112623GB-I00 project funded by MCIN/AEI/10.13039/501100011033/ and by "ESF Investing in your future". The support from the Consellería de Educación, Universidades e Formación Profesional (Xunta de Galicia) and ERDF/FEDER program through grants ED431G2019/04 and ED431C2022/19 is also acknowledged.

II. Related work

A. Responsible Process mining

Multiple research initiatives exist on methods that can be grouped under the name of responsible data science. These initiatives seek to create awareness about potential negative consequences of data and knowledge misuse, and propose a series of principles to ensure its responsible use [13]. Responsible Process Mining (RPM) aims on developing a framework to make process mining more trustworthy, fair and transparent. Based on the FACT (Fairness, Accuracy, Confidentiality, Transparency) principles, RPM poses a series of challenges that process mining has to face and (potentially) solve in the present or immediate future; granting the fairness, trustworthiness and transparency of process mining [13]. One of the challenges in the RPM framework is focused on the explainability and understandability of process models and process mining results i.e., *"How to ensure accurate interpretation of complex results by domain experts?"*.

B. Natural Language in Process Mining

Previous to the proposal of the RPM framework, the need for better ways of conveying process mining results was already identified [7], [18]. As potential solutions, different approaches try to harness the strengths of natural language for enhancing the understandability, usability and ubiquity of process mining.

1) Process Mining Querying: Works based in the idea of Natural Language Interfaces to Databases propose querying process information through natural language. In [9], authors extend the ontology-based ATHENA system [19] with a method that derives a domain ontology from process data to support natural language queries over event logs stored in an ElasticSearch database. Authors in [3] take a similar approach, proposing a natural language interface that allows to query existing process mining tools. They introduce a reference architecture for a conversational interface to process mining (only addressing the querying part) that translates natural language queries into logical representations that are ultimately translated into API calls to process mining tools.

2) Natural Language Generation: Classical NLG systems are pipeline-based, and use domain knowledge as well as linguistic knowledge in order to automatically generate texts that summarize the most insightful aspects of some non-linguistic input data [16]. In particular, D2T systems are concerned with natural language generation from numeric input data [15]. In the process mining domain, authors in [11], [12] propose a pipeline for generating natural language descriptions of process models in order to maintain stable representations of process models during their life-span. Focusing on supporting process model validation and inconsistency detection through natural language.

Previous to NLG approaches, Fuzzy Set Theory [23] was introduced with the aim of modeling and managing the underlying uncertainty naturally present in natural language. In this approach, the semantics of linguistic terms are modeled through fuzzy sets. Then, these fuzzy sets are integrated into predefined structures or templates, known as protoforms, in a way that allows to summarize data (with some degree of uncertainty present on them) in a linguistic manner, in the form of fuzzy quantified sentences [21], [22]. In the literature, authors in [4], [20] propose the use of fuzzy quantified sentences to summarize event log data. By modeling the semantics of event log data through fuzzy sets, linguistic summaries can be generated. However, as no other process mining techniques are applied (discovery, replay, etc.), the summarized information is constrained to that present on an event log; lacking control-flow or time related information that may be of utter interest.

Combining both perspectives, authors in [6], [7], propose a framework for the generation of natural language descriptions of processes based on a series of protoforms introduced in [5]. Leveraging process mining techniques, not only control-flow (process model) information is used, but descriptions are enriched with quantitative and qualitative information about past process execution extracted by mining and replaying event logs. These descriptions were evaluated as valid, understandable and insightful by domain experts, effectively helping non-technical users to better understand processes.

We can conclude that current proposals are limited to the static generation of natural language descriptions of a process, and predominantly focused on process model or event log descriptions, which do not provide a complete understanding of what is actually happening in a process. To the best of our knowledge, the approach presented in [7] is the only one supporting all process mining perspectives [1] and thus providing users with complete, informative descriptions of a process. Based on these ideas, in this paper, we propose a new reference ontology-based architecture leveraging process mining techniques and fuzzy set theory for the interactive generation of natural language explanations of processes. The use of an ontology allows to define semantic entities and their relationships in the process mining and fuzzy sets domain, allowing non-technical users to query the system for natural language explanations of processes.

III. An Ontology-Based Architecture for the Explanation of Processes

This section introduces our proposal, centered around an ontology-based architecture designed for the interactive generation of natural language explanations of processes. The core of our system is this ontology, which encapsulates the semantic relationships and properties within the process domain, integrating concepts from fuzzy set theory to handle linguistic uncertainty. The proposed architecture, depicted in Figure 1, supports both online and offline modes of interaction through a user-friendly web-based interface. Online interaction allows users to input queries and receive natural language explanations in real-time. Conversely, in the offline mode, expert users can enhance the system by adding new semantics

[1] Each of the dimensions in which a process can be analyzed: organizational, control-flow, case, and time perspectives [2].

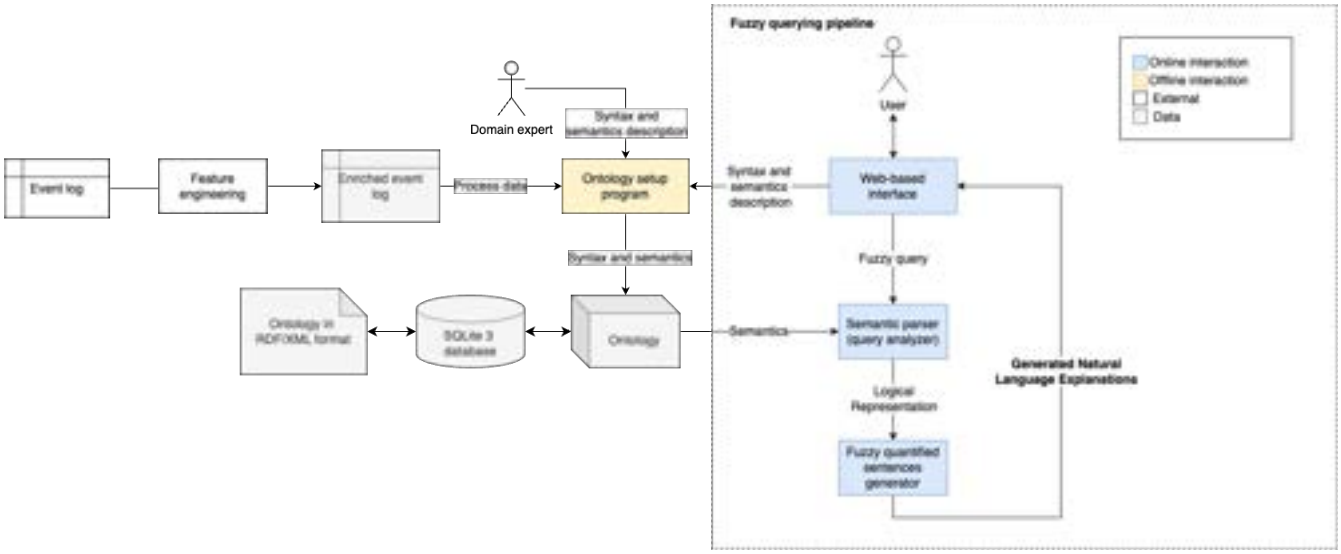


Figure 1. Architecture overview of the proposed system

to the ontology, further refining the explanation capabilities of the system.

A. Ontology

The domain ontology is a central piece in the architecture, as it captures a semantic description of the process and fuzzy quantification domains in terms of relevant entities and their relationships. Based in a simplified version of the XES meta-model [1], we extend the ontology with concepts from fuzzy set theory, which allow to capture the semantics of linguistic terms (words or sentences) via fuzzy sets. Having the semantics of these linguistic terms modeled in the ontology is what allows the system to ”understand” the terms used in a query, and then use these semantics for generating fuzzy quantified sentences from an event log. The semantics of these terms are modeled in the ontology by a domain expert during the design phase, or by a user during the so-called offline interaction. Furthermore, on the ontology building stage, the system automatically populates the ontology from the event log data.

B. Feature engineering: Event log enrichment

As a previous stage to the proposed system, event log enrichment may take place. Applying different process mining techniques (process discovery, replay, etc.) new features of a process can be discovered and be used as new process, case or event attributes. Through feature engineering and feature selection, the original event log is enriched for example with temporal or control-flow related information derived from the original log. This would definitely allow for the generation of more interesting natural language explanations.

C. Fuzzy Quantified Sentences

Fuzzy quantified sentences provide a way of handling uncertainty and expressing it in natural language, allowing for the quantification of the most relevant characteristics of some data.

1) Linguistic Variable: Summarizer, qualifier and quantifier are defined as linguistic variables, used to model the partitioning of the domain of a non-linguistic variable into several linguistic values. Being a linguistic value a linguistic term with a label (word or sentence) and some semantics modeled through a fuzzy set. A fuzzy set A in a referential X is identified by a membership function μ_A, that assigns to each element $x \in X$ a grade of membership in [0,1] i.e. the degree to which x satisfies the property indicated by A.

Among many others, linguistic quantifiers can be relative or absolute: a relative linguistic quantifier relates the proportion of elements of the referential that fulfill the property declared by the summarizer e.g., ”most”, ”almost all”, and is represented as a fuzzy set $Q \in [0, 1]$; absolute linguistic quantifiers denote absolute quantities in an imprecise way e.g., ”about three”, ”much more than five”, and are defined as fuzzy sets in the non-negative real line $Q \in R^+$.

2) Generation of Fuzzy Quantified Sentences: The generation of fuzzy quantified sentences consists in the instantiation of a *protoform* with some linguistic values for the quantifier, qualifier and summarizer, and computing the truth degree of the sentence to the summarized data. We will be using Zadeh’s quantification model [22] for the truth degree evaluation, although any other valid model could be used [8]. Assume Z is some case attribute which can take values (numeric or non-numeric) in $V = \langle v_1, v_2, ...\rangle$ e.g. cost, sex, etc. $X = \langle x_1, x_2, ..., x_n\rangle$ is the set of n cases from an event log that show attribute Z, we use z_i to indicate the value of attribute Z for case x_i. Then, $Z = \{z_1, z_2...\}$ is the collection of observations of attribute Z for the cases in Y. If we define a summarizer A in V, and a quantifier Q, they can be used to linguistically summarize attribute Z for the referential X. T indicates the truth of the statement that Q elements of the

collection X satisfy the sentence S.

For a type-I sentence, Zadeh's quantification model is defined as:

$$T = \mu_Q \left(\frac{1}{n} \sum_{i=1}^{n} \mu_A(z_i) \right) \quad (1)$$

Inside the parentheses, the membership degree sum of all elements of a referential to a fuzzy set is called absolute cardinality; if divided by the total number of elements in the referential n, is called relative cardinality.

Notice, than more than one attribute can be used as summarizer e.g. "most students are *tall and young*". In this case, the membership degrees to both attributes are aggregated, in our case t-norm minimum, denoted with the "$\wedge$" operator.

For type-II sentences, a qualifier B of V is added. For this type of sentence, Zadeh's quantification model is as follows:

$$T = \mu_Q \left(\frac{\sum_{i=1}^{n} \mu_A(z_i) \wedge \mu_B(z_i)}{\sum_{i=1}^{n} \mu_B(z_i)} \right) \quad (2)$$

As before, more than one attribute can be used as qualifier e.g. "most students that are *tall and young* play basketball". Also, it may be interesting to not use a quantifier at all; we will refer to sentences like this as *qualitative sentences*, and use as its truth value, the relative cardinality of the summarizer. The challenge then is to automatically generate true fuzzy quantified sentences from an enriched event log. Usually, this is done by performing an exhaustive search through the semantic space[2]. However, this search can be computationally prohibitive, thus why we employ Kacprzyk and Zadrozny's proposal of an interactive approach for the generation of fuzzy quantified sentences [10].

3) Fuzzy Quantified Sentences in the NLG pipeline: The NLG pipeline consists of 3 modules each of which is subdivided in different tasks [16]. The content determination task is focused on deciding what information should be conveyed in the final generated text. As protoforms are relatively fixed, the purpose of a protoform-based approach is mainly to determine appropriate linguistic values for the summarizer, qualifier and quantifier (and any other possible element) i.e. what information is relevant and should be conveyed. In this sense, the generation of fuzzy quantified sentences is mainly concerned with this task. Nevertheless, fuzzy quantification protoform-based approaches are valid in the framework of more complex NLG systems, as fuzzy quantification techniques provide ways of handling uncertainty in natural language, they can be used in the content determination stage, aiding on the summarization of the most relevant aspects of the non-linguistic input data [14]. Furthermore, through the use of a human-computer interaction interface (as we propose), and with the use of protoforms, interactive and more sophisticated texts can be generated without the need of a heavy user-interaction.

[2]The power set of all the defined semantics: quantifiers, qualifiers and summarizers

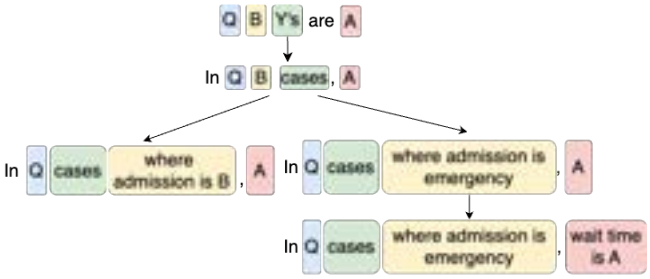


Figure 2. Example of a taxonomy of protoforms.

D. Querying the system

To query the system, the user directly employs linguistic terms whose semantics are modeled in the ontology; then a semantic parsing module is in charge of retrieving the semantics associated to each term. Before a linguistic term may be used in a query, it has to be modeled and "stored". This modeling is performed beforehand by a domain expert, however, users can add its own syntax and semantics through the offline-interaction with the system. By modeling the semantics of different linguistic terms in the ontology, their reutilization is possible. For modeling the queries, we consider the summarizer and qualifier as abstract fuzzy sentences like "X is A", where X is a linguistic variable describing an attribute of referential Y and A is one of its possible values e.g., X = *age*, A = *young* yield "*age* is *young*". The value of A can also be left empty, only indicating which linguistic variable is of interest and querying for its possible values. The derivation of protoforms from a query then proceeds as:

1) The user formulates a query as a protoform of interest and assigns linguistic variables and values as desired, leaving the elements of the protoform to which is looking for, empty. For example: "In Q cases *where admission is emergency*, A". This query would allow to quantify how *emergency admission* affects any other features of the process.
2) The system retrieves the semantics associated to each term in the query from the ontology and parses the query. The fuzzy set associated to each of the input terms is retrieved.
3) Then, instantiates the input protoform in all possible ways, replacing abstract symbols i.e. A and Q in the example query, with the semantics that are in the ontology. Following Zadeh's calculus of fuzzy quantified propositions, all fuzzy quantified sentences compatible with the query are generated and its truth value is computed.
4) All valid sentences are returned to the user ordered by truth value.

So, when the user inputs a query, the system has to instantiate all possible protoforms in the semantic space that comply with the query, giving place to a protoform taxonomy, in which at each level, more protoform elements are instantiated and protoforms are less abstract. An example can be seen in Figure

Table I
CLASSIFICATION OF POSSIBLE QUERIES.

Query	Given	Sought	Intent
In ___ cases, attribute *waiting time* is *long*.	S	Q	Summaries through ad-hoc queries
In ___ cases where attribute *waiting time* is *short*, attribute *age* is *much older than 80*.	S B	Q	Conditional summaries through ad-hoc queries
In *many* cases, attribute *waiting time* is ___.	Q $S_{structure}$	S_{value}	Simple value oriented summaries
In *many* cases, attribute ___ is ___.	Q	S	Exploration value oriented summaries
In *some* cases where attribute *admission* is *emergency*, attribute *waiting time* is ___.	Q B $S_{structure}$	S_{value}	Conditional value oriented summaries
In *some* cases where attribute *admission* is *emergency*, attribute ___ is ___.	Q B	S	Conditional consequence oriented summaries
In *some* cases where attribute ___ is ___, attribute *waiting time* is *long*.	Q S	B	Cause of given feature oriented summaries
In ___ cases, attribute ___ is ___.	Nothing	S B Q	General fuzzy rules

2. At the top of we have a completely abstract protoform, in the next level, we have instantiated the referential as the cases in a process.

On the right branch qualifier B is instantiated to "*admission* is *emergency*", below, linguistic variable "*waiting time*" is selected for the summarizer, but no linguistic value is chosen. In the left branch, only the linguistic variable has been selected for the qualifier.

The more abstract the query is, the less we assume, the more sentences are returned to the user. To limit the search and reduce the computational cost, we limit the generation to sentences with a maximum of two aggregated attributes per qualifier and summarizer, and with a truth degree of 0.7 or above. Also, we define a set of proportional quantifiers e.g. "almost none, some, many, etc." as a series of positive, non-monotonous fuzzy sets (trapezoids), reducing the redundancy of the generated sentences.

The input query then is constructed through a wizard-like web-based interface, were the user parts from an empty protoform (he/she can choose what type of protoform to use) and assigns linguistic variables and values where desired. Table I shows a series of queries, with the given inputs, the sought elements that the system has to complete, and a description of the user intent for each type of query. This classification is based on the classification proposed by Kacprzyk and Zadrozny's in [10].

IV. PROOF-OF-CONCEPT

As a proof-of-concept we illustrate how some explanations already validated can be generated. These explanations were generated in the context of a pipeline for the generation of natural language descriptions of healthcare processes. Particularly, of the Aortic Stenosis Integrated Care Process (AS ICP) implemented in the Cardiology Department of the University Hospital of Santiago de Compostela [7]. First, we identify the underlying protoforms from the generated explanations:

- In approximately half (52.72%) cases, patient had a wait time between its CT scanning and its intervention lower than expected.
- In approximately half (56.28%) cases, where patients were intervened via TAVI, a CT scan is performed short after its corresponding heart team meeting.

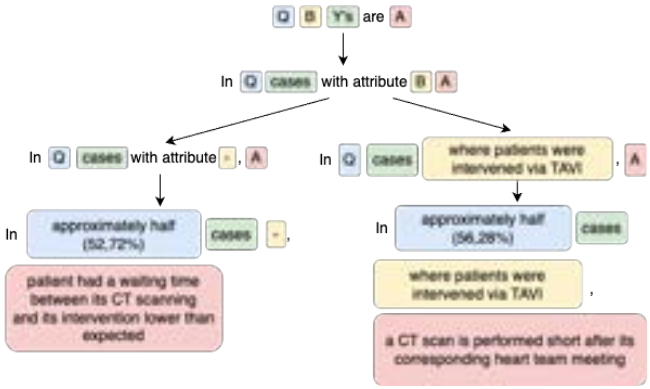


Figure 3. Generation of explanations of the AS ICP based on its underlying protoforms.

From these examples, we can identify type-I "In Q cases, R" and type-II "In Q cases with attribute C, R" protoforms. Where R makes reference to a temporal relation between two activities. These temporal relations are additional attributes computed in the log enrichment phase, where through process discovery and replay, temporal and frequency information can be discovered. In these examples, more complex summarizers than those seen until now are used; these kind of non-trivial, more elaborated summarizers are the most interesting. As their semantics are still human-modeled, and domain semantics are usually shared among users in a particular domain (for example, in the healthcare domain, clinical guides shared by all experts allow to define stable semantics for linguistic variables and values), these more complex summarizers are still human-consistent and allow for the generation of richer explanations. In Figure 3 we illustrate how the underlying protoforms of the generated informative explanations are generated following the taxonomy of protoforms that the system generates with an input query.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a process-mining-based, ontology-driven architecture and a system for the interactive generation of natural language explanations of processes through fuzzy quantified sentences. The ontology supporting

the architecture models the semantics of the process mining and fuzzy quantification domains, allowing to reason with process related data in a linguistic manner. Furthermore, utilizing fuzzy linguistic terms allows for the modeling of the inherent imprecision present in natural language, a really useful tool for summarizing data. The presented ontology is initially modeled by domain experts, however, through an web-based interface, expert users can add their own syntax and semantics to the ontology, making the system user-extendable, reusable and easily maintainable. We provide a novel ontology-driven algorithm for the interactive generation of natural language explanations. A proof-of-concept is presented, showing the capabilities of this approach to generate useful, understandable and informative explanations of processes from different domains.

As future work, the proposed architecture should be extended, both from the processing and generation sides. A deeper natural language processing pipeline as in [3], [9] can be implemented, giving support to queries in natural language. As we have seen, protoforms represent a powerful concept to formulate multiple types of sentences in a uniform way. However, this structure can be enriched with additional information: by modeling the semantics of protoforms into the system's underlying ontology, users could modify protoforms in a more flexible way. Adding the rest of the NLG pipeline stages to the system would also allow for richer natural language explanations: different phrasing of the same information, referring expression generation or aggregating phrases to generate more complex summaries. From the process standpoint, the ontology proposed in [17] can be incorporated into the proposed ontology, giving support to the event log enrichment stage. Furthermore, supporting natural language queries relating not yet discovered features, effectively computing the new feature (attribute) the user is interested in, and generating a natural language explanation summarizing it.

Acknowledgments

Thanks to Dr. Carlos Peña and Dr. Violeta González from the Department of Cardiology, University Clinical Hospital of Santiago de Compostela, SERGAS, Biomedical Research Center in the Cardiovascular Diseases Network (CIBER-CV), for providing the anonymized data.

References

[1] IEEE standard for extensible event stream (XES) for achieving interoperability in event logs and event streams. IEEE Std 1849-2016 pp. 1–50 (2016)
[2] van der Aalst, W.M.P.: Process Mining: Data Science in Action. Springer (2016)
[3] Barbieri, L., Madeira, E.R.M., Stroeh, K., van der Aalst, W.M.P.: Towards a natural language conversational interface for process mining. In: Process Mining Workshops. pp. 268–280. Springer International Publishing, Cham (2022)
[4] Dijkman, R.M., Wilbik, A.: Linguistic summarization of event logs - A practical approach. Inf. Syst. **67**, 114–125 (2017)
[5] Fontenla-Seco, Y., Bugarin, A., Lama, M.: Fuzzy temporal protoforms for the quantitative description of processes in natural language. In: 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–6 (2021)
[6] Fontenla-Seco, Y., Lama, M., Bugarín, A.: Process-to-text: A framework for the quantitative description of processes in natural language. In: Trustworthy AI - Integrating Learning, Optimization and Reasoning. pp. 212–219 (2021)
[7] Fontenla-Seco, Y., et al.: A framework for the automatic description of healthcare processes in natural language: Application in an aortic stenosis integrated care process. Journal of Biomedical Informatics **128**, 104033 (2022)
[8] Fuentes, A.C., Ramos-Soto, A., Diz, A.J.B.: An experimental study on the behaviour of fuzzy quantification models. In: Giacomo, G.D., Catalá, A., Dilkina, B., Milano, M., Barro, S., Bugarín, A., Lang, J. (eds.) ECAI 2020. vol. 325, pp. 267–274. IOS Press (2020)
[9] Han, X., et al.: Bootstrapping natural language querying on process automation data. In: 2020 IEEE International Conference on Services Computing (SCC). pp. 170–177 (2020)
[10] Kacprzyk, J., Zadrozny, S.: Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries, and natural-language generation. IEEE Transactions on Fuzzy Systems **18**(3), 461–472 (2010)
[11] Leopold, H.: Natural language in business process models. In: Lecture Notes in Business Information Processing (2013)
[12] Leopold, H., Mendling, J., Polyvyanyy, A.: Generating natural language texts from business process models. Lecture Notes in Computer Science **7328**, 64–79 (2012)
[13] Mannhardt, F.: Responsible Process Mining, pp. 373–401. Springer International Publishing, Cham (2022)
[14] Ramos-Soto, A., Bugarín, A., Barro, S.: On the role of linguistic descriptions of data in the building of natural language generation systems. Fuzzy Sets and Systems **285**, 31–51 (2016)
[15] Reiter, E.: An Architecture for Data-to-Text Systems. In: Proc. 11th European Workshop on Natural Language Generation. pp. 97–104. ACL, USA (2007)
[16] Reiter, E., Dale, R.: Building Natural Language Generation Systems. Cambridge University Press, USA (2000)
[17] del Río-Ortega, A., Resinas, M., Ruiz-Cortés, A.: Defining process performance indicators: An ontological approach. In: Meersman, R., Dillon, T., Herrero, P. (eds.) On the Move to Meaningful Internet Systems: OTM 2010. pp. 555–572. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
[18] Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. Journal of Biomedical Informatics **61**, 224–236 (2016)
[19] Saha, D., Floratou, A., Sankaranarayanan, K., Minhas, U.F., Mittal, A.R., Özcan, F.: Athena: An ontology-driven system for natural language querying over relational data stores. Proc. VLDB Endow. **9**(12), 1209–1220 (aug 2016)
[20] Wilbik, A., Dijkman, R.M.: Linguistic summaries of process data. 2015 IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE) pp. 1–7 (2015)
[21] Yager, R.R.: A new approach to the summarization of data. Information Sciences **28**(1), 69–86 (1982)
[22] Zadeh, L.A.: Fuzzy logic = computing with words. IEEE Transactions on Fuzzy Systems **4**(2), 103–111 (May 1996)
[23] Zadeh, L.: Fuzzy sets. Information and Control **8**(3), 338–353 (1965)

Sistemas de reglas difusas para el diseño de modelos de decisión en perfiles de pasajero

David Muñoz-Valero
Escuela de Ingeniería Industrial y Aeroespacial
Departamento de Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Toledo, España
david.munoz@uclm.es

Enrique Adrian Villarrubia-Martin
Escuela Superior de Informática
Departamento de Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
enrique.villarrubia@uclm.es

Julio Alberto López-Gómez
Escuela de Ingeniería Minera e Industrial
Departamento de Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Almadén, Ciudad Real, España
julioalberto.lopez@uclm.es

Juan Moreno-Garcia
Escuela de Ingeniería Industrial y Aeroespacial
Departamento de Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Toledo, España
juan.moreno@uclm.es

***Abstract*—El método por el que un pasajero compra un billete de tren es crucial para que las compañías ferroviarias puedan ofertar servicios adecuados, tanto al pasajero como al operador de la infraestructura. En este trabajo se estudia cómo se pueden utilizar los Sistemas de Reglas Difusas para modelar el proceso de toma de decisiones de un pasajero de tren a la hora de comprar un billete. Más concretamente, se han utilizado un método de generación de sistemas de reglas difusas que permite al experto diseñar estos sistemas. Se presentan dos ejemplos que ha permitido definir el comportamiento de dos tipos de pasajeros muy distintos en una primera aproximación muy preliminar. Estos modelos serán incorporados a nuestro simulador ROBIN para poder estudiar su comportamiento en detalle y el ajuste de los mismos.**

***Index Terms*—Modelo difuso de pasajeros, modelos de decisión, sistemas de reglas difusas, conjuntos difusos**

I. Introducción

La Unión Europea ha propuesto la liberalización de los mercados ferroviarios como medio de mejorar la utilización de las infraestructuras. Se trata de permitir que varias empresas ferroviarias operen en la misma red ferroviaria. Inicialmente, la mayoría de los países europeos mantenían una estructura monopolística del mercado ferroviario. Sin embargo, en las últimas tres décadas, la legislación europea, empezando por la Directiva 91/140 de 1991 [1], ha animado a varios países a adoptar un enfoque de mercado ferroviario compartido. En el caso concreto de España, el proceso de liberalización del mercado ferroviario se inició con la promulgación de la Ley 39/2003 del sector ferroviario. Sin embargo, no fue hasta 2021 [2] cuando la primera empresa ferroviaria extranjera comenzó a operar en un corredor ferroviario español de alta velocidad. Desde entonces, España ha implantado un modelo de mercado verticalmente separado, con un administrador de infraestructuras (ADIF) encargado de gestionar la capacidad de infraestructura. Simultáneamente, varias empresas ferroviarias (*Renfe*, *Ouigo* e *Iryo*) gestionan sus servicios de forma independiente. Para más detalles sobre las distintas estructuras de mercado [3].

Esta liberalización obliga a las empresas a estudiar y modelizar el comportamiento del mercado para poder realizar las mejores propuestas al administrador de infraestructuras de forma que le resulten más beneficiosas. Dentro de este entorno resulta de interés la obtención de modelos que sean capaces de representar el comportamiento de los pasajeros en la compra de billetes, lo que permite realizar las ofertas más adecuadas a lo que el pasajero desea en función de las características que considera en el proceso de compra del billete, tales como: precio, lugares de origen y destino, comodidad, etc. Dado que cada tipo de pasajero realiza la compra del billete considerando diferentes prioridades, por ejemplo, no es lo mismo la forma de comprar un billete de un estudiante que la de un CEO de una empresa. Para modelar esto se utilizará, se utilizará el concepto de "perfil de pasajero", que indica las preferencias consideradas por el pasajero en la compra del billete. En el diseño de estos perfiles se considerarán un conjunto de características que serán modeladas por un grupo de variables que permiten modelar la forma en que un pasajero toma la decisión de compra de un billete.

Para obtener estos modelos se pueden utilizar diferentes técnicas de modelización. Una forma común de hacerlo con-

Este trabajo ha contado con el apoyo de las ayudas PID2020-112967GB-C32 y PID2020-112967GB-C33 financiadas por MCIN/AEI/10.13039/501100011033 y por FEDER Una forma de hacer Europa y, el Departamento de Tecnologías de la Información y Sistemas y el Vicerrectorado de Investigación de la Universidad de Castilla-La Mancha. Se realizó cuando E.A. Villarrubia-Martín era becario predoctoral en la Universidad de Castilla-La Mancha financiado por el Fondo Social Europeo Plus (FSE+)..

Fig. 1. Logo del simulador ROBIN.

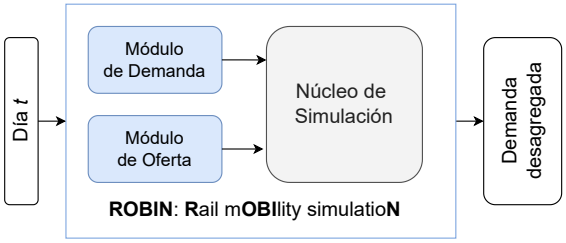


Fig. 2. Funcionamiento del kernel del simulador ROBIN.

siste en utilizar funciones matemáticas que calculen un valor de utilidad para el pasajero [4]. Otra posibilidad puede ser el uso de la lógica difusa (FL) [5], que se ha utilizado con éxito en un gran número de problemas y dominios diferentes. Además, la FL es muy útil para modelar perfiles de pasajeros gracias a la riqueza semántica que permiten los conjuntos difusos y las etiquetas lingüísticas [6]. Por todo ello, proponemos el uso de Sistemas de Reglas Difusas (FRS) para la creación de un modelo de decisión de pasajeros. Estos sistemas son eficaces, representan correctamente la incertidumbre inherente al comportamiento de los pasajeros y son fáciles de entender. Además, son explicables, lo que es absolutamente necesario en los modelos de Inteligencia Artificial utilizados actualmente en los sistemas automáticos.

ROBIN (Rail mOBIlity simulatioN) es un simulador microscópico para simular la movilidad ferroviaria en régimen de competencia. El simulador es parametrizable y lo suficientemente general como para simular flujos de pasajeros en sistemas ferroviarios en competencia. Además, permite modelizar de forma desagregada el comportamiento de los viajeros y sus elecciones de viaje en un mercado ferroviario competitivo. Dentro de este simulador, se utiliza una función de utilidad para modelizar la compra de un billete. Las Figuras 1 y 2 muestran su logo y la estructura de su kernel respectivamente. En este trabajo se propone utilizar FRS tipo TSK. Por ello, hemos diseñado un método para generar automáticamente un modelo de decisión que sirva de apoyo a una entidad en el proceso de toma de decisiones [7]. Admite variables de entrada que pueden tomar distintos valores (crisp, difusos o categóricos). El procedimiento genera automáticamente un conjunto de reglas difusas (FRS) que simulan el funcionamiento de la entidad en relación con la decisión a tomar. Este modelo se comporta de una forma similar a los modelos del tipo TSK. Las reglas de este modelo indican la importancia de las variables de entrada a la hora de generar el valor de salida. Finalmente, destacar que los modelos obtenidos tienen una alta interpretabilidad.

Se presentarán un par de ejemplos tipo de los modelos generados por nuestro método, analizando su comportamiento con el fin de mostrar que representan lo que se pretendía que modelaran. También se mostrará cómo un experto con pocos conocimientos de FL puede inducir un modelo a partir del método propuesto.

El resto del trabajo se estructura de la siguiente manera. En la Sección II se detallan algunas variables a considerar para la realización del modelo del pasajero y cómo pueden ser representadas para que el método presentado pueda generar el modelo. Se mostrarán dos ejemplos de FRS inducidos para la toma de decisiones sobre el perfil del pasajero. Finalmente, en la Sección III se proponen las conclusiones y el trabajo futuro.

II. Diseño de FRS para el modelo del pasajero

En el diseño de cualquier sistema de modelado es muy importante representar correctamente la entidad que se pretende modelar. Para esto se deben estudiar las características que definen la entidad y cómo se pueden definir. En nuestro caso, se deben considerar las características para modelar el proceso de toma de decisiones de un determinado perfil de pasajero a la hora de la compra de un billete de tren.

Dentro de este problema, el billete es un concepto clave ya que es lo que valorará el modelo de decisión. Para modelizar correctamente el billete, se pueden considerar algunas características definitorias. Cada característica estará representada por una o varias variables. En esta primera aproximación se utilizará una única variable para cada una de las características. Se utilizarán las siguientes características:

1) Temporal: seleccionan las fechas preferidas de viaje.
2) Económica: se refieren principalmente al coste de la plaza. Vendrán definidos por la capacidad económica del pasajero, así como por la imagen que quiera transmitir.
3) Imagen y reputación de las compañías: consideran la imagen de la compañía para el pasajero.

Para cada una de las variables que forman parte de las características, se diseñarán conjuntos difusos que se utilizarán en las reglas FRS. Cada uno de estos conjuntos se define dentro de un dominio. Por simplicidad, se utilizarán conjuntos difusos trapezoidales. Este tipo de conjuntos son ampliamente utilizados porque son computacionalmente eficientes en la implementación de estos sistemas. Además, son bastante intuitivos en su definición y en la interpretación de su significado. Las variables utilizadas en esta primera aproximación son el horario (SC - *schedule*) y el precio del viaje (P - *price*), junto con la reputación de la empresa para el pasajero (RE - *reputation*).

Para trabajar con el horario de viaje, proponemos el uso de conjuntos difusos que representan el intervalo de tiempo en el que el pasajero desea realizar el viaje. Por ejemplo, supongamos un pasajero que desea realizar un viaje de unas 2 horas de duración y prefiere salir temprano por la mañana

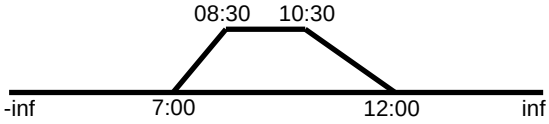


Fig. 3. Conjunto difuso para modelar las preferencias de horario del tren.

Fig. 4. Conjunto difuso utilizado para modelar las preferencias de precio.

y llegar con tiempo suficiente para comer al destino. Esto se puede modelizar mediante un conjunto difuso trapezoidal (denominado PAS_{SC}) que define, en un extremo, un intervalo de tiempo de salida y, en el otro, el intervalo de tiempo de llegada. La Figura 3 muestra un conjunto difuso que puede modelar este caso. Como puede observarse, el pasajero prefiere realizar el viaje entre las 8:30 y las 10:30 (instantes con pertenencia máxima), y admite salir sobre las 7:00 como muy temprano (primera hora con pertenencia mayor que 0), y llegar, como muy tarde, sobre las 12:00, última hora con pertenencia mayor que 0. Las horas estarán representadas por la tupla $fecha = (year, month, day, hour, minute, second)$. Este conjunto admite dos tipos de pertenencia, a un valor real v y a un intervalo $schedule = [init, end]$. En este trabajo utilizaremos la pertenencia al intervalo.

El valor de pertenencia de $[init, end]$ a SC se obtiene mediante la Ecuación 1 que calcula el grado de pertenencia medio de los valores v_i tomados de n en n minutos en el intervalo $[init, end]$ a SC.

$$\mu_{SC}([init, fin]) = \frac{\sum_{i=inic}^{end} \mu_{SC}(v_i)}{end - init} \quad (1)$$

Por ejemplo, dejemos que $schedule = [init, fin] =$ [08:15,10:45] sea la tupla que define el horario de un viaje concreto, donde $init$ es la hora de inicio del viaje y end es la hora de finalización del viaje. En este caso, el tren sale a las 08:15 y llega a destino a las 10:45. Por ejemplo, si se toman valores cada 15 minutos, se obtiene la siguiente secuencia de valores de grado de pertenencia $[0.83, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0.83]$ para cada elemento de la lista de horas [8:15, 8:30, 8:45, 9:00, 9:15, 9:30, 9:45, 10:00, 10:15, 10:30, 10:45] que da el valor agregado 0.97.

Para la modelización del precio, se utiliza la variable *Precio* (P), que toma un valor real definido en $[0, 100]$ donde los valores cercanos a 0 representan los precios más baratos y 100 será el billete más caro posible. Por lo tanto, el precio del billete debe normalizarse entre 0 y el precio máximo del billete considerado. En el modelo, esta variable será un conjunto difuso que modela los precios admitidos por el pasajero. Por ejemplo, supongamos que un pasajero de bajos ingresos desea viajar, su preferencia en relación con la variable P puede

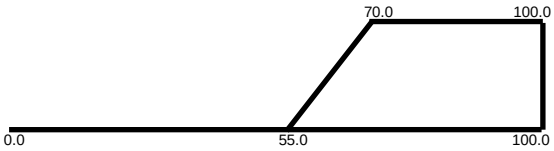


Fig. 5. Conjunto difuso para modelar las preferencias de la reputación de la compañía.

modelizarse mediante el conjunto difuso que se muestra en la Figura 4. Como puede verse, sus preferencias se aproximan al valor 0, lo que indica que desea un billete barato.

Por último, con respecto a la reputación de la empresa, se definirá un conjunto difuso que define la reputación de la empresa definida con un soporte de 0 a 100. Un valor más alto indica una mejor reputación para el pasajero, por ejemplo, el conjunto representado en la Figura 5 muestra una empresa que tiene una reputación alta para el pasajero.

Tras detallar cómo modelar las preferencias de los pasajeros, se mostrará la modelización de dos tipos de pasajeros muy diferentes: un estudiante y un director general de una gran empresa. Se han seleccionado los FIS (sistemas de inferencia difusa) TSK [8]. Se ha utilizado un método de inducción que genera reglas similares a las reglas de los TSK FIS que constan de un fuzzificador, un conjunto de reglas TSK y un motor de inferencia. El motor de inferencia realiza procesos de inferencia basados en las variables de entrada utilizando el conjunto de reglas TSK para generar el valor de salida final. La clave para la inducción automática de TSK FIS es la generación de reglas [9]. Se pueden encontrar muchos métodos en la literatura, como el presentado por Salimi y Mehdi [10]. Pensamos que estos sistemas pueden ser adecuados para ser generados automáticamente ya que el valor del consecuente es una función de las variables de entrada. Esta dependencia puede ayudar a la inducción de estos sistemas de reglas. Concretamente, hemos utilizado el método presentado en [7] para la inducción de los modelos que se presentará a continuación.

TABLE I
DISEÑO DE PERFILES DE PASAJERO

STD **(estudiante)**		
STD_{SC}	STD_P	STD_{RE}
[00:00, 08:00, 19:00, 24:00]	[0, 0, 15, 20]	[0, 0, 100, 100]
CEO **(CEO)**		
STD_{SC}	STD_P	STD_{RE}
[7:00, 8:30, 10:30, 12:00]	[0, 0, 100, 100]	[80, 90, 100, 100]

SC (horario), P (Precio), RE (reputación).

El perfil del estudiante se caracteriza por una baja capacidad económica y disponibilidad de tiempo para realizar el viaje ya que tiene una residencia en el destino donde permanecerá un tiempo. Además, no le importa la reputación de la compañía si le ofrece un billete barato. Por todo esto, a nuestro método de inducción se le indica la aparición de variables de forma individual, no con otras en el antecedente y que el consecuente

debe tener más peso en la regla relativa a P. El modelo inducido que representa a un estudiante es el siguiente:

R_1 : IF SC is STD_{SC} THEN 20
R_2 : IF P is STD_P THEN 80

Por otro lado, deben diseñarse los conjuntos STD_P y STD_{SC} para cada uno de los estudiantes que utilicen el modelo. La Tabla I muestra un ejemplo concreto que refleja las características expuestas anteriormente: poca capacidad económica y le importa poco el horario. STD_P debe tomar valores cercanos a $0, 0$, por ejemplo, puede ser $STD_P = [0, 0, 15, 20]$. En cuanto a STD_{SC} dependerá del tiempo que el alumno quiera estudiar, por ejemplo, supongamos que no quiere madrugar demasiado y llegar a destino antes de que anochezca. La duración del viaje no le importa mucho. En este caso podría ser $STD_{SC} =$ [00:30, 08:00, 19:00, 24:00]. El FRS no cuenta con ninguna regla asociada a la reputación debido a que no le importa al estudiante.

El perfil del CEO está marcado por una gran capacidad económica, no le importa el coste del billete y tiene un gran interés en que el viaje se realice a las horas que él desea. Además, le importa la imagen que tiene de la empresa, si no confía en la empresa no quiere viajar con ella, no puede fallar en su viaje. En este caso ocurre algo similar al caso anterior, no interesa la aparición conjunta de variables en los antecedentes, y, la variable más destacada para el CEO es SC, luego tendrán un peso más destacado en el consecuente. Este puede ser su modelo:

R_1 : IF SC is CEO_{SC} THEN 70
R_2 : IF RE is CEO_{RE} THEN 30

En este caso, el CEO utiliza las variables SC y RE, la variable P no se utiliza porque no le importa (Tabla I). El CEO viaja temprano y supongamos que el viaje dura aproximadamente 2 horas. En este caso podría ser $STD_{SC} =$[7:00, 8:30, 10:30, 12:00]. Además, quiere una reputación alta de la empresa, por lo que STD_{RE} debería tomar valores cercanos a $100, 0$, por ejemplo, podría ser $STD_{RE} = [80, 90, 100, 100]$.

Se han diseñado dos FRS con el método propuesto para proporcionar algunas ideas para modelizar la compra de un billete de tren de acuerdo al "perfil del pasajero". Estas consideraciones son:

1) El FRS generado utiliza en el antecedente las variables en las que está interesado el pasajero según su perfil. Las variables que no le interesan no se hace uso de ellas en el antecedente.
2) El consecuente ofrece un mayor valor de salida en las reglas que utilizan las variables que más interesan al pasajero y está ponderado entre 0 y 100 (si consideramos las reglas) por la forma de funcionar del método de inducción utilizado. Esto indica claramente las preferencias del pasajero estudiando el consecuente.
3) La combinación de las variables de entrada para calcular automáticamente el valor de salida es importante en los FRS generados. Si bien en este documento no se ha mostrado ningún ejemplo por la temprana fase de desarrollo en la que nos encontramos.

III. Conclusiones

Este artículo se ha presentado la utilización de un método de inducción para generar FRS que modelan perfiles de pasajeros a la hora de la compra de un billete. Los modelos generados pueden ser útiles para ayudar a las compañías a representar los perfiles de los distintos pasajeros. Pueden ser muy útiles en simuladores del mercado ferroviario para la modelización de los diferentes tipos de usuarios de trenes.

Las TSK FIS son adecuadas para este fin, ya que el valor del consecuente de las reglas es función de las variables del antecedente, lo que permite calcularlas automáticamente. Además, se han identificado tres consideraciones que ayudarán en el diseño de este sistema automático.

Como trabajo futuro, pretendemos la creación de FRS más detallados que consideren más variables para definir de una forma más precisa el comportamiento del pasajero. Además, debemos integrar dichos modelos en nuestro simulador [4] reemplazando la función de utilidad que utilizamos en la actualidad. El entorno versátil que ofrece el simulador permitirá comprobar la eficacia de los modelos obtenidos para cada perfil de pasajero, así como su ajuste a las diferentes situaciones.

References

[1] EC. 1991. Council Directive 91/440/EEC of 29 July 1991 on the Development of the Community'srailways, European Commission.
[2] ADIF AV. Adif y Adif AV aprueban el Horario de Servicio 2021-2022, que arrancará el próximo 12 de diciembre. ADIF AV Press Releases. 1st October (2021) https://www.adifaltavelocidad.es/w/adif-y-adif-av-aprueban-el-horario-de-servicio-2021-2022-que-arrancar%C3%A1-el-pr%C3%B3ximo-12-de-diciembre (Last visited 10 Jan. 2024)
[3] Ait Ali, A.: Methods for Capacity Allocation in Deregulated Railway Markets Doctoral thesis,comprehensive summary, Linköping University Electronic Press (2020)
[4] Castillo-Herrera E., García-Ródenas R., Jimenez-Linares L., López-García M.L, López-Gómez J.A., Martín-Baos J.A. , Moreno-Garcia J., Munoz-Valero D., Rodriguez-Benitez L., Villarrubia-Martin E.A. ROBIN: Rail mOBIlity simulatioN, Transportation Research Procedia, 2023.
[5] Zadeh L.A.. Fuzzy sets, Information and Control, volume 8, Issue 3. 1965. Pages 338-353.
[6] Zadeh L.A.. The concept of a linguistic variable and its application to approximate reasoning—I, Information Sciences, volume 8, Issue 3, 1975, 199-249.
[7] Muñoz-Valero D., Moreno-Garcia J., López-Gómez J.A., Villarrubia-Martin E.A., Jimenez-Linares L., A framework for the automatic generation of fuzzy decision models based on expert knowledge to support the decision-making process, Enviada a IEEE Transactions on Fuzzy Systems, 2024.
[8] Takagi T., Sugeno M.. Fuzzy identification of systems and its applications to modeling and control,IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-15, no. 1, pp. 116-132, Jan.-Feb. 1985.
[9] Chen T., Shang C., Su P., Shen Q., Induction of accurate and interpretable fuzzy rules from preliminary crisp representation, Knowledge-Based Systems, volume 146, pp.152-166, 2018.
[10] Salimi-Badr A., Mehdi Ebadzadeh M., A novel learning algorithm based on computing the rules'desired outputs of a TSK fuzzy neural network with non-separable fuzzy rules, Neurocomputing, volume 470, pp.139-153, 2022.

A new decision-making method for the assessment of investments and companies based on fuzzy mid-points

Carles Mulet-Forteza
Dept. of Business Economics
Balearic Islands University
Palma de Mallorca, España
carles.mulet@uib.es

Miquel A. Serra-Moll
Dept. of Mathematics
and Computer Science
Balearic Islands University
and
Health Research Institute
of the Balearic Islands (IdISBa)
Hospital Universitari Son Espases
Palma de Mallorca, España
m.serra@uib.cat

Oscar Valero
Dept. of Mathematics
and Computer Science
Balearic Islands University
and
Health Research Institute
of the Balearic Islands (IdISBa)
Hospital Universitari Son Espases
Palma de Mallorca, España
o.valero@uib.es

***Abstract*—In decision-making, we face a situation where we have to decide whether it is worth performing an investment or not, based on a multitude of variables. The Net Present Value (NPV) indicates the profit that our investment will produce over a series of time periods within a fixed time horizon. Its computation begins with certain data or variables from which we can obtain a cash-flow for each time period under consideration. Some of these variables involved in the aforementioned computation can take values within an interval of possibilities whose end-points are their pessimistic forecast (PF) and their optimistic one (OF). From these values, the expert subjectively generates two distinct cash-flow projections: the Optimistic Cash Flow (OCF) and the Pessimistic Cash Flow (PCF). The Net Present Value (NPV) is then computed using a cash-flow obtained as the arithmetic mean of both projections. However, there are infinite possible intermediate cash-flows that could be considere to obtain the NPV. Given the lack of an analytical reason to select the arithmetic mean cash-flow as the most appropriate one, and considering that such a choice could be derived from an appropriate aggregation function, in this paper we introduce a new decision-making method for the assessment of investments and companies. This method is based on the use of fuzzy mid-points and aggregation functions. Moreover, the new methodology incorporates a penalization, fixed following an analytical procedure and thus reducing subjectivity, for the elapsed time and for the discrepancy between OCF and PCF in order to generate the NPV. Finally, all methodologies are tested by applying them to a paradigmatic example where real data is considered and the NPV of a five-year hotel assessment is computed. Here, it is illustrated that the proposed methodologies could be appropriate for the assessment of investments and companies, particularly when operating under uncertainty.**

***Index Terms*—Aggregation Function, Decision-Making, Company Assessment, Investment, Mid-Point, Time-Dispersion Penalization.**

This research is part of project PID2022-139248NB-I00 funded by MICIU/AEI/10.13039/501100011033 and "ERDF A way of making Europe".

I. Introduction

The Net Present Value indicates the profit that our investment will produce over a series of time periods within a fixed time horizon. First of all, we need some data or variables from which we can obtain a cash-flow for each time period under consideration. These variables can take values in an interval of possibilities whose end-points are the pessimistic forecast (PF) and the optimistic one (OF).

Many times, the expert or decision maker provides two distinct projections of cash-flows: the Optimistic Cash-Flow (OCF) and the Pessimistic Cash-Flow (PCF). Now, we are in front of infinite different possible forecasts, all those values between the OCF and the PCF, and the expert in charge needs to determine which data is more appropriate or representative among them in order to make his/her final decision, i.e., in order to fix a concrete value as suitable forecasting for that cash-flow. The NPV is then computed using an intermediate cash-flow determined through subjective judgments. Certainly, this process can be approached in various ways and generally, there is no analytical method to determine the aforementioned intermediate cash-flow value. Typically, such a value is set heuristically according to the expert's criteria. In order to avoid this handicap, a typical approach is to set such a value by taking the arithmetic mean between the PCF and OCF.

The classical methodology, based on the Certainty Equivalent, assumes that the generation of each flow involved in the NPV computation is independent, on the one hand, of the year in which it is generated and, on the other, of the dispersion or discrepancy between the values PCF and OCF. The classical methodology operates by estimating some reduction coefficients, which are subjectively determined by the expert in charge. These coefficients are used to penalize the forecasts of the intermediate cash-flow corresponding to

more distant periods, due to the uncertainty associated with them. Depending on the estimation of the aforementioned reduction coefficients, it can lead to very different results and thus, the final assessment is highly sensitive to the choice of the coefficients. Observe that such a NPV value may not be the most reasonable and credible forecasting, since it has been fixed by means of subjective judgments.

Motivated by exposed facts, the objective of this study is to propose an analytical technique based on the use of L-fuzzy sets in the sense of [1], aggregation functions and mid-points, and thus applying fuzzy logic. This new methodology operates similarly to the classical one. It considers two distinct projections of cash-flows obtained for each year under consideration: the Optimistic Cash-Flow (OCF) and the Pessimistic Cash-Flow (PCF). Consequently, we are in front of infinite different possible intermediate cash-flow forecasts and we need to determine which one is more appropriate among them in order to make the final decision. The mentioned intermediate cash-flow is stated by means of the use of aggregation functions. Moreover, in contrast to the classical methodology, the new technique incorporates two penalizations: one for the elapsed time and the second one for the discrepancy between the PCF and the OCF in order to generate the NPV. The aforementioned penalizations are incorporated in the mathematical computation by means of weights that, unlike the case of the Certainty Equivalent, are constructed analytically and not through subjective estimations.

The originality of our study lies in its application of fuzzy logic to the field of financial forecasting, offering an analytical and structured alternative to the subjective estimations used traditionally. Consequently, it introduces a new methodology based on the use of aggregation functions and fuzzy mid-points for dealing with the inherent uncertainties of financial forecasting.

The presented methodology is tested applying it to a paradigmatic example where real data is considered and the NPV of a five-year hotel assessment is computed. Such an example illustrates that the proposed methodologies could be appropriate for the assessment of investments and companies, particularly when operating under conditions of uncertainty.

The structure of this study is as follows: Section II presents a brief literature review on the subject in question. Section III explains the key concepts necessary to understand investment theory. Section IV presents our time-dispersion penalization methodology. This is followed by Section V, which showcases a practical application to a company assessment. Finally, in Section VI closing conclusions are provided.

II. Literature Review

The necessity of handling the aforesaid uncertainty in problem-solving has generated a large amount of literature. Over the years, research has shown that traditional tools issued from probability theory were felt to be insufficient to handle uncertainty [2]. In contrast, the fuzzy set theory developed by Zadeh [1], [3] presents a new point of view when operating with uncertainty. This theory was particularly designed to mathematically represent uncertainty and vagueness. Moreover, it provides formalized tools for dealing with many real problems where uncertainty takes place [4].

Since investment decision-making implies uncertainty, the fuzzy set theory appears to be an appropriate framework. A wide literature review on fuzzy investment decision-making is given in [5]. In the aforementioned reference, most of the works on fuzzy investment decision-making are related to the usage of discounted fuzzy cash-flows.

In this light, our study contributes to and increases the number of researches conducted based on the aforementioned premises. It demonstrates how investment criteria, which incorporate fuzzy logic through aggregation functions for decision-making, allow us to deal better with uncertainty. This is because the computation relies entirely on an analytical procedure, thereby subjectivity is reduced.

III. Investment theory

A. Net present value

When we face an assessment of an investment or company, it is essential to determine its profitability. For this reason, we need a measure so as to have a reference for the assessment or to decide whether it is profitable to proceed or not with the investment.

The Net Present Value (NPV) of an investment is defined as the sum of the present values of incoming (benefits) and outgoing (costs) cash-flows over a period of time. The computation of NPV takes into account the cash-flow chronology. This is accomplished by using a discount rate in order to homogenize the different cash-flows at different moments of time. In other words, the NPV is the value of an investment discounted to the present and therefore, it is useful to determine how much an investment is worth [6]. The computation of NPV is performed by the following formula: $NPV = -A + \sum_{i=1}^{n} \frac{Q_i}{(1+k)^i}$, where A is the Initial disbursement, Q_i is the cash-flow or difference between payments and collections generated by the investment associated to period i, n ($n \in \mathbb{N}$) is the Duration of the investment measured in years and k is the Cost of capital (discount rate).

B. Terminal Value

When we are in front of an investment project or a company assessment, the first step is to define the period of time that will be studied.

The terminal value is the project value at the end of the defined period of time, in other words, the price by which we understand that we could sell or transfer it. For this reason, to assess our project we must take into account not only the total cash-flows generated through the project life, but also the terminal value at its end (see [6]).

Regarding a company assessment, we usually consider that it will continue to generate income sine die, meaning that its end is not defined. In that case, it is impossible to make such long-term projections so we must define the terminal value as follows: $TV = \frac{Q_n \cdot (1+q)}{(k-q)}$, where q is the Expected Cash-flow growth.

C. Classical methodology: the certainty equivalent

The classical methodology uses a technique that allows introducing risk and uncertainty in the analysis of investment projects. This methodology is based on the use of the Certainty Equivalent. Remember that when we obtain a cash-flow, it incorporates a certain level of risk because it is obtained from expected variables, as mentioned before, set in many cases by subjective expert judgments. Following [7], in view of the preceding fact, the methodology consists in reducing the expected cash-flows of an investment treating its risk level individually.

To apply this technique, we must assign a reduction coefficient α_n for each period of time n (β for the terminal value). The coefficient values α_n are obtained by a subjective estimation of an expert and they will oscillate between 0 and 1 ($\alpha_n \in (0,1]$). Moreover, such coefficients are inversely proportional to the risk perceived by the expert for the period in which the cash-flow is estimated. Finally, we multiply each of the estimated net cash-flows at different periods of time by its associated reduction coefficient as follows:

$$NPV = -A + \frac{\alpha_1 \cdot Q_1}{(1+k)^1} + \frac{\alpha_2 \cdot Q_2}{(1+k)^2} + \cdots + \frac{\alpha_n \cdot Q_n}{(1+k)^n} + \frac{\beta \cdot TV}{(1+k)^n}.$$

IV. Time-dispersion penalization methodology

Before presenting the core concepts of our new methodology, it is crucial to recognize that it follows the fundamental principles of fuzzy logic philosophy. Each cash-flow involved in the computation of the NPV takes values in a closed interval and, hence, according to [1], they are considered as L-fuzzy sets, where L is a closed interval of non-negative real numbers. On account of [8], the considered aggregation functions are defined over an interval of non-negative real numbers which are not necessarily the unit interval. Notice that aggregation functions deal with interpolation within a range of values and such an interpolation provides the mid-point between OCF and PCF as L-fuzzy sets.

A. Aggregation functions

As exposed in Sections I and III-C, in order to compute the NPV value, the expert provides two distinct projections of cash-flows: the Optimistic Cash-Flow (OCF) and the Pessimistic Cash-Flow (PCF). This is achieved by establishing an optimistic forecast (OF) and a pessimistic forecast (PF) for those variables which inherently involve uncertainty. Later he/she computes the NPV by means of an intermediate cash-flow, between the OCF and the PCF, fixed by means of subjective judgments. Of course, this can be done in many ways and, in general, there is not an analytical method to determine the aforesaid intermediate NVP value. Typically, such a value is set heuristically according to the expert's criteria. So this provides that the estimation can be set heuristically as any value between the PCF and the OCF. In order to reduce the subjectivity to some extend, sometimes this value is fixed as the arithmetic mean value between OCF and PCF. In any case, the value taken by the expert as the most reasonable or credible forecasting is in fact a mid-point between PCF and OCF. Here, aggregation functions arise as a natural mathematical tool in order to merge PCF and OCF. Notice that aggregation functions capture mathematically the mid-point notion. Taking this into account it seems natural to consider aggregation functions as a suitable candidate to express mid-point selections in the estimation of the intermediate cash-flows and, hence, in the computation of NPV.

In the remainder of this section, we introduce the notion of aggregation function, which will play a central role in our time-dispersion penalization methodology. Let us recall that, according to [8] (see also [9]), an aggregation function is a function $f : [a,b]^2 \to [a,b]$ satisfying the following properties:

(i) $f(a,a) = a$ and $f(b,b) = b$,

(ii) $f(x,y) \leq f(z,w)$ provided that $x \leq z$ and $y \leq w$.

Of course, in the preceding definition we denote by $\mathbb{R}^+$ the set of nonnegative real numbers and, in addition, we are assuming that $a, b \in \mathbb{R}^+$ with $a < b$.

Notice that we are considering aggregation functions only with two arguments because this is enough for our purpose. For a fuller treatment of averaging aggregation functions we refer the reader to [8] and [9].

Among aggregation functions, we will focus our attention on the so-called averaging aggregation functions which generalize the classical arithmetic mean. In fact, the arithmetic mean is a concrete example of this kind of functions. An aggregation function f is said to be averaging provided that the following inequality holds for all $x, y \in \mathbb{R}^+$:

$$\min(x,y) \leq f(x,y) \leq \max(x,y).$$

In the computation of the aforesaid mid-points we will use only a few celebrated instances of averaging aggregation functions with the aim of illustrating our new methodologies for estimating the NPV: the Hurwicz Operator, the Minimum Operator and the Maximum Operator. Furthermore, the Arithmetic Mean will be used as a reference when comparing all numerical obtained results. Let us recall such aggregation functions for the convenience of the reader.

- The Arithmetic Mean M, where, given $x, y \in \mathbb{R}^+$,
$$M(x,y) = \frac{x+y}{2}.$$
- Maximum operator Max and Minimum operator Min, where, given $x, y \in \mathbb{R}^+$, we have
$$Max(x,y) = x_{\max}, \quad Min(x,y) = x_{\min}.$$
Notice that $x_{\max}$ and $x_{\min}$ denote the greatest and the lowest values of the vector input (x,y), respectively. Observe that, in this scenario, both the Maximum Operator and the Minimum Operator provide a mid-point equal to the one that an expert would give applying directly the Optimistic and Pessimistic Decision-Making Criteria which provide the extreme estimations, i.e., OCF and PCF respectively (see [10]).
- Hurwicz Operator HC, where, given $x, y \in \mathbb{R}^+$, this aggregation function finds a value between the optimistic

$x_{\max}$ and pessimistic $x_{\min}$ positions, from a coefficient α ($\alpha \in [0,1]$) known as coefficient of realism. This coefficient measures the degree of optimism of the decision maker, so it depends on the expert inclination. It ranges between 0 and 1 being 0 when the decision maker is totally pessimistic and 1 when he/she is totally optimistic (see [10] and [9]). Therefore, the mathematical expression is the following one:

$$HC(x,y) = \alpha x_{max} + (1-\alpha)x_{min}.$$

Observe that for $\alpha = \frac{1}{2}$, the Hurwicz operator becomes the Arithmetic Mean. Moreover for $\alpha = 1$ and $\alpha = 0$, it retrieves the maximum and the minimum operators, respectively.

B. Time-dispersion cash-flow penalization methodology

For each year n we compute two cash-flows. Both are computed as in the classical way. The first one OCF_n is obtained considering only the optimistic forecasts of those variables that admit both pessimistic and optimistic estimations. The second one PCF_n is obtained considering only the pessimistic forecasts of those variables that admit both pessimistic and optimistic estimations. Those variables that do not admit two estimations are treated as in the classical methodology.

Taking into account the exposed facts, the new methodology is based in the following four steps:

1) For each year n, we compute the Cash-Flows PCF_n and OCF_n.
2) For each year n, from both Cash-Flows (PCF_n, OCF_n), we obtain a weight w_n which only depends on the elapsed time n, PFC_n and OCF_n. A few examples proposed for analyzing the numerical example will be detailed later. Observe that this weight can be interpreted as a time-dispersion penalty of both cash-flows estimations, PCF_n and OCF_n.
3) Consider the value $HC(\text{PCF}_n,\text{OCF}_n)$ as the mid-point (realism estimation) of the Cash-Flow associated to year n, where the realism coefficient is fixed as w_n.
4) The NPV is obtained as in the classical case using the new mid-point calculated cash-flows and setting the reduction coefficients equal to 1.

Notice that the value of $HC(\text{PCF}_n,\text{OCF}_n)$ will decay following the time-dispersion penalization. However, this applies only to variables which are related to an income. In contrast, variables which represent an expense, will behave in an opposite manner since when the penalization is applied, the $HC(\text{PCF}_n,\text{OCF}_n)$ will be increased over time and dispersion.

C. A few possible realism weights

We end this section exposing different possible weights that can be applied to get the realism coefficient mentioned above. Of course, they are only instances that captures time-dispersion penalty. However, other weights could be set by the expert as realism coefficients for the assessment so that they adjust in the most appropriate way to the investments or company under consideration.

If (v_p, v_o, n) is a vector where v_p represents a pessimistic value, v_o an optimistic value and n the elapsed time period (years), the following examples are instances of time-dispersion penalties:

$$w_1(v_p,v_o,n) = \begin{cases} 1 & \text{if } n = 1 \\ \min\left\{\frac{\alpha}{n-1}, \frac{1}{\frac{|v_p - v_o|}{v_o}+1}\right\} & \text{if } n > 1 \end{cases},$$

$$w_2(v_p,v_o,n) = \begin{cases} 1 & \text{if } n = 1 \\ \frac{\alpha}{n-1} \cdot \frac{1}{\frac{|v_p - v_o|}{v_o}+1} & \text{if } n > 1 \end{cases},$$

$$w_3(v_p,v_o,n) = \frac{1}{(n-1)+(\frac{|v_p - v_o|}{v_p})+1}.$$

V. Practical application to the assessment of companies

Now that the methodologies to be tested are already explained, we propose a practical example with real data so that we can apply all the different methodologies, the classical one (arithmetic mean) and the time-dispersion cash-flow penalization methodology. Concretely, it is an example of a five-year hotel assessment.

The main objective is to obtain different NPVs with the different explained methodologies and to confirm that the estimation provided by the new methodology is reasonable and could be considered as suitable as that provided by the classical methodology, but now with the advantage of avoiding the high dependence of the NPV estimation caused by the reduction weights imposed by the certainty equivalent and chosen subjectively by the expert in charge.

Table I summarizes the values of the different variables that describe the company in the example. Notice that the cells represent, for each variable, an interval whose limits are the pessimistic and the optimistic forecasts respectively provided by an expert in charge and, thus, such a variable can be considered as an L-fuzzy set. Let us note again that those variables involved in the cash-flows computation that do not admit pessimistic and optimistic forecasts are not exposed in the aforementioned table because they are treated as in the classical case (see the GitHub repository).[1].

TABLE I
VARIABLES SUMMARY.

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
Pax/room	[2.8, 3.1]	[2.795, 3.11]	[2.79, 3.12]	[2.78, 3.13]	[2.77, 3.15]
OR[1]	[0.78, 0.81]	[0.77, 0.82]	[0.765, 0.825]	[0.76, 0.83]	[0.75, 0.835]
AIS[2]	[35.5, 37]	[36, 38]	[36.25, 38.5]	[36.5, 39]	[36.75, 39.5]
GOP/Sales	[0.225, 0.27]	[0.22, 0.28]	[0.215, 0.285]	[0.213, 0.29]	[0.21, 0.3]
Taxes/Sales[3]	[0.04, 0.02]	[0.04, 0.02]	[0.04, 0.02]	[0.04, 0.02]	[0.04, 0.02]
IG[4]	[0.04, 0.01]	[0.04, 0.01]	[0.04, 0.01]	[0.04, 0.01]	[0.04, 0.01]

[1] Occupancy Rate. [2] Average Income per Stay. [3] Fees, Insurances, other Taxes / Sales. [4] Investment Growth.

Once we have determined the PF and OF values for each variable exposed in Table I, it is the time to calculate the

[1]All necessary data regarding the company description, the values of all involved variables and all calculations required for the methodologies considered, both the classical and the newly introduced one, can be consulted in the documentation which can be downloaded from the following GitHub repository: https://github.com/mserrauibcat/A-new-decision-making-method-for-the-assessment-of-investments-and-companies-.git

profit-and-loss account for each year, which is necessary to obtain the various Q_n values. Depending on the methodology, such values match up with the PCF_n, the OCF_n or $M(PCF_n, OCF_n)$. All the calculations which are necessary to apply the different methodologies can be consulted step by step in the aforementioned Excel source code (see the GitHub repository). From now on, we will only present the results as a summary to facilitate reading.

The results

As extreme estimations, in the following we show the results after computing the NPV of the investment applying the Maximum and Minimum Operators (special cases of the Hurwicz operator). The values of the variables used to perform the aforementioned computations are shown in Table II for the Minimum Operator and Table III for the Maximum Operator.

A. Minimum operator

Notice that this is the equivalent to the pessimistic scenario, in other words, the pessimistic value of each variable have been used to calculate each cash-flow, resulting in the following estimation NPV $= 50642.14$ €. All PCFs (Q_{min}) are summarized in Table IV.

TABLE II
VARIABLES SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
Pax/room	2.800	2.795	2.790	2.780	2.770
OR[1]	0.780	0.770	0.765	0.760	0.750
AIS[2]	35.500	36.000	36.250	36.500	36.750
GOP/Sales	0.225	0.220	0.215	0.213	0.210
Taxes/Sales[3]	0.040	0.040	0.040	0.040	0.040
IG[4]	0.040	0.040	0.040	0.040	0.040

[1] Occupancy Rate. [2] Average Income per Stay. [3] Fees, Insurances, other Taxes / Sales. [4] Investment Growth.

B. Maximum operator

Notice that this is the equivalent to the optimistic scenario, in other words, the optimistic value of each variable have been used to calculate each cash-flow, resulting in the following estimation NPV $= 15412651.12$€. All OCFs (Q_{max}) are summarized in Table IV.

TABLE III
VARIABLES SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
Pax/room	3.100	3.110	3.120	3.130	3.150
OR[1]	0.810	0.820	0.825	0.830	0.835
AIS[2]	37.000	38.000	38.500	39.000	39.500
GOP/Sales	0.270	0.280	0.285	0.290	0.300
Taxes/Sales[3]	0.020	0.020	0.020	0.020	0.020
IG[4]	0.010	0.010	0.010	0.010	0.010

[1] Occupancy Rate. [2] Average Income per Stay. [3] Fees, Insurances, other Taxes / Sales. [4] Investment Growth.

C. Arithmetic mean

Having presented the pessimistic and the optimistic scenarios, it is also interesting to present, as a reference with which to compare, the mid-scenario which corresponds with the application of the Arithmetic Mean between the OCF and the PCF (the mid-point provided by the classical methodology and, again, a particular case of the Hurwicz operator). The obtained NPV estimation is the following one: NPV $= 7731646.63$ €. Thus, from each OCF and PCF a Q_{AM} is obtained. These are summarized in Table IV.

TABLE IV
PESSIMISTIC, OPTIMISTIC AND MID CASH-FLOWS SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
Q_{min}	913 877.50€	880 388.41€	845 869.20€	820 617.35€	788 593.32€
Q_{max}	1 318 076.06€	1 436 587.55€	1 499 351.62€	1 563 860.84€	1 666 695.19€
Q_{AM}	1 115 976.78€	1 158 487.98€	1 172 610.41€	1 192 239.09€	1 227 644.26€

D. Time-dispersion cash-flow penalization methodology

Notice that, from now on, H denotes the mid-point value obtained through the Hurwicz operator, in other words, the application of the Hurwicz to the optimistic and pessimist forecasts for the cash-flow under consideration using as the coefficient of realism one of the weights proposed in Subsection IV-C. Next, we apply the methodology introduced in Subsection IV-B testing the aforesaid weights. Recall that in this methodology we compute the Pessimistic Cash-Flow (Q_{min}) and the Optimistic Cash-Flow (Q_{max}) and, later, we apply the Hurwicz operator to get a mid-point Q_H. The coefficient of realism is one of the proposed weights in Subsection IV-C. Table IV summarizes the optimistic and pessimistic cash-flows obtained for each year.

1) Weight 1 (w_1): After the application of the methodology with the current weight, we have obtained the cash-flows shown in Table V. The obtained NPV estimation is the following one: NPV $= 4583834.59$ €.

TABLE V
MID-POINTS CASH-FLOWS SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
w_1	1.000	0.721	0.500	0.333	0.250
Q_H	1 318 076.06€	1 281 348.92€	1 172 610.41€	1 068 365.18€	1 008 118.79€

2) Weight 2 (w_2): After the application of the methodology with the current weight, we have obtained the cash-flows shown in Table VI. The obtained NPV estimation is the following one: NPV $= 3286214.27$ €.

TABLE VI
MID-POINTS CASH-FLOWS SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
w_2	1.000	0.721	0.348	0.226	0.164
Q_H	1 318 076.06€	1 281 348.92€	1 073 429.68€	988 552.16€	932 369.83€

3) Weight 3 (w_3): After the application of this methodology with the current weight, we have obtained the cash-flows shown in Table VII. The obtained NPV estimation is the following one: NPV $= 2945672.16$ €.

TABLE VII
MID-POINTS CASH-FLOWS SUMMARY

	Year 1	*Year 2*	*Year 3*	*Year 4*	*Year 5*
w_3	0.693	0.380	0.265	0.204	0.164
Q_H	1 194 125.34€	1 091 729.09€	1 019 089.21€	972 123.06€	932 226.48€

E. Summary results and discussion

To facilitate comparison of the different NPVs obtained through the application of the methodologies under discussion, we compile them in Table VIII.

TABLE VIII
NPV ESTIMATIONS SUMMARY.

Methodology	Net Present Value
Minimum Operator	50 642.14 €
Maximum Operator	15 412 651.12 €
Arithmetic Mean	7 731 646.63 €
Hurwicz Operator: w_1	4 583 834.59 €
Hurwicz Operator: w_2	3 286 214.27 €
Hurwicz Operator: w_3	2 945 672.16 €

Table VIII shows that the minimum and maximum operator results delimitate the extreme values that our analysis can give us, which have been fixed by the expert. Among these values, we find the results obtained after the computation of the proposed methodologies. This fact indicates that the new NPV values provided by our new methodology are reasonable and they could be equally appropriate for the assessment of investments and companies as those induced by the classical one (specially that given by the arithmetic mean).

Recall that the classical methodology operates under the estimation of some reduction coefficients fixed subjectively by the expert in charge for penalizing those forecasts of the cash flow corresponding to more distant periods. Of course, depending on the estimation of the aforementioned reduction coefficients, it can lead to very different results. This fact means that the estimation is highly sensitive to the choice of the aforementioned reduction coefficients. So there is not any analytical reason that allows us to conclude that the NPV estimation provided by the classical method (in particular by means of the Arithmetic mean) is the most appropriate. In fact, such an appropriateness is validated by the expert based on his/her experience and feeling.

Observe that w_1 gives us an estimation close to that of the Arithmetic Mean. However, the remainder estimations given by w_2 and w_3 could also provide us realistic perspectives (even more realistic than those obtained by the classical one) of our investment or company assessment. This is because the classical methodology operates regardless of the discrepancy between the optimistic and pessimistic forecasts.

Remember that results given by the Arithmetic Mean have been obtained considering the reduction coefficients equal to 1 but these reduction coefficients α_n for each period of time n (β for the terminal value) oscillate between 0 and 1 ($\alpha_n \in (0, 1]$). Moreover, such coefficients are inversely proportional to the risk perceived by the expert for the period in which the cash-flow is estimated.

VI. CONCLUSIONS AND FURTHER RESEARCH

We have introduced an analytical technique based on the use of aggregation functions and fuzzy mid-points to select an estimation of the NPV of an investment or company assessment. In contrast to the classical methodology, based on the certainty equivalent, the new one incorporates a penalization, fixed following an analytical procedure and thus reducing subjectivity, for the elapsed time and for the discrepancy between pessimistic PFC and optimistic forecasts OFC in order to generate the NPV. This fact constitutes an advantage in front of the classical methodology, where the reduction coefficients are determined by the subjective estimation of an expert and, as a consequence, the final assessment is highly sensitive to the choice of the coefficients. The new technique has been tested processing real data and computing NPV estimations of a five-year hotel assessment which are between the pessimistic and optimistic forecasts and, hence, being reasonable and could be appropriate for the assessment of investments and companies in general.

As a further work a study about what kind of weights could be more useful according to the investment type will be made. It also seems interesting to develop a software tool, implemented in Python, which could be incorporated in the expert's control panel with the aim of analyzing various scenarios by simply changing the desired value of variables and weights and, thus, giving the expert more support for decision-making. Moreover, another idea is to consider pre-aggregation functions and their utilities in order to apply them in our study. These functions generalize aggregation functions, which fulfill the same boundary conditions but demand only directional monotonicity [11]. Finally, it is important to note that in the literature, not only are fuzzy sets used to deal with uncertainty, but interval-valued fuzzy sets are also well-suited. The latter are a generalization of classical fuzzy sets where the membership values are intervals [12]. Hence, it could be interesting to analyze how to rewrite our study in this interval-valued setting.

REFERENCES

[1] J. Goguen, "L-fuzzy sets," *Journal of Mathematical Analysis and Applications*, vol. 18, no. 1, pp. 145–174, 2015.
[2] D. Dubois and H. Prade, "Possibility theory: qualitative and quantitative aspects," in *Quantified representation of uncertainty and imprecision*. Springer, 1998, pp. 169–226.
[3] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.
[4] C. Kahraman, "Investment decision making under fuzziness," *Journal of Enterprise Information Management*, vol. 24, no. 2, pp. 126–129, 2011.
[5] ——, *Fuzzy engineering economics with applications*. Springer, 2008, vol. 233.
[6] J. Berk, P. DeMarzo, and D. Stangeland, "Corporate finance (3rd canadian ed.)," 2015.
[7] A. A. Robichek and S. C. Myers, *Optimal financing decisions*. Prentice-Hall, 1965.
[8] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation functions*. Academic Press, 2009.
[9] G. Beliakov, H. B. Sola, and T. C. Sánchez, *A practical guide to averaging functions*. Springer, 2016.
[10] B. Render, R. M. Stair Jr., and M. E. Hanna, *Quantitative Analysis for Management*, 14th ed. United States: Pearson, 2022.
[11] H. Bustince, J. A. Sanz, G. Lucca, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, and G. Ochoa, "Pre-aggregation functions: definition, properties and construction methods," in *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2016, pp. 294–300.
[12] A. Bouchet, M. Sesma-Sara, G. Ochoa, H. Bustince, S. Montes, and I. Díaz, "Measures of embedding for interval-valued fuzzy sets," *Fuzzy Sets and Systems*, vol. 467, p. 108505, 2023.

El Rol de la Lógica Difusa en la Era de la Inteligencia Artificial de Propósito General.

Daniel Díaz[1,*], Juan José Herrera[1], Francisco Herrera[1], Isaac Triguero[1,2]
[1] Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence, Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada
[2] School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, United Kingdom
[*]danieldiaz@correo.ugr.es

***Resumen*—En Inteligencia Artificial existe una creciente necesidad de contar con modelos capaces de lidiar con un amplio espectro de tareas de aprendizaje sobrepasando las limitaciones de los sistemas diseñados para una única tarea. La reciente aparición de los Sistemas de Inteligencia Artificial de Propósito General *(General-Purpose Artificial Intelligence Systems, GPAIS)* plantea desafíos de configuración y adaptabilidad del modelo a escalas de complejidad mucho mayores que el diseño de los modelos tradicionales de aprendizaje automático. La Lógica Difusa ha sido una herramienta muy práctica para el diseño y la optimización de modelos de aprendizaje automático, dotándolos de la capacidad de adaptarse a la incertidumbre de la tarea considerada. Por tanto, su aplicación a GPAIS es una elección natural. En esta contribución se analiza la importancia de la Lógica Difusa en el diseño de GPAIS. Se examinan varios casos donde la Lógica Difusa puede ayudar a diseñar estos sistemas u optimizarlos para mejorar su rendimiento en un mundo abierto. También se presentan diferentes estrategias seguidas para este fin, discutiendo áreas tangenciales, contribuciones recientes, identificando nichos y delineando potenciales líneas de investigación. Como caso particular, se discute la integración de la Lógica Difusa con el Aprendizaje de Cero Ejemplos que muestra como la potencia combinada de las fortalezas de ambas impacta directamente en el rendimiento mejorado del sistema resultante. Esta combinación promete revolucionar la Inteligencia Artificial, desplegando un panorama nuevo de nuevas aplicaciones que aprovechan la sinergia entre la Lógica Difusa y los GPAIS.**

Palabras clave — Inteligencia Artificial, Propósito General, Lógica Difusa, Aprendizaje de Cero Ejemplos, Grandes Modelos de Lenguaje, Aprendizaje Automático.

I. Introducción

La principal característica del ser humano que le diferencia del resto de seres vivos es la capacidad de razonar. Dada una serie de premisas, el ser humano es capaz de extraer conclusiones en base al uso de operadores lógicos para deducir la veracidad de afirmaciones que han sido construidas a partir de las anteriores. Podemos considerar esos operadores dentro de la lógica clásica y trabajar con enunciados que son ciertos o falsos [1]. Cuando se desea representar el razonamiento lógico sobre aspectos dinámicos de la vida real, en ocasiones, necesitamos otras lógicas [2]. La Lógica Difusa, ideada por Lofti Zadeh [3], es una extensión de la lógica clásica que gestiona información imprecisa, siendo útil en contextos inciertos. Se basa en conjuntos con grados de pertenencia variables, lo que permite expresar valores de verdad más allá del simple binomio verdadero/falso. Esta lógica es aplicada en Inteligencia Artificial (IA) para crear sistemas que pueden manejar y aprender de la ambigüedad y ha dado lugar a sistemas expertos de tipo difuso, como p. ej. los sistemas de control automático [1].

El Aprendizaje Automático (*Machine Learning*, ML) es un subcampo de la IA enfocado en el desarrollo de modelos capaces de aprender patrones a partir de los datos. La optimización de estos modelos es un área de insistente investigación. La diversidad de objetivos considerados hasta la fecha refleja que las capacidades de optimización de los modelos de ML dependen de diferentes criterios. Avances recientes en diferentes áreas de investigación de ML, como Aprendizaje Profundo [4], Procesamiento del Lenguaje Natural [5] y los Grandes Modelos de Lenguaje (*Large Language Models*, LLMs) [6], incluyendo chatbots de rendimiento sin precedentes como ChatGPT [7] indican un cambio notable hacia sistemas de IA con mayor capacidad de generalización. Estos Sistemas de IA de Propósito General (GPAIS) han ganado especial importancia debido a que son capaces de ejecutar tareas para las que no fueron explícitamente entrenados [8]. Es importante remarcar que potencialmente, estos sistemas pueden presentar ciertas limitaciones como, por ejemplo, el sesgo presente en los datos de entrenamiento, la falta de comprensión profunda real del contenido que producen, las alucinaciones generadas gramaticalmente coherentes, pero factualmente incorrectas, etc. Por consiguiente, siempre que sea factible, su implementación y uso deberían estar supervisados por un humano.

El objetivo de la presente contribución es mostrar el potencial de la integración de la Lógica Difusa con GPAIS para mejorar la capacidad de generalización y manejo de la incertidumbre en sistemas de IA. A través de la revisión de la literatura reciente, ejemplos aplicados y un análisis crítico, se busca proporcionar una perspectiva innovadora sobre cómo la combinación de ambas puede superar las limitaciones actuales y abrir nuevas posibilidades en el campo de la IA. Dicha sinergia promete revolucionar la forma en la que las máquinas comprenden y procesan información, abriendo nuevas posibilidades a la IA del futuro.

A continuación, la Sección 2 realiza un breve recorrido por las diferentes aplicaciones de la Lógica Difusa en IA, la Sección

3 muestra las fundaciones de los GPAIS y la Sección 4, el alineamiento entre ambas. En la Sección 5, se pone foco en un caso de estudio para mostrar potenciales sinergias entre la Lógica Difusa y el Aprendizaje de Cero Ejemplos (*Zero-Shot Learning*, ZSL). Terminamos la contribución exponiendo las principales ventajas de combinar la Lógica difusa y GPAIS.

II. PRELIMINARES

A. *La Lógica Difusa en Machine Learning*

La Lógica Difusa es particularmente útil en situaciones donde la información es ambigua o incierta, como en la toma de decisiones humanas o en la percepción sensorial. La Lógica Difusa puede manejar la imprecisión inherente a muchos problemas del mundo real y proporcionar soluciones más flexibles y adaptativas. Sus principales ventajas son:

1) Maneja la **incertidumbre e imprecisión**, facilita la representación de conocimiento impreciso o vago, ofrece flexibilidad en la toma de decisiones al considerar grados de pertenencia, permite una interpretación más intuitiva y cercana a la forma en que los humanos razonan, es útil en sistemas donde las reglas son difíciles de definir de manera precisa.

2) Modela **sistemas complejos**, puede usar reglas lingüísticas para crear grados de pertenencia, reduce la complejidad de sistemas no lineales, integra con sistemas expertos para mejorar la toma de decisiones.

3) Permite una mayor **robustez** de las aplicaciones frente a ruido y variabilidad de los datos.

Existen dos tipos principales de Lógica Difusa: Tipo I y Tipo II. La lógica difusa de Tipo I asigna grados de pertenencia entre 0 y 1 a elementos en conjuntos difusos con funciones de membresía fijas, facilitando el manejo de incertidumbre en diversas aplicaciones. El Tipo II añade más incertidumbre al permitir funciones de membresía difusas, útiles cuando los datos o reglas son aún más imprecisos. Ambos tipos son valiosos para sistemas que procesan verdades parciales en lugar de absolutas.

La Lógica Difusa ha sido empleada en diferentes casos de uso de ML. Los sistemas basados en reglas difusas son los más ampliamente extendidos, de entre los que destacan Sistemas Difusos Genéticos, Sistemas Difusos Jerárquicos, Sistemas Neuro Difusos y Sistemas Evolutivos Difusos [9]. No obstante, lo lógica difusa se puede usar en otros contextos que no sean reglas. A modo de ejemplo, podemos encontrar modelos difusos para mejorar modelos no supervisados como el k-medias, o PCA, para modelos tradicionales supervisados, como el algoritmo de los vecinos más cercanos, entre muchos otros.

Además del uso de la Lógica Difusa para mejorar el rendimiento de muchos de los algoritmos, se puede destacar su uso para mejorar la explicabilidad de algoritmos de caja negra es ampliamente estudiado [10], lo cual puede llegar a ser aún más relevante en GPAIS.

B. *GPAIS: definición y taxonomía.*

De acuerdo con [8], los GPAIS se definen como "*sistemas avanzados de IA capaces de realizar eficazmente una serie de tareas distintas. Su grado de autonomía y habilidad viene determinado por varias característica clave, como la capacidad de adaptarse o realizar bien nuevas tareas que surjan en el futuro, la demostración de competencia en ámbitos para los que no ha sido intencionada y específicamente entrenado, la capacidad de aprender a partir de datos limitados y el reconocimiento proactivo de sus propias limitaciones con el fin de mejorar su rendimiento, la habilidad para aprender de datos limitados y el reconocimiento proactivo de sus propias limitaciones para mejorar su rendimiento*".

Los GPAIS se convierten en un paso adelante en la optimización del aprendizaje automático y se clasifican en dos categorías diferentes [8], que se ilustran en la Figura 1:

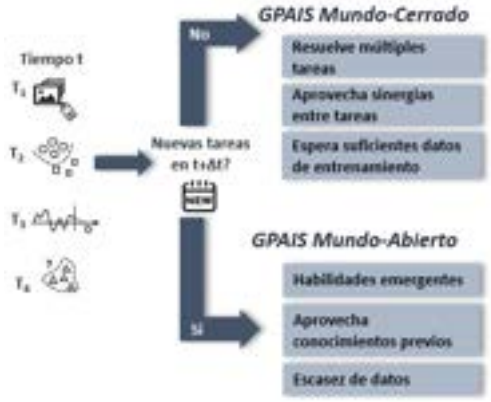


Fig. 1. GPAIS: mundo abierto vs mundo cerrado (adaptado de [8])

- *GPAIS de mundo cerrado*: asume que todas las tareas que realiza el sistema están predeterminadas en la fase de entrenamiento. Es la forma más sencilla en la que el GPAIS es capaz de acometer más de una tarea. El GPAIS reentrena todo su modelo ML cuando surge una nueva tarea. Esto se hace normalmente poniendo foco en las variables de diseño del modelo, tomando la asunción de disponer de suficientes datos para el reentrenamiento. Este tipo de sistemas pueden estar limitados a la hora de generalizar a nuevas tareas que estén fuera de su ámbito de entrenamiento.
- *GPAIS de mundo abierto*: se refiere a escenarios donde el sistema de IA opera en un entorno más dinámico y en evolución. Estos sistemas reconocen la presencia de tareas desconocidas e imprevistas que pueden surgir con el tiempo y se adaptan a ellas mediante el ajuste del modelo o modelos existentes sin asumir que se disponen de suficientes datos. De este modo, el GPAIS se centra en la diversidad de modelos más que en el ajuste de la configuración del modelo, por lo que puede utilizarse en distintos problemas, incluso en ausencia de datos suficientes para las nuevas tareas. Así, estos sistemas tienen un alto grado de generalización, flexibilidad y capacidad de aprendizaje en entorno de incertidumbre en la cantidad de datos. Para dicha generalización, el GPAIS aprovecharía lo que aprendió antes para adaptarse más rápidamente a las nuevas tareas.

Las propiedades intrínsecas de los GPAIS pueden dar lugar a posibles problemas o conflictos durante su diseño o incluso

mientras el sistema está funcionando. Por lo tanto, debemos asegurarnos de que existan mecanismos específicos para superar estas dificultades. Aquí radica la justificación del paradigma IA potenciada por IA (*AI-powered AI*) que pretende paliar la dificultad inherente al diseño, la construcción y la adaptación de los GPAIS para abordar eficazmente nuevas tareas emergentes. La IA potenciada por IA implica aprovechar una técnica de IA para diseñar o mejorar otro modelo de IA [8].

III. GPAIS: TAXONOMÍA DE MÉTODOS

Atendiendo a la taxonomía propuesta en [8] describiremos alguna de las principales tendencias de investigación en que se están desarrollando GPAIS. Es importante tener en cuenta que puede existir cierto grado de solapamiento entre categorías, y que ciertas estrategias podrían facilitar la transición entre categorías, por ejemplo, añadir ciertos elementos a un sistema para evolucionarlo de mundo cerrado a mundo abierto.

Con el objetivo de la IA sea más autosuficiente y capaz de aprender sin intervención humana, la comunidad de IA ha seguido muchas estrategias para proporcionar capacidades de generalización. En términos generales, se observan dos enfoques distintos: IA potenciada por IA y enfoque de modelo único.

- ***IA potenciada por IA*:** Consiste en la aplicación de técnicas de inteligencia artificial para mejorar y optimizar el desarrollo y el rendimiento de otros sistemas de IA [11]. Esto puede incluir, por ejemplo, el uso de algoritmos de aprendizaje automático para seleccionar y optimizar hiperparámetros de modelos de IA, o el uso de técnicas de procesamiento de lenguaje natural para mejorar la interacción humano-máquina.
- ***Modelo único:*** No todos los avances en GPAIS necesitarían siempre utilizar un modelo de IA adicional para ayudar a generalizar. En lugar de ello, sus capacidades de generalización proceden del aprendizaje a partir de diversas tareas y/o grandes cantidades de datos. Se incluirían en esta categorización los modelos fundacionales y el aprendizaje multitarea.

La Figura 2 representa gráficamente la taxonomía multidimensional propuesta en [8] para las diferentes categorías de GPAIS, aproximaciones y áreas de investigación.

1) IA potenciada por IA

Se pueden categorizar las formas en las que una IA puede potenciar otros sistemas de IA, teniendo en cuenta su objetivo: diseñar un algoritmo de IA o enriquecer un algoritmo de IA para mejorar su aprendizaje/rendimiento:

a) IA para diseñar IA

Es habitual que las etapas del ciclo de vida del ML se vuelvan algo repetitivas, por lo que los expertos basan sus decisiones en su experiencia con problemas anteriores. Los algoritmos de IA pueden utilizarse para ayudar con todas esas decisiones de diseño, extrayendo conocimientos generales sobre cómo implementar un proceso de ML de forma más fácil y rápida dentro de un conjunto de tareas de ML. Dentro de la fase de construcción del algoritmo, cabe destacar *AutoML-Zero* que es un enfoque de aprendizaje automático que busca automatizar todo el proceso de desarrollo de modelos de IA, desde la selección de algoritmos y la ingeniería de características hasta la optimización de hiperparámetros y la evaluación de modelos [12]. La idea es democratizar el acceso a la IA, permitiendo que personas sin experiencia en aprendizaje automático puedan crear y desplegar modelos de IA de alta calidad.

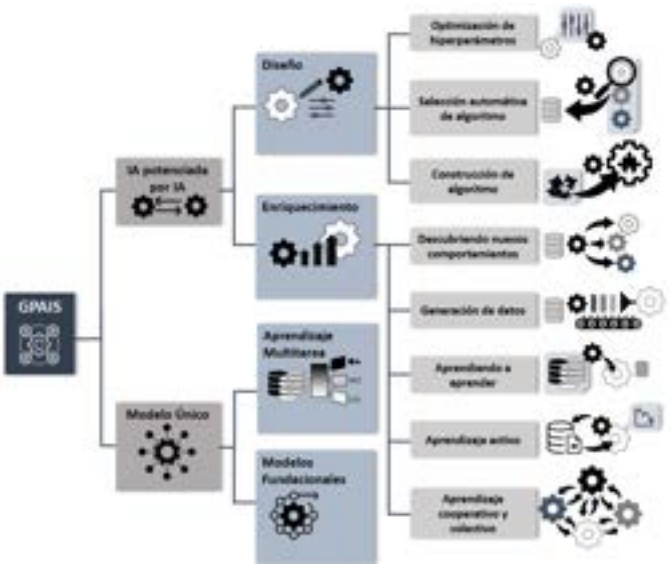


Fig. 2. Taxonomía de aproximaciones a GPAIS (adaptado de [8])

b) IA para enriquecer IA

A menudo se combina la IA con otras técnicas de IA para enriquecer a ayudar en su proceso de aprendizaje. Aprovechando la potencia de varias técnicas de IA y su potencial colaboración, podemos potenciar sistemas de IA para que se conviertan en sistemas inteligentes más sólidos y capaces [13]. En este aspecto, cabe destacar diferentes técnicas utilizadas para el enriquecimiento: descubrimiento de nuevos comportamientos, generación de datos, aprender a aprender, aprendizaje activo, aprendizaje cooperativo y colectivo.

2) Modelo Único

Ejemplos de modelos únicos de IA que generalizan a múltiples tareas son:

a) Aprendizaje Multitarea

Es una técnica de aprendizaje automático en la que un modelo se entrena para realizar múltiples tareas simultáneamente, aprovechando las similitudes y diferencias entre las tareas para mejorar el rendimiento general [8]. El conocimiento y representación se optimizan en tareas múltiples para mejorar el rendimiento, a diferencia del aprendizaje automático tradicional que se enfoca en tareas únicas. Los modelos de aprendizaje multitarea, que suelen ser GPAIS de mundo cerrado, requieren reentrenamiento o ajustes para nuevas tareas. Estrategias como el preentrenamiento y ajuste fino facilitan la adaptación a entornos de mundo abierto, necesitando datos suficientes para la nueva tarea. Aproximaciones como metaheurística, aprendizaje por refuerzo y meta-transformer han tenido mucho éxito a la hora de reducir el riesgo de error y la baja eficiencia, cuando se comparan estos sistemas con modelos entrenados para tareas individuales.

b) Modelos Fundacionales

Los modelos fundacionales se basan en el aprendizaje profundo estándar y en el aprendizaje por transferencia, y se definen como un modelo que *"se entrena con gran cantidad de datos (generalmente usando autosupervisión a escala) que pueden ser adaptados (p.ej. mediante ajuste fino) a una amplia gama de tareas posteriores"* [14]. Se argumenta en [8] que los GPAIS van más allá del aprendizaje por transferencia y el aprendizaje profundo, y que, por lo tanto, los modelos fundacionales se convierten en un subconjunto de los GPAIS.

A modo de ejemplo, se entrenaría un modelo con grandes conjuntos de datos y, posteriormente, podría ser ajustado para realizar tareas específicas, como clasificación de imágenes, reconocimiento de voz y procesamiento de lenguaje natural, y por tanto son inherentemente multitarea. Son modelos que se han vuelto muy populares en los LLMs mostrando grandes capacidades de propósito general. Dentro de esta categoría, cabe destacar la IA Generativa que es un tipo de IA que puede crear ideas y contenidos nuevos, como conversaciones, historias, imágenes, videos y música. A diferencia de otros enfoques de IA que se centran en el análisis y la predicción, la IA generativa se enfoca en la creación de nuevos datos a partir de modelos aprendidos. Los modelos fundacionales pertenecen inherentemente a la clasificación de mundo abierto, pero con algunas limitaciones [8], dado que requiere de grandes cantidades de datos de calidad durante la fase de entrenamiento. Si no se dispone de esa cantidad de datos, para que puedan acometer nuevas tareas, deben hibridarse con otros modelos de IA, como por ejemplo el aprendizaje activo.

IV. Integración de la Lógica Difusa y GPAIS: una visión general

La integración de la Lógica Difusa puede ofrecer varias ventajas. En esta sección, sugerimos algunas posibilidades. Primero, la Lógica Difusa puede proporcionar un marco para manejar la incertidumbre en la información auxiliar utilizada en GPAIS [12]. Por ejemplo, las descripciones textuales de las clases pueden ser vagas o ambiguas, y la Lógica Difusa puede ayudar a modelar estas incertidumbres de manera efectiva.

En segundo lugar, en vez de considerar los atributos como presentes o ausentes (binarios), la Lógica Difusa permite grados de pertenencia, lo que puede reflejar mejor la variabilidad y la riqueza de las características del mundo real. En [15] se muestran diferentes aplicaciones reales de sistemas que combinan la Lógica Difusa con aprendizaje evolutivo y por refuerzo. Teniendo en cuenta las fortalezas de la Lógica Difusa, integrarla en modelos de GPAIS permitiría que realizaran predicciones más precisas, mejorando su interpretación.

Tercero, la combinación de ambas técnicas puede facilitar la creación de sistemas de IA más robustos y generalizables. La Lógica Difusa puede suavizar las decisiones de clasificación en GPAIS, lo que podría ser especialmente útil en escenarios de Aprendizaje Generalizado de Cero Ejemplos (*Generalized Zero-Shot Learning*, GZSL) [16] donde el modelo debe funcionar bien tanto en clases vistas como no vistas.

Finalmente, la Lógica Difusa puede ayudar a mejorar la interpretación y explicabilidad de los modelos de GPAIS. Al proporcionar un marco para razonar sobre la incertidumbre y la ambigüedad, los sistemas que integran Lógica Difusa pueden ofrecer explicaciones más intuitivas y comprensibles para las decisiones tomadas por el modelo.

Integrando la Lógica Difusa dentro de GPAIS, se consigue manejar la incertidumbre y la imprecisión de los datos, lo que permite realizar predicciones y tomar decisiones más precisas en los sistemas de IA. La Tabla 1 expone puntos en común entre la Lógica Difusa y algunas de las principales técnicas de GPAIS.

TABLA I – SINERGIAS ENTRE LA LÓGICA DIFUSA Y GPAIS

Técnica / Ejemplo GPAIS	**Puntos en común con la Lógica Difusa**
IA potenciada por IA	Puede utilizarse para detectar cambios en el entorno y adaptarse a ellos, transferir conocimientos de forma eficaz de un conjunto de tareas a una nueva tarea y transferir conocimientos o representaciones específicas de una tarea de origen para mejorar el rendimiento en una tarea de destino.
Modelos Fundacionales	Puede utilizarse para transferir conocimientos de forma eficaz de un conjunto de tareas a una nueva tarea, mejorando su capacidad de adaptación a nuevas tareas.
IA Generativa	Puede utilizarse para que un modelo se entrene simultáneamente en diferentes tareas, mejorando la calidad y la diversidad de las muestras generadas.
Aprendizaje Multitarea	Puede utilizarse para transferir conocimientos o representaciones específicas de una tarea de origen para mejorar el rendimiento en una tarea de destino.
AutoMl-Zero	Puede utilizarse para detectar cambios en el entorno y adaptarse a ellos, transferir conocimientos de forma eficaz de un conjunto de tareas a una nueva tarea y transferir conocimientos o representaciones específicas de una tarea de origen para mejorar el rendimiento en una tarea de destino.

La Tabla II muestra posibles aportaciones, basadas en Lógica Difusa, a las limitaciones actuales de los sistemas GPAIS.

TABLA II – POSIBLES APORTACIONES DE LA LÓGICA DIFUSA

Limitación	**Potencial Aportación de la Lógica Difusa**
Dependencia de grandes conjuntos de datos.	Puede ayudar a resolver incertidumbre cuando existen pocos datos de los que aprender.
Sesgos en los datos de partida/alucinaciones.	Reduciendo sesgos usando métricas difusas para cuantificar sesgos.
Dificultad para manejar la incertidumbre y la imprecisión.	Mejorando el rendimiento de los algoritmos de IA al proporcionar una representación más precisa del dominio del problema.
Capacidad limitada para adaptarse a nuevas tareas.	Mejorando la adaptabilidad de los modelos de fundamento único a nuevas tareas.
Dificultad para gestionar tareas complejas.	Manejando la complejidad de las tareas proporcionando una representación más precisa del dominio del problema.
Dificultad para generar muestras de alta calidad.	Mejorando la calidad de las muestras generadas al proporcionar una representación más precisa del dominio del problema.
Control limitado sobre las muestras generadas.	Proporcionando un mayor control sobre las muestras generadas al permitir una especificación más precisa del resultado deseado.

Dificultad para gestionar tareas con diferentes niveles de complejidad.	Mejorando la capacidad del aprendizaje multitarea para gestionar tareas con diferentes niveles de complejidad al proporcionar una representación más precisa del dominio del problema.
Capacidad limitada para transferir conocimientos entre tareas.	Mejorando la capacidad del aprendizaje multitarea para transferir conocimientos entre tareas al permitir una especificación más precisa de las relaciones entre tareas.
Dificultad para gestionar espacios de búsqueda complejos.	Mejorando la capacidad de AutoML-cero para manejar espacios de búsqueda complejos al proporcionar una representación más precisa del dominio del problema.
Capacidad limitada para incorporar el conocimiento del dominio.	Mejorando la capacidad de AutoML-cero para incorporar el conocimiento del dominio al permitir una especificación más precisa del dominio del problema.

En resumen, la integración de la Lógica Difusa con GPAIS tiene el potencial de crear sistemas de IA que no solo son capaces de reconocer clases no vistas, sino que también manejan la incertidumbre de manera más natural y proporcionan soluciones más flexibles y comprensibles. Esto representa un paso adelante en el desarrollo de sistemas de IA que pueden operar de manera efectiva en entornos del mundo real, donde la ambigüedad y la falta de datos completos son comunes.

V. LÓGICA DIFUSA PARA ZERO-SHOT LEARNING

Dentro del mundo abierto en GPAIS, hay diversos paradigmas que intentan lidiar con situaciones desconocidas. Entre ellos destaca lo que se conoce como Reconocimiento de Conjuntos Abiertos, (*Open Set Recognition*, OSR). El OSR es una metodología dentro del ML que se encarga de lidiar con situaciones desconocidas, las cuales no fueron aprendidas por los modelos durante el proceso de entrenamiento. Se espera que los modelos de OSR no solamente sean capaces de rendir correctamente bajo situaciones conocidas, sino que también sean capaces de manejar otros sucesos con total incertidumbre [17].

Hasta ahora, los modelos de OSR operan discerniendo qué situaciones son conocidas previamente, gracias a su entrenamiento, y cuales no, marcando estas últimas con una opción de rechazo sin intentar profundizar más [17]. Sin embargo, el ZSL y su variante más general, Aprendizaje de Cero Ejemplos Generalizado, son otros paradigmas que se encargan de enfrentar con más detenimiento estas situaciones inexploradas, bajo la asistencia de información semántica auxiliar de las situaciones que ya se conocen y las que no. Mientras que el ZSL, en la fase de prueba, trabaja únicamente en un entorno con el que no ha sido entrenado, el GZSL opera en un entorno más general al encontrarse con situaciones con las que fue entrenado y otras con las que no.

La información semántica juega un papel fundamental ya que, pese a que haya situaciones desconocidas para el modelo, brinda una descripción de éstas permitiendo establecer un nexo de unión entre la información que ya el modelo conoce y la que no, en aras de establecer una transferencia del conocimiento de las clases de un tipo a otro poder tratarlas posteriormente.

A continuación, destacamos algunos ejemplos comentando potenciales aplicaciones de la Lógica Difusa en ZSL:

- **Complemento en la Incertidumbre**: En ZSL existe una incertidumbre inherente debido a la existencia de clases no vistas, la cual se transfiere al espacio semántico por ser el elemento sobre el que se apoya la clasificación. La Lógica Difusa puede ayudar a manejar esta incertidumbre tanto en dichos espacios como en los algoritmos que operan bajos éstos, por ejemplo, el vecino más cercano.
- **Explicabilidad**: La explicabilidad en los modelos de ZSL no está muy desarrollada, sobre todo de cara a los modelos que emplean técnicas generativas. Por eso la interpretación intuitiva de la Lógica Difusa puede ser beneficiosa para explicar mejor los resultados.
- **Generalización y Robustez**: La robustez y la generalización en ZSL son unos aspectos críticos dado que los modelos son entrenados con una cantidad limitada de información, siendo muy sensibles al resto de información que no han visto. De hecho, la mayoría de los modelos de ZSL no son capaces de generalizar bien. Además, los modelos de GZSL presentan cierto sesgo hacia las clases vistas. La Lógica Difusa puede mitigar dicha sensibilidad haciendo estos modelos más robustos y con una mayor capacidad de generalización frente a las clases no vistas.

En literatura actual, se están empezando a ver algunos esfuerzos para combinar estas dos técnicas en aplicaciones prácticas. Destacamos los siguientes ejemplos:

- La **detección de objetos** destacados en imágenes que contienen información de color y profundidad sobre las escenas es fundamental para la automatización que conduce a aplicaciones como la conducción autónoma de vehículos y la manipulación robótica. Se integran ambas técnicas mediante un enfoque de fusión de datos de color y profundidad para la detección de objetos destacados que aprovecha el aprendizaje profundo y una novedosa estrategia de entrenamiento basada en la teoría de conjuntos difusos que elimina la necesidad de ejemplos para el entrenamiento [18]. Este ejemplo de integración de ambas técnicas resuelve, entre otras, en la Tabla II, la dependencia de grandes cantidades de datos, la dificultad para generar muestras de calidad.
- El **diagnóstico de fallos de cero ejemplos** ayuda a identificar fallos no vistos mediante la predicción de atributos. La integración realizada comprende un método de aprendizaje jerárquico difuso de cero ejemplos para resolver estos problemas. En primer lugar, los atributos se dividen en diferentes capas según la granularidad de grueso a fino a través del conocimiento experto, en lugar de tratarse por igual. A continuación, se diseña una estrategia de transferencia de conocimiento para transferir el conocimiento de los atributos de grano grueso a los de grano fino, lo que puede mejorar la precisión de la predicción de atributos. Por último, se desarrolla una estrategia de inferencia difusa para distinguir el efecto de atributos con distinta granularidad en la inferencia de fallos. Esta estrategia puede identificar los fallos paso a paso en un

orden de grano grueso a fino. La eficacia del método propuesto se verifica mediante el proceso de una central térmica real [19]. Este ejemplo resuelve la incertidumbre e imprecisión en los datos y la transferencia de conocimiento indicados en la Tabla II.

- Replanteamiento de **clasificadores**, reemplazando los atributos binarios muy utilizados en ZSL, por otros que implementen la Lógica Difusa, denominados atributos difusos. De esta forma se eliminan las limitaciones en el aprendizaje y mejorando notablemente el rendimiento del clasificador. En [20] se propone un algoritmo "*Fuzzy Zero-Shot Learning*" y el uso de atributos difusos para expresar las características, especialmente las de la naturaleza. Un atributo difuso usa un numero para expresar el grado del atributo que tiene cada objeto. El uso de atributos difuso impacta directamente en el mejor rendimiento del modelo.

La integración entre la Lógica Difusa y ZSL es un campo prometedor con aplicaciones prácticas que se benefician de la capacidad de manejar la incertidumbre y la ambigüedad, características inherentes al razonamiento humano y a muchos procesos del mundo real. La investigación futura en esta dirección no solo profundizará en el entendimiento teórico de la integración de estas técnicas, sino que también explorará nuevas aplicaciones prácticas en la resolución de problemas complejos.

Conclusiones

Las conclusiones extraídas tras llevar a cabo este estudio confirman el potencial transformador que supone combinar técnicas avanzadas como la Lógica Difusa con GPAIS en el campo de la inteligencia artificial. La Lógica difusa permite manejar la incertidumbre y la imprecisión inherentes a muchos problemas del mundo real, lo que permite un razonamiento más flexible y cercano al pensamiento humano. Por otro lado, los GPAIS permiten reconocer clases o conceptos no vistos durante el entrenamiento, lo que permite generalizar el conocimiento adquirido a nueva categorías o conceptos, favoreciendo la versatilidad y escalabilidad de los sistemas. Aprovechando las fortalezas de ambos enfoques y superando sus limitaciones, esta fusión estratégica no solo muestra cómo mejorar significativamente la capacidad predictiva y generalización de los modelos, sino que también allana el camino hacia sistemas más adaptables e inteligentes capaces de abordar desafíos complejos con mayor eficiencia y precisión. A modo de resumen, las principales ventajas son:

- Manejo de incertidumbre y generalización en los datos utilizados por los modelos GPAIS, mejorando su generalización y adaptación a nuevas categorías.
- Interpretabilidad y transferencia de conocimiento, debido a que la Lógica Difusa proporciona una mejor interpretación de los conceptos y relaciones aprendidos por GPAIS facilitando la transferencia de conocimiento entre dominios.
- Robustez y flexibilidad, al favorecer la creación de sistemas capaces de adaptarse a cambios en el entorno, nuevas situaciones y conceptos con mayor eficacia.

Agradecimientos

Cabe destacar que esta iniciativa se realiza en el marco de los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation).

Referencias

[1] D. L. Palacios, "Introducción a la lógica difusa y sus aplicaciones.", Asociación Nacional de Estudiantes de Matemáticas, 2017.

[2] C. G. Morcillo, "Lógica difusa, una introducción práctica," Técnicas de softcomputing, vol. 29, 2011.

[3] L. Zadeh, "Fuzzy sets," Information and Control, 1965.

[4] F. Wang and M. Shen, "Automatic kernel generation for large language models on deep learning accelerators," in 2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD), 2023.

[5] J. Just, "Natural language processing for innovation search – reviewing an emerging non-human innovation intermediary," Technovation, 2024.

[6] M. A. K. Raiaan, M. S. H. Mukta, K. Fatema, N. M. Fahad, S. Sakib, M. M. J. Mim, J. Ahmad, M. E. Ali, and S. Azam, "A review on large language models: Architectures, applications, taxonomies, open issues and challenges," IEEE Access, vol. 12, 2024.

[7] C. Gao, F. Howard, N. Markov, E. Dyer, S. Ramesh, Y. Luo, and A. Pearson, "Comparing scientific abstracts generated by chatgpt to real abstracts with detectors and blinded human reviewers," npj Digital Medicine, 2023.

[8] I. Triguero, D. Molina, J. Poyatos, J. Del Ser, and F. Herrera, "General purpose artificial intelligence systems (GPAIS): Properties, definition, taxonomy, societal implications and responsible governance," Information Fusion, vol. 103, 2024.

[9] A. K. Varshney and V. Torra, "Literature review of the recent trends and applications in various fuzzy rule-based systems," International Journal of Fuzzy Systems, vol. 25, no. 6, 2023.

[10] M. Yeganejou, K. Honari, R. Kluzinski, S. Dick, M. Lipsett, and J. Miller, "Dcnfis: Deep convolutional neuro-fuzzy inference system," ArXiv, 2023.

[11] H. Song, I. Triguero, and E. Ö zcan, "A review on the self and dual interactions between machine learning and optimisation," Progress in Artificial Intelligence, vol. 8, 04 2019.

[12] M. Tsiakmaki, G. Kostopoulos, S. Kotsiantis, and O. Ragos, "Fuzzy-based active learning for predicting student academic performance using automl: a step-wise approach," Journal of Computing in Higher Education, vol. 33, 12 2021.

[13] W. Lin, "Deep reinforcement learning based haptic enhancement for tele-diagnosis," in 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE Press, 2022.

[14] Rishi Bommasani, et al. On the opportunities and risks of foundation models, 2022. arXiv:2108.07258

[15] Gu, X., Han, J., Shen, Q. et al. Autonomous learning for fuzzy systems: a review. Artificial Intelligence Review 56, 7549–7595 (2023)

[16] K. Huang, S. Mckeever, and L. Miralles-Pechuán, "Generalized zero-shot learning for action recognition fusing text and image gans," IEEE Access, vol. 12, 2024.

[17] C. Geng, S.-J. Huang, and S. Chen, "Recent advances in open set recognition: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, 2018.

[18] S. Bhuyan, A. Kar, D. Sen, and S. Deb, "Rgb-d fusion through zero-shot fuzzy membership learning for salient object detection," IEEE Transactions on Artificial Intelligence, 2024.

[19] X. Chen, B. Zhang, C. Zhao, J. Ding, and W. Wang, "From coarse to fine: Hierarchical zero-shot fault diagnosis with multi-grained attributes," IEEE Transactions on Fuzzy Systems, 2024.

[20] C. Liu, Z. Shang, and Y. Y. Tang, "Zero-shot learning with fuzzy attribute," in 2017 3rd IEEE International Conference on Cybernetics (CYBCONF), 2017.

Análilsis de los Cambios en los Patrones de Temperatura mediante Técnicas de *Stream Clustering*

1st Asier Urio-Larrea
Universidad Pública de Navarra
Pamplona, Spain
asier.urio@unavarra.es

2nd Graçaliz Pereira Dimuro
Universidade Federal do Rio Grande do Sur
Rio grande do sur, Brasil
gracalizdimuro@furg.br

3rd Javier Andreu-Pérez
University of Essex
Coclchester, United Kingdom
j.adreu-perez@essex.ac.uk

4th Heloisa Camargo
Universidade Federal do Sao Carlos
Sao Carlos, Brasil
heloisa@dc.ufscar.br

5th Javier Aguirre
Universidad Pública de Navarra
Pamplona, Spain
jaguirreeraso@gmail.com

6th Humberto Bustince
Universidad Pública de Navarra
Pamplona, Spain
bustince@unavarra.es

***Abstract*—El cambio climático afecta a las condiciones medioambientales de las distintas regiones. La capacidad de constatar estos cambios es una eficaz herramienta para adaptarse a la evolución de las condiciones. Los datos meteorológicos se generan continuamente en múltiples estaciones de todo el mundo, proporcionando una valiosa información sobre la variabilidad en el tiempo de los patrones climáticos. El estudio de este flujo de datos nos permite comprender mejor los nuevos patrones climáticos. Este trabajo explora, mediante un algoritmo de agrupamiento de flujos de datos (*stream clustering*), el potencial de emplear datos meteorológicos obtenidos en diferentes localizaciones geográficas para rastrear el cambio en los patrones climáticos en la Comunidad Foral de Navarra durante los últimos 20 años. El estudio de caso mostró la aplicabilidad de los métodos de flujos de datos a la segmentación incremental de regiones geográficas en función de sus factores climatológicos.**

***Index Terms*—data stream, clustering, clima**

I. Introduction

El cambio climático hace referencia al aumento de la temperatura global de la Tierra, que se está produciendo con mayor rapidez que antes. Este cambio está causado principalmente por las actividades humanas, que producen grandes cantidades de CO_2 [1]. El cambio climático tiene un impacto directo en el medio ambiente: el aumento de la desertificación, más olas de calor, incendios forestales y sequías son sólo algunas de las repercusiones del cambio climático en el medio ambiente [2], [3]. Comprender estos cambios es una valiosa herramienta para hacer frente a las nuevas condiciones meteorológicas.

El clustering ha sido empleado por diferentes autores para trabajar con problemas climáticos y meteorológicos. Por ejemplo, En [4], los autores emplean un método de clustering de consenso para identificar las diferentes regiones climáticas de Estados Unidos. Calculan los clusters de temperatura y precipitación de forma independiente y, a continuación, crean los agrupamientos de consenso finales basándose en ellos. En [5], emplean un método de agrupación en dos etapas con el objetivo de determinar regiones homogéneas de temperatura en la India. En [6], desarrollan un algoritmo de co-clustering para analizar patrones espacio-temporales, concretamente en relación con la temperatura de las estaciones meteorológicas holandesas. En [7], examinan los cambios en la temperatura diaria del aire de Europa Central mediante regresión cuantílica y clustering jerárquico. Esta técnica permite estimar tendencias en la distribución de los datos y patrones espaciales. En [8], combinan el método de la función ortogonal empírica rotada y la agrupación k-means para identificar regiones homogéneas de temperaturas máximas diarias en China. En [23], proponen un método de clustering jerárquico basado en una medida de similitud para datos intervalo valorados, y lo emplean para agrupar regiones españolas en función de sus temperaturas.

Los datos climáticos son producidos constantemente por estaciones meteorológicas con una gran diversidad de frecuencias. Estas estaciones registran continuamente diferentes parámetros meteorológicos. Con el objetivo de identificar los cambios en las diferentes zonas climáticas de Navarra, en este estudio proponemos el uso de *stream clustering*. El *stream clustering* no se limita a agrupar el conjunto de datos disponibles en un momento concreto, sino que también permite tratar con un flujo de datos en constante evolución. Estos algoritmos están diseñados para tratar conceptos que evolucionan, es decir, que la distribución de los datos cambia con el tiempo. Esta capacidad nos permitirá estudiar la evolución de las regiones climáticas a partir de una serie de datos meteorológicos.

El objetivo principal de este trabajo es estudiar los cambios en los patrones de temperatura en Navarra en los últimos 20 años con respecto a la temperatura media diaria en la estación de verano. Para ello, utilizamos un método de *stream clustering* recientemente publicado llamado TFS-DBSCAN, que es una versión difusa y adaptada a flujos de datos temporales del

Fig. 1: Ubicación de las estaciones meterologicas selecionadas en la región de Navarra.
Imagen modificada de https://commons.wikimedia.org/wiki/File:Spain_Navarre_relief_location_map.svg CC BY-SA Mguillen

algoritmo DBSCAN, un método de clustering de densidad que tiene una baja sensibilidad a las influencias de valores atípicos.

Este artículo está organizado como sigue: En la sección II, presentamos los datos meteorológicos que se utilizarán en el estudio y una descripción del algoritmo seleccionado junto con el marco específico para el experimento. La sección III muestra los resultados obtenidos y una discusión de los mismos. Para finalizar, en la Sección IV, destacamos las principales conclusiones de este trabajo y proponemos algunas líneas futuras de investigación.

II. Datos y Metodología

A. Datos

La provincia de Navarra está situada en el norte de España; su superficie es de 10.391 km^2. Con respecto al clima, en el norte de la región, que tiene una geografía montañosa y está cerca del Golfo de Vizcaya, presenta un clima oceánico. En el centro, el clima es mediterráneo, y en el sur, es fresco y semiárido.

Las series temporales de la temperatura media diaria se obtienen de la agencia meteorológica pública de Navarra (http://meteo.navarra.es). Entre las 153 estaciones meteorológicas actuales, sólo se seleccionaron para este estudio 56 estaciones que tenían algún dato de temperatura media entre 2004 y 2023. Estas estaciones seleccionadas se pueden ver en la Figura 1. Teniendo en cuenta que en esta región las temperaturas presentan una gran variabilidad entre las distintas estaciones, hemos seleccionado únicamente información relativa a la estación estival (julio y agosto).

Estos datos, compuestos por las coordenadas de longitud y latitud y la temperatura media, se normalizan, y cada uno de los 70.203 puntos se introduce cronológicamente en el sistema.

B. Stream Clustering

Stream clustering es un conjunto de algoritmos de agrupamiento que tratan con un flujo continuo de datos en lugar de un conjunto de puntos estáticos.

Según Zubarovglu y Atalay [9], hay diferentes familias de algoritmos, como los jerárquicos (ODAC [10], E-Stream [11]) basados en particiones (CluStream [12], incremental k-means [13]), basados en rejillas (GCHDS [14], GSCDS [15]), basados en la densidad (DBSCAN incremental [16], D-Stream [17]), y basados en modelos (CluDiStream [18], SWEM [19]). Se han desarrollado varias adaptaciones difusas basadas en estos algoritmos. Estas modificaciones permiten una mayor flexibilidad en los límites de los clústeres y las distancias entre los puntos. Por ejemplo, d-FuzzStream [20] se basa en el algoritmo CluStream y TSF-DBSCAN [22] en DBSCAN. Otra clasificación de los algoritmos de aprendizaje automático está relacionada con la disponibilidad de etiquetas de salida conocidas para las entradas; si la etiqueta existe, lo denominamos supervisado, mientras que si no hay etiqueta de salida en el conjunto de datos de entrenamiento, se trata de un método de aprendizaje no supervisado.

Entre los distintos algoritmos de *stream clustering*, seleccionamos el algoritmo TSF-DBSCAN, desarrollado por Bechini et. al. en 2022. Este algoritmo basado en la densidad determina las diferentes zonas climáticas sin necesidad de el conocimiento previo del número de clusters. Y su naturaleza difusa también permite clusters con bordes difusos y solapados. El algoritmo es no supervisado, ya que no sabemos de antemano qué clusters corresponden a los datos de entrada; de hecho, esa es la información que buscamos en este caso de aplicación.

El algoritmo adopta el paradigma online-offline y un modelo *damped window*, como se muestra en la Figura 2. Este paradigma online-offline divide el procesamiento del flujo de datos en dos etapas. La primera de ellas recibe cada uno de los elementos y construye una estructura con ellos. Después, la etapa offline genera los clusters cuando se le solicita, basándose en la estructura generada en la etapa online. El modelo de *damped window* asocia a cada objeto un peso que disminuye con el tiempo; los objetos cuyo peso es inferior a un umbral dado dejan de tenerse en cuenta, por lo que pueden eliminarse de la memoria.

En este algoritmo, las dos etapas tratan el flujo de datos de la siguiente manera:

- Etapa en línea: En esta etapa, el algoritmo mantiene una lista de los puntos que está considerando. Con la llegada de cada nuevo punto de datos, éste se añade a la lista. Cada uno de esos puntos tiene dos listas:
  - Kernel: Que almacena los puntos de su vecindad que están más cerca que σ_{min}
  - Shell: Que almacena los puntos de su vecindad que están entre σ_{min} y σ_{max}.

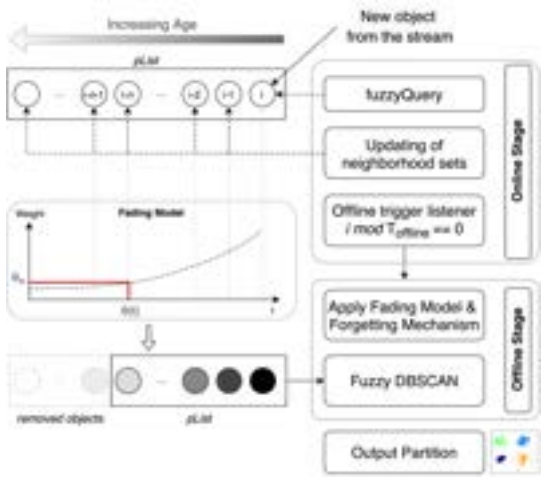


Fig. 2: Pasos principales de TSF-DBSCAN [22]. La estructura de datos *plist* almacena los objetos más recientes. Cada objeto p está asociado a dos conjuntos, con objetos en su vecindad kernel y en su vecindad shell, respectivamente.

Para cada uno de los puntos añadidos a Kernel (o Shell), su Kernel (o Shell) se actualiza con el nuevo punto.

- Etapa offline: Esta etapa se ejecuta cada $T_{offline}$ puntos de entrada y genera las particiones de datos basándose en la información recogida en la etapa online. Sin embargo, para hacer frente a la deriva de conceptos (*concept drift*) en los datos recopilados en línea, se emplea el modelo *damped window* para disminuir la importancia de los datos más antiguos con respecto a los puntos más nuevos. Al comienzo de la etapa fuera de línea, se aplica el factor de decadencia, α, a cada uno de los puntos; esto disminuye el peso de los más antiguos y los elimina si su peso está por debajo del umbral especificado, θ_w. Para determinar los conglomerados basados en la densidad, un punto se identifica como kernel si la suma de los pesos de los puntos pertenencientes a su kernel es mayor que el peso mínimo. W_{min}

C. Marco experimental

Para realizar el análisis del comportamiento de las regiones climáticas, seleccionamos los siguientes atributos para el flujo de datos:

- Temperatura media diaria registrada en la estación meteorológica.
- Latitud de la estación meteorológica.
- Longitud de la estación meteorológica.

Hemos seleccionado únicamente los datos de los meses de julio y agosto entre los años 2004 y 2023, lo que supone un total de 70.204 puntos de datos. No todas las estaciones meteorológicas están activas a lo largo del periodo de tiempo del estudio, y otras empiezan a funcionar después de la fecha inicial de los datos recogidos. Los datos se normalizan antes de introducirlos en el algoritmo. La fase offline del algoritmo se calcula cada 7.000 puntos de datos, lo que equivale a unos dos años de datos.

La tabla I muestra los parámetros seleccionados para el algoritmo TSF-DBSCAN.

Parámetro	Simbolo	Valor
Periodo de evaluación offline	$T_{offline}$	7000
Radio de los vecinos del kernel	ε_{min}	0.10
Radio de los vecinos de shell	ε_{max}	0.35
Umbral de peso	θ_w	0.0015
Factor de decadencia	α	0.3
Miníma suma de pesos	W_{min}	1

TABLE I: Parámetros del algorítmo TSF-DBSCAN

El código completo de este experimento puede consultarse en: https://github.com/asieriko/ClimateStreamNavarra

III. Resultados

Los clusters producidos por el algoritmo están formados por los datos de flujo procedentes de las estaciones meteorológicas. Para una mejor visualización, hemos coloreado la superficie de Navarra, donde cada punto corresponde al color correspondiente al cluster de las estaciones meteorológicas más cercanas. Los pequeños pentágonos que aparecen en las imágenes representan la ubicación de las estaciones meteorológicas.

Antes de comenzar el análisis de los resultados, es necesario tener en cuenta que tal y como se observa en la Figura 1, debido a las condiciones geográficas, principalmente montañas y variaciones de altitud, algunas regiones del norte pueden aparecer divididas aunque compartan las mismas condiciones climáticas.

Para analizar los cambios en las regiones climaticas realizamos el paso offline del algoritmo 10 veces a lo largo de la evolución del flujo de datos, en concreto para los años 2004, 2006, 2008, 2010, 2013, 2915, 2017, 2019, 2021 y 2023. Mostramos los resultados que destacan los cambios de agrupación más relevantes en las siguientes figuras.

Inicialmente, tras la primera evaluación en el año 2004, el algoritmo detecta 12 áreas diferentes, entre las cuales destaca una región con una gran superficie que abarca desde el centro de la región hasta el sur (Figura 3).

No hay cambios en el resultado hasta el año 2008, en el que podemos ver cómo un grupo de regiones del oeste se fusionan en una más grande (Figura 4). También se puede obsercar como una pequeña región al sur de esta que estaba incluida en la gran región central se diferencia de ambas en este momento de evaluación.

Si bien no se ha mostrado en las figuras por ser un cambio menor, en la evaluación correspondiente al año 2010 la pequeña región situada al sur del área recientemente expandida y que en 2004 pertenecia a la gran región central se une al área occidental.

A continuación, en 2013, como puede verse en la Figura 5, esta gran región occidental combinada anteriormente se fusiona con la gran región inicial del sur. Esta nueva zona ocupa la mitad de la región de estudio.

Fig. 3: Resultados de agrupamiento para datos anteriores a 2004

Fig. 5: Resultados de agrupamiento para datos anteriores a 2013

Fig. 4: Resultados de agrupamiento para datos anteriores a 2010

Fig. 6: Resultados de agrupamiento para datos anteriores a 2023

A medida que avanza el tiempo, no se producen grandes cambios y se consolida la gran región formada anteriormente. Sin embargo, en 2015 y 2021 comienzan a funcionar nuevas estaciones meteorológicas en el norte de Navarra. En la Figura 6, se puede observar que dos regiones del norte se dividen en 4.

Resumiendo los resultados obtenidos, hemos visto como las zonas climáticas de la región tienden a homogeneizarse. En la Tabla II, vemos que hay una tendencia a reducir el número de clusters hasta 2013. A partir de ese año, podríamos interpretar erróneamente la tabla debido a que el aumento del número de clusters en los años siguientes fue causado por las nuevas estaciones meteorológicas introducidas en la zona norte de

Año	2004	2006	2008	2010	2013
Clusters	12	12	11	10	9
Año	2015	2017	2019	2021	2023
Clusters	10*	10*	10*	11**	12**

TABLE II: Número de clusters identificados por el algoritmo

Navarra y no por cambios en los patrones climáticos. Esta disponibilidad adicional de datos de arroyos permite dividir dos clusters que antes eran más grandes debido a la falta de datos.

IV. Conclusiones y trabajo futuro

En este trabajo nos hemos propuesto emplear el clustering de flujos de datos como herramienta para analizar la evolución de la temperatura a través de las estaciones meteorológicas de Navarra. En particular, se ha seleccionado para la tarea un algoritmo difuso basado en la densidad, denominado TSF-DBSCAN.

Con nuestra metodología, hemos demostrado cómo han convergido las diferentes zonas climáticas de Navarra. En concreto, las regiones sur y este, que se detectaron como diferentes en cuanto a las temperaturas estivales a principios de la década de 2000, se han fusionado en una más grande en la década de 2010. El resto de las zonas han tenido un comportamiento más consistente, que podría estar relacionado con la geografía de sus respectivos entornos. Este estudio muestra la utilidad de los algoritmos de *stream clustering* para la información geolocalizada y cómo la relación entre los distintos puntos de localización se agrupa a lo largo del tiempo. Esta investigación sirve de base para el estudio del impacto del cambio climático mediante algoritmos de clustering de flujos de datos.

Para este trabajo, sólo hemos estudiado la temperatura media diaria de los meses de verano de julio y agosto. Sin embargo, pueden estudiarse otros parámetros, como las temperaturas máximas y mínimas y el nivel de precipitaciones. Para un estudio más exhaustivo, habría que tener en cuenta más datos. La medida de distancia es otro aspecto importante del algoritmo, y puede modificarse para que pondere los atributos de coordenadas de forma diferente a los climáticos. En futuros trabajos, estudiaremos cómo afecta a los resultados el hecho de utilizar distintos métodos de *stream clustering*, así como las posibles fuentes de desacuerdo entre los diversos métodos posibles. Asimismo, otro trabajo futuro es la evaluación de los resultados integrando factores relacionados con la localización, como la comprobación de efecto de los factores geográficos en los agrupamientos.

Agradecimientos

El trabajo está parcialmente financiado por del proyecto PID2022-136627NB-I00 de la Agencia Estatal de Investigación financiado por MCIN/AEI/10.13039/501100011033/FEDER,UE. Graçaliz Pereira agradece a la Agencia de Financiación Brasileña CNPq (Proc. 407206/2023-0, 304118/2023-0) (CNDCT Brasil). Asier Urio-Larrea es beneficiario de los Contratos predoctorales Santander-UPNA 2021.

References

[1] Lynas, M, Houlton, B.Z. and Perry, S., Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature, Environmental Research Letters, Issue 11, Vol 16, (2021) https://dx.doi.org/10.1088/1748-9326/ac2966

[2] Shukla, P. R.; Skea, J.; Calvo Buendia, E.; Masson-Delmotte, V.; et al. (eds.). IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse gas fluxes in Terrestrial Ecosystems, Intergovernmental Panel on Climate Change

[3] Seneviratne, Sonia I.; Zhang, Xuebin; Adnan, M.; Badi, W.; et al. (2021). Chapter 11: Weather and climate extreme events in a changing climate. Intergovernmental Panel on Climate Change, AR6 WG1 2021.

[4] Fovell, R. G. (1997). Consensus Clustering of U.S. Temperature and Precipitation Data. Journal of Climate, 10(6), 1405-1427.

[5] Bharath, R., Srinivas, V.V. and Basu, B. (2016), Delineation of homogeneous temperature regions: a two-stage clustering approach. Int. J. Climatol., 36: 165-187. https://doi.org/10.1002/joc.4335

[6] Xiaojing Wu, Raul Zurita-Milla and Menno-Jan Kraak (2015) Co-clustering geo-referenced time series: exploring spatio-temporal patterns in Dutch temperature data, International Journal of Geographical Information Science, 29:4, 624-642, DOI: 10.1080/13658816.2014.994520

[7] Barbosa, S. M., Scotto, M. G., and Alonso, A. M.: Summarising changes in air temperature over Central Europe by quantile regression and clustering, Nat. Hazards Earth Syst. Sci., 11, 3227-3233, https://doi.org/10.5194/nhess-11-3227-2011, 2011

[8] Y. Yu, Q. Shao, Z. Lin, Regionalization study of maximum daily temperature based on grid data by an objective hybrid clustering approach J. Hydrol., 564 (2018), pp. 149-163, 10.1016/j.jhydrol.2018.07.007

[9] Zubaroğlu, A., Atalay, V. Data stream clustering: a review. Artif Intell Rev 54, 1201-1236 (2021). https://doi.org/10.1007/s10462-020-09874-x

[10] P. P. Rodrigues, J. Gama and J. Pedroso, Hierarchical Clustering of Time-Series Data Streams, in IEEE Transactions on Knowledge and Data Engineering, vol. 20, no. 5, pp. 615-627, May 2008, doi: 10.1109/TKDE.2007.190727.

[11] Udommanetanakit K, Rakthanmanon T, Waiyamai K (2007) E-stream: Evolution-based technique for stream clustering. vol 4632, pp 605-615. https://doi.org/10.1007/978-3-540-73871-8_58

[12] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, kdd, vol. 96, no. 34, pp. 226-231, 1996.

[13] Ordonez C (2003) Clustering binary data streams with k-means. In: Proceedings of the 8th ACM SIGMOD workshop on research issues in data mining and knowledge discovery, DMKD '03, Association for Computing Machinery, New York, NY, USA, pp 12-19, https://doi.org/10.1145/882082.882087

[14] Lu, Y., Sun, Y., Xu, G., and Liu, G. (2005). A grid-based clustering algorithm for high-dimensional data streams. In Advanced Data Mining and Applications: First International Conference, ADMA 2005, Wuhan, China, July 22-24, 2005. Proceedings 1 (pp. 824-831). Springer Berlin Heidelberg.

[15] Sun Y, Lu Y (2006) A grid-based subspace clustering algorithm for high-dimensional data streams. In: Feng L, Wang G, Zeng C, Huang R (eds) Web information systems-WISE 2006 workshops. Springer, Berlin, pp 37-48

[16] Ester M, Kriegel HP, Sander J, Wimmer M, Xu X (1998) Incremental clustering for mining in a data warehousing environment. In: Proceedings of the 24rd international conference on very large data bases, VLDB '98, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 323-333

[17] Chen Y, Tu L (2007) Density-based clustering for real-time stream data. In: Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining, KDD '07, pp 133-142

[18] Zhou A, Cao F, Yan Y, Sha C, He X (2007) Distributed data stream clustering: a fast em-based approach. In: 2007 IEEE 23rd international conference on data engineering, pp 736-745

[19] Dang XH, Lee VCS, Ng WK, Ong KL (2009) Incremental and adaptive clustering stream data over sliding window. In: Bhowmick SS, Kung J, Wagner R (eds) Database and expert systems applications. Springer, Berlin, pp 660-674

[20] L. Schick, P. de Abreu Lopes and H. de Arruda Camargo, d-FuzzStream: A Dispersion-Based Fuzzy Data Stream Clustering, 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, Brazil, 2018, pp. 1-8, doi: 10.1109/FUZZ-IEEE.2018.8491534.

[21] C. Nordahl, V. Boeva, H. Grahn, and M. Persson Netz, EvolveCluster: an evolutionary clustering algorithm for streaming data, Evolving Systems, vol. 13, no. 4, pp. 603-623, Aug. 2022, doi: 10.1007/s12530-021-09408-y.

[22] A. Bechini, F. Marcelloni, and A. Renda, 'TSF-DBSCAN: A Novel Fuzzy Density-Based Approach for Clustering Unbounded Data Streams', IEEE Transactions on Fuzzy Systems, vol. 30, no. 3, pp. 623–637, Mar. 2022, doi: 10.1109/TFUZZ.2020.3042645.

[23] N. Rico, P. Huidobro, A. Bouchet, and I. Díaz, 'Similarity measures for interval-valued fuzzy sets based on average embeddings and its application to hierarchical clustering', Information Sciences, vol. 615, pp. 794–812, Nov. 2022, doi: 10.1016/j.ins.2022.10.028.

Espacios métricos parciales acotados y el problema de agregación

A. Mir-Fuentesl*†
*Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
†Institut d'Investigació Sanitària Illes Balears (IdISBa)
Hospital Universitari Son Espases.
Email: a.mir@uib.cat

O. Valerol*†
*Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears,
† Institut d'Investigació Sanitària Illes Balears (IdISBa)
Hospital Universitari Son Espases.
Email: o.valero@uib.es

***Resumen*—En 1981, Borsík y Doboš investigaron el problema de la agregación para espacios métricos. Así, caracterizaron aquellas funciones que permiten combinar una colección de métricas, cada una de ellas definidas en un conjunto, para obtener una sola métrica definida en el espacio producto como resultado.**

Posteriormente, en 1994, Matthews introdujo la noción de espacio métrico parcial con el objetivo de proporcionar una herramienta matemática adecuada para modelar ciertos procesos que surgen de modo natural en ciencias de la computación y en inteligencia artificial. Inspirados por la aplicabilidad mencionada de los espacios métricos parciales y por el hecho de que existen métricas parciales útiles en dicho campo que pueden ser inducidas mediante agregación, en 2015 Alghamdi, Shahzad y Valero exploraron el problema planteado por Borsík y Doboš en el marco de los espacios métrico parciales. Sin embargo, muchas de las métricas parciales empleadas en las aplicaciones son acotadas. Motivados por el hecho de que dicho caso no ha sido tratado desde el punto de vista de la agregación en la literatura, en este trabajo ofrecemos una descripción general de cómo agregar, mediante una función, una colección finita de espacios métricos parciales acotados para obtener una métrica parcial acotada definida en el espacio producto como resultado. Así, una caracterización de dichas funciones es proporcionada.

***Index Terms*—Espacio métrico parcial, agregación, monotonía, acotación**

I. Introducción y preliminares

La agregación de información es de suma importancia ya que nos permite tratar una gran cantidad de datos, pudiendo ser estos de naturaleza muy distinta, reduciéndolos a un único dato con el objetivo de poder llevar a cabo toma de decisiones que permitan resolver el problema bajo estudio. Al combinar diversas fuentes, potenciamos nuestra comprensión de los datos y la extracción de patrones significativos que nos faciliten la toma de decisiones. En última instancia, la agregación nos capacita para obtener conocimientos útiles a partir de piezas de información dispersas, impulsando la innovación y el progreso en diversos campos (véase [1]).

En muchas ocasiones, los datos a agregar representan diferentes distancias entre objetos. Con el objetivo de poder cuantificar una medida de distancia que englobe la información proporcionada por cada una de ellas, parece razonable trasladar las ideas procedentes de la teoría de la agregación de la información al caso métrico. Al combinar estas distancias, se puede obtener una visión más completa y representativa de la relación entre los elementos que intervienen en el problema tratado. Esto puede ser útil en diversas aplicaciones, como en la clasificación y detección de patrones en datos, donde la evaluación de la disimilitud entre objetos juega un papel fundamental, de modo que la integración de múltiples métricas puede proporcionar una medida más robusta y precisa de la distancia entre los objetos mencionados.

Varios autores han examinado minuciosamente qué funciones permiten combinar una colección de distancias cada una de ellas definidas en un conjunto, para obtener una sola distancia definida en el espacio producto como resultado. De hecho, en 1981 Borsík y Doboš investigaron a fondo dicho problema en el caso métrico. En concreto, caracterizaron aquellas funciones que permiten agregar una colección (no necesariamente finita) de espacios métricos a partir de lo que definieron como tripletes triangulares [2]. Dicho problema suscitó tanto interés que posteriormente se publicó una monografía tratándolo en profundidad (véase [3] y la referencias que contiene).

Desde que el problema de agregación de espacios métricos fue tratado por Borsík y Doboš, diversos autores han planteado y analizado el mismo problema en contextos más generales (véase, entre otros, [4]–[11]).

En 1994, Matthews introdujo la noción de espacio métrico parcial con el objetivo de proporcionar una herramienta matemática adecuada para modelar ciertos procesos que surgen de modo natural en ciencias de la computació y en inteligencia artificial ([12]–[17]). Recordemos que, dado un conjunto no vacío X, una métrica parcial sobre X es una función $p : X \times X \to \mathbb{R}^+$ ($\mathbb{R}^+$ denota el conjunto de los números reales no negativos) que cumple para todo $x, y, z \in X$ las siguientes condiciones:

(P1) $p(x, y) = p(y, x)$
(P2) $p(x, x) \leq p(x, y)$
(P3) $p(x, x) = p(x, y) = p(y, y) \Leftrightarrow x = y$
(P4) $p(x, z) + p(y, y) \leq p(x, y) + p(y, z)$

El par (X, p) se denomina espacio métrico parcial (EMP en lo sucesivo).

Esta investigación forma parte del proyecto PID2022-139248NB-I00 financiado por MICIU/AEI/10.13039/501100011033 y "FEDER Una manera de hacer Europa".

Inspirados por la aplicabilidad mencionada de los espacios métricos parciales y por el hecho de que existen métricas parciales útiles en dicho campo que pueden ser inducidas mediante agregación, Alghamdi, Shahzad y Valero plantearon el problema de Borsík y Doboš en el caso métrico parcial en [18] (véase también [19]). Concretamente, obtuvieron una caracterización de las funciones que agregan métricas parciales en términos de lo que llamaron tripletas triangulares parciales. Sin embargo, muchas de las métricas parciales empleadas en las aplicaciones son acotadas. El estudio de cómo agregar una colección de espacios métricos parciales donde cada una de las métricas son acotadas no ha sido tratado desde el punto de vista de la agregación en la literatura. Motivados por este hecho, en este trabajo nos centramos en cómo agregar, mediante una función, una colección finita de espacios métricos parciales acotados para obtener una métrica parcial acotada definida en el espacio producto como resultado.

Con el objetivo de proporcionar la mencionada caracterización definamos algunos conceptos que jugarán un papel clave.

Definición 1. *Diremos que un EMP (X, p) está acotado si para algún $c \in \mathbb{R}^+$ con $c > 0$ se tiene que, para todo $x, y \in X$, se satisface que $p(x, y) \leq c$. Para un EMP acotado, el menor valor (ínfimo) de c para el cual se satisface la desigualdad anterior se denomina constante de acotación, y entonces se dice que (X, p) es un espacio métrico parcial c-acotado.*

El siguiente ejemplo proporciona una instancia de espacio métrico parcial acotado que no es una métrica.

Ejemplo 1. *Considérese el conjunto de todas las sucesiones finitas e infinitas Σ^∞ definidas a partir de un alfabeto no vacío Σ. Dada $v \in \Sigma^\infty$, denotaremos por $\mathrm{l}(v)$ la longitud de v. De este nodo $\mathrm{l}(v) \in \mathbb{N} \cup \{\infty\}$ para toda $v \in \Sigma^\infty$. Fijemos $\Sigma_F = \{v \in \Sigma^\infty : \mathrm{l}(v) \in \mathbb{N}\}$ y $\Sigma_\infty = \{v \in \Sigma^\infty : \mathrm{l}(v) = \infty\}$. Claramente se tiene que $\Sigma^\infty = \Sigma_F \cup \Sigma_\infty$. Definamos la función $p_\Sigma : \Sigma^\infty \times \Sigma^\infty \to [0, 1]$ como $p_\Sigma(v, w) = 2^{-\mathrm{l}(v,w)}$ para todo $v, w \in \Sigma^\infty$, donde $\mathrm{l}(v, w)$ denota la longitud del prefijo común más largo entre v y w (obsérvese que consideramos $l(v, w) = 0$ cuando v y w no tiene un prefijo en común). Nótese que $l(u, u) = l(u)$ para toda $u \in \Sigma^\infty$. Obviamente hemos asumido que $2^{-\infty} = 0$. No es difícil verificar que $(\Sigma^\infty, p_\Sigma)$ es un EMP $\frac{1}{2}$-acotado.*

A continuación introducimos el concepto de función de agregación de espacios métricos parciales acotados.

Definición 2. *Dado $c \in \mathbb{R}^+$ con $c > 0$, diremos que una función $B : [0, c]^n \to \mathbb{R}^+$ es una función de agregación de espacios métricos parciales acotados (abreviada como FAEMPA) siempre que $P_B : X \times X \to \mathbb{R}^+$ sea una métrica parcial acotada para cada familia de espacios métricos parciales c-acotados $\{(X_i, p_i) \mid i = 1, 2, \ldots, n\}$, donde*

$$P_B(\bar{x}, \bar{y}) = B(p_1(x_1, y_1), p_2(x_2, y_2), \ldots, p_n(x_n, y_n))$$
$$= B((p_i(x_i, y_i))_{i=1}^n)$$

para todo $\bar{x} = (x_i)_{i=1}^n, \bar{y} = (y_i)_{i=1}^n \in X = \prod_{i=1}^n X_i$.

El resto del artículo está organizado del modo siguiente. En la Sección II se presentan algunas propiedades básicas que toda función de agregación de espacios métricos parciales acotados debe satisfacer. La Sección 1 está dedicada a proporcionar una caracterización de dichas funciones. Finalmente, en la Sección IV, se proporcionan alguna conclusiones y se indican algunas propuestas de trabajo futuro.

II. Funciones de agregación de espacios métricos parciales acotados: algunas propiedades básicas

Es interesante observar que la Definición 2 es una adaptación de la dada para el caso no acotado en [18] (véase también [19]). Nótese que en dicha definición tiene lugar un proceso de agregación de métricas parciales acotadas de tal modo que el resultado también es una métrica parcial acotada. Conservar la propiedad de acotación es un hecho valioso en aplicaciones prácticas donde necesitamos controlar la magnitud de las distancias entre objetos.

La siguiente proposición garantiza un hecho que parece natural que toda FAEMPA debería cumplir, la acotación.

Proposición 1. *Si B es una FAEMPA entonces B está acotada.*

Demostración. Sea $c \in \mathbb{R}^+$ con $c > 0$. Supongamos que $B : [0, c]^n \to \mathbb{R}^+$ es una FAEMPA. Consideremos el EMP acotado $([0, c], \text{máx})$. Entonces $([0, c]^n, P_B)$ es un EMP acotado donde P_B viene dada, para todo $\bar{x}, \bar{y} \in [0, c]^n$, por

$$P_B(\bar{x}, \bar{y}) = B((\text{máx}\{x_i, y_i\})_{i=1}^n).$$

De este modo existe $d \in \mathbb{R}^+$ con $d > 0$ tal que $P_B(\bar{x}, \bar{y}) \leq d$ para todo $\bar{x}, \bar{y} \in [0, c]^n$. Ahora, para todo $\bar{a} \in [0, c]^n$, consideremos $\bar{0} = (0, 0, \ldots, 0)$, entonces

$$B(\bar{a}) = B((\text{máx}\{a_i, 0\})_{i=1}^n) = P_B(\bar{a}, \bar{0}) \leq d$$

Deduciendo que B está acotada. □

La proposición siguiente proporciona dos propiedades de las FAEMPA que jugarán un papel crucial en la caracterizaciόm que posteriormente será introducida.

Proposición 2. *Sea $c \in \mathbb{R}^+$ con $c > 0$ y sea B una FAEMPA. Las siguientes afirmaciones se cumplen:*

1. *Si $B(\bar{a}) = 0$, entonces $\bar{a} = \bar{0}$.*
2. *B es monótona no decreciente.*

Demostración. Consideremos el EMP acotado $([0, c], \text{máx})$. Supongamos que $B(\bar{a}) = 0$, entonces

$$P_B(\bar{a}, \bar{0}) = B\left((\text{máx}\{a_i, 0\})_{i=1}^n\right) = B(\bar{a}) = 0.$$

Dado que P_B es una métrica parcial, se tiene que $\bar{a} = \bar{0}$.

Veamos a continuación la monótona de B . Sean $\bar{a}, \bar{b} \in [0, c]^n$ tal que $\bar{a} \leq \bar{b}$ ($a_i \leq b_i$ para todo $i = 1, \ldots, n$). Consideremos, de nuevo, el EMP acotado $([0, c], \text{máx})$. Dado que P_B es una métrica parcial, tenemos que

$$P_B(\bar{a}, \bar{a}) \leq P_B(\bar{a}, \bar{b}).$$

Así,

$$B\left((\text{máx}\{a_i, a_i\})_{i=1}^n\right) \leq B\left((\text{máx}\{a_i, b_i\})_{i=1}^n\right).$$

Luego

$$B(\bar{a}) = B\left((a_i)_{i=1}^n\right) \leq B\left((b_i)_{i=1}^n\right) = B(\bar{b}).$$

Por lo tanto, B es monótona no decreciente. □

Nótese que la propiedad de monotonía asegura que al aumentar las distancias entre puntos individuales, la distancia agregada también aumentará, lo cual está intimamente relacionado con la propiedad (P2) de una métrica parcial. Además, el hecho de que $B(\bar{a}) = 0$ está conectado con la propiedad de separación (P3) que toda métrica parcial debe satisfacer.

III. Funciones de agregación de espacios métricos parciales acotados: Una caracterización

En esta sección se proporciona una descripción completa de las funciones de aquellas funciones que son útiles para agregar los espacios métricos parciales acotados. Para ello resulta necesario introducir el concepto de triplete parcial.

Definición 3. *Sea* $\mathbb{R}_+^3 = \{(x_1, \ldots, x_3) : x_1, x_2, x_3 \in \mathbb{R}^+\}$. *Diremos que* $(a_1, a_2, a_3) \in \mathbb{R}_+^3$ *es un triplete parcial sobre* $(b_1, b_2, b_3) \in \mathbb{R}_+^3$ *si se cumplen las siguientes condiciones:*

(i) $a_1 \geq \text{máx}\{b_2, b_3\}$ *y* $a_1 + b_1 \leq a_2 + a_3$
(ii) $a_2 \geq \text{máx}\{b_1, b_3\}$ *y* $a_2 + b_2 \leq a_1 + a_3$
(iii) $a_3 \geq \text{máx}\{b_1, b_2\}$ *y* $a_3 + b_3 \leq a_1 + a_2$

Obsérvese que las condiciones dadas en la defición anterior pueden ser escritas de manera compacta para $k = 1, 2, 3$ del modo siguiente:

$$a_k \geq \max_{i \neq k}\{b_i\} \quad \text{y} \quad a_k + b_k \leq \sum_{i \neq k} a_i.$$

Ahora que hemos definido el concepto de triplete parcial, vamos a construir una métrica parcial acotada basada en éstos. Esta construcción va a ser crucial para identificar y demostrar las propiedades que determinan si una función es apta para agregar espacios métricos parciales acotados.

Proposición 3. *Sea* $\bar{a} = (a_1, a_2, a_3) \in \mathbb{R}_+^3$ *un triplete parcial sobre* $(b_1, b_2, b_3) \in \mathbb{R}_+^3$. *Entonces, se puede inducir un espacio métrico parcial acotado* $(X, p_{C\bar{a}\bar{b}})$.

Demostración. Dado que $a_k \geq \text{máx}_{i \neq k}\{b_i\}$ para todo $k = 1, 2, 3$, tenemos que $a_k \geq \text{mín}_{i \neq k}\{b_i\}$ para todo $k = 1, 2, 3$. A continuación, distinguimos tres casos:

Caso 1. $a_k > \text{mín}_{i \neq k}\{b_i\}$ para todo $k = 1, 2, 3$. Consideremos $X = \{x, y, z\}$, donde x, y, z son todos diferentes dos a dos. Definimos la función $p_{C\bar{a}\bar{b}} : X \times X \to \mathbb{R}^+$ de la siguiente manera:

$$\begin{aligned} p_{C\bar{a}\bar{b}}(x, y) = p_{C\bar{a}\bar{b}}(y, x) &= a_1 \qquad & p_{C\bar{a}\bar{b}}(z, z) &= b_1 \\ p_{C\bar{a}\bar{b}}(x, z) = p_{C\bar{a}\bar{b}}(z, x) &= a_2 \qquad & p_{C\bar{a}\bar{b}}(y, y) &= b_2 \\ p_{C\bar{a}\bar{b}}(y, z) = p_{C\bar{a}\bar{b}}(z, y) &= a_3 \qquad & p_{C\bar{a}\bar{b}}(x, x) &= b_3 \end{aligned}$$

No es difícil comprobar que $(X, p_{C\bar{a}\bar{b}})$ es un espacio métrico parcial acotado.

Caso 2. Supongamos que $a_k = \text{mín}_{i \neq k}\{b_i\}$ para dos índices $k \in \{1, 2, 3\}$. Podemos asumir sin pérdida de generalidad que la igualdad anterior se cumple para $k = 1, 2$. Entonces se sigue que $a_2 = a_1 = a_3 = b_1 = b_2 = b_3$. Establecemos $X = \{x\}$. Definimos la función $p_{C\bar{a}\bar{b}} : X \times X \to \mathbb{R}^+$ de la siguiente manera:

$$p_{C\bar{a}\bar{b}}(x, x) = a_1$$

Claramente, $(X, p_{C\bar{a}\bar{b}})$ es un espacio métrico parcial acotado.

Caso 3. Supongamos que $a_k = \text{mín}_{i \neq k}\{b_i\}$ solo para un índice $k \in \{1, 2, 3\}$. Podemos asumir sin pérdida de generalidad que la igualdad anterior se cumple para $k = 1$. Entonces $a_1 = \text{mín}\{b_2, b_3\}$, $a_2 > \text{mín}\{b_1, b_3\}$ y $a_3 > \text{mín}\{b_1, b_2\}$. De aquí se sigue que $a_1 = b_2 = b_3$. Fijemos $X = \{x, y\}$ con x e y diferentes. Definamos la función $p_{\bar{c}\bar{a}\bar{b}} : X \times X \to \mathbb{R}^+$ de la siguiente manera:

$$p_{C\bar{a}\bar{b}}(x, y) = p_{C\bar{a}\bar{b}}(y, x) = a_2,$$

$$p_{C,\bar{a}\bar{b}}(x, x) = b_3, \quad p_{C,\bar{a}\bar{b}}(y, y) = b_1.$$

Un cálculo sencillo muestra que $(X, p_{C,\bar{a}\bar{b}})$ es un espacio métrico parcial acotado. □

En lo sucesivo, el espacio métrico parcial $(X, p_{C\bar{a}\bar{b}})$ introducido en el resultado anterior será llamado espacio métrico parcial canónico inducido por el triplete parcial $\bar{a} = (a_1, a_2, a_3)$ sobre $\bar{b} = (b_1, b_2, b_3)$.

Los siguientes resultados, que se basan en la proposición anterior, nos proporcionan ciertas construcciones que nos ayudarán en la anunciada caracterización.

Lema 1. *Sea* $c \in \mathbb{R}^+$ *con* $c > 0$. *Entonces, para todo* $a, b, e, f \in [0, c]$ *tales que* $a + b \leq e + f$ *y* $b \leq e, b \leq f$, *existe un espacio métrico parcial acotado (X, p) de modo que existen* $x, y, z \in X$ *tales que* $p(x, y) = a', p(z, z) = b, p(x, z) = e$ *y* $p(y, z) = f$, *donde* $a' = \text{mín}\{e + f - b, c\}$.

Demostración. Consideremos $a' = \text{mín}\{e+f-b, c\}$. Tenemos que $a' \geq a$ y $a' + b \leq e + f$. A continuación mostramos que $a' + f - e \geq 0$. Si $e + f - b < c$, entonces $a' = e + f - b$ y así $a' + f - e = 2f - b \geq 0$. De lo contrario, $e + f - b \geq c$ da $a' = c$. Así, $a' + f - e = c + f - e \geq 0$ debido a que $c \geq e$.

Por lo tanto, tenemos $a' + f - e \geq 0$. De manera similar, se puede mostrar que $a' + e - f \geq 0$. Entonces es fácil ver que $\bar{a} = (a', e, f)$ es un triplete parcial sobre $\bar{b} = (b, f, e)$. Nótese que $a', b_2, b_3 \in [0, c]$. Por la Proposición 3, concluimos que el espacio métrico parcial canónico $(X, p_{C\bar{a}\bar{b}})$ satisface que $p(x, y) = a'$, $p(z, z) = b$, $p(x, z) = e$ y $p(y, z) = f$. □

Lema 2. *Sea* $c \in \mathbb{R}^+$ *con* $c > 0$. *Entonces, para todo* $a, b_1, b_2 \in [0, c]$ *con* $a \geq b_1$ *y* $a \geq b_2$, *existe un espacio métrico parcial acotado* (X, p) *de modo que existen* $x, y \in X$ *tales que* $p(x, y) = a$, $p(x, x) = b_1$ *y* $p(y, y) = b_2$.

Demostración. Es fácil ver que $\bar{a} = (a, a, a)$ es un triplete parcial sobre $\bar{b} = (0, b_1, b_2)$. La Proposición 3 establece que

el espacio métrico parcial canónico $(X, p_{C\bar{a}\bar{b}})$ satisface que $p(x,y) = a$, $p(x,x) = b_2$ y $p(y,y) = b_1$. □

A la vista de lo expuesto tenemos todas las herramientas necesarias para poder introducir la caracterización de las funciones de agregación de espacios métricos parciales acotados.

Teorema 1. *Sea $c \in \mathbb{R}^+$ con $c > 0$. Entonces, la función $B : [0,c]^n \to \mathbb{R}^+$ es una FAEMPA si y solo si satisface las siguientes condiciones:*

(1) *Si $\bar{a}, \bar{b}, \bar{e}, \bar{f} \in [0,c]^n$ tal que $\bar{b} \leq \bar{e}, \bar{b} \leq \bar{f}, \bar{a} + \bar{b} \leq \bar{e} + \bar{f}$, entonces $B(\bar{a}) + B(\bar{b}) \leq B(\bar{e}) + B(\bar{f})$.*

(2) *Si $\bar{a}, \bar{b}_1, \bar{b}_2 \in [0,c]^n$ con $\bar{b}_1 \leq \bar{a}$ y $\bar{b}_2 \leq \bar{a}$ y $B(\bar{a}) = B(\bar{b}_1) = B(\bar{b}_2)$, entonces $\bar{a} = \bar{b}_1 = \bar{b}_2$.*

Demostración. Supongamos que $B : [0,c]^n \to \mathbb{R}^+$ es una FAEMPA.

(1) Consideremos $\bar{a}, \bar{b}, \bar{e}, \bar{f} \in [0,c]^n$ tales que $\bar{a} + \bar{b} \leq \bar{e} + \bar{f}$, $\bar{b} \leq \bar{e}$ y $\bar{b} \leq \bar{f}$. Entonces para cada $i \in \{1, 2, \ldots, n\}$ tenemos que $a_i + b_i \leq e_i + f_i$, $b_i \leq e_i$ y $b_i \leq f_i$, donde $a_i, b_i, e_i, f_i \in [0,c]$. Por el Lema 1, existe un espacio métrico parcial acotado (X_i, p_i) y existen $x_i, y_i, z_i \in X_i$ tales que $p_i(x_i, y_i) = a'_i$, $p_i(z_i, z_i) = b_i$, $p_i(x_i, z_i) = e_i$ y $p_i(y_i, z_i) = f_i$, donde $a'_i = e_i + f_i - b_i \geq a_i$. Para el espacio métrico parcial $X = \prod_{i=1}^n X_i$ con la métrica parcial P_B, tenemos que $P_B(\bar{x}, \bar{y}) = B(\bar{a'})$, $P_B(\bar{z}, \bar{z}) = B(\bar{b})$, $P_B(\bar{x}, \bar{z}) = B(\bar{e})$ y $P_B(\bar{y}, \bar{z}) = B(\bar{f})$. Dado que (X, P_B) es un espacio métrico parcial, tenemos que $P_B(\bar{x}, \bar{y}) + P_B(\bar{z}, \bar{z}) \leq P_B(\bar{x}, \bar{z}) + P_B(\bar{y}, \bar{z})$. Por lo tanto,

$$B(\bar{a'}) + B(\bar{b}) \leq B(\bar{e}) + B(\bar{f})$$

Dado que $\bar{a} \leq \bar{a'}$ y la Proposición 2 proporciona la monotonía de B, obtenemos

$$B(\bar{a}) \leq B(\bar{a'}) \leq B(\bar{e}) + B(\bar{f}) - B(\bar{b})$$

(2) Supongamos que para algunos $\bar{a}, \bar{b}_1, \bar{b}_2 \in [0,c]^n$ con $\bar{b}_1 \leq \bar{a}$ y $\bar{b}_2 \leq \bar{a}$ tenemos que

$$B(\bar{a}) = B(\bar{b}_1) = B(\bar{b}_2)$$

Ahora, para cada $i \in \{1, 2, \ldots, n\}$, tenemos $b_{1i} \leq a_i$ y $b_{2i} \leq a_i$. Por el Lema 2, existe un espacio métrico parcial (X_i, p_i) tal que

$$p_i(x_i, y_i) = a_i, \quad p_i(x_i, x_i) = b_{2i}, \quad p_i(y_i, y_i) = b_{1i}.$$

Para el espacio métrico parcial P_B definido en $X = \prod_{i=1}^n X_i$, tenemos

$$P_B(\bar{x}, \bar{y}) = B(\bar{a}), \quad P_B(\bar{x}, \bar{x}) = B(\bar{b}_2),$$
$$P_B(\bar{y}, \bar{y}) = B(\bar{b}_1).$$

De donde se deduce que

$$B(\bar{a}) = B(\bar{b}_1) = B(\bar{b}_2).$$

Luego

$$P_B(\bar{x}, \bar{y}) = P_B(\bar{x}, \bar{x}) = P_B(\bar{y}, \bar{y}).$$

Por tanto, $\bar{x} = \bar{y}$ y, así, $\bar{a} = \bar{b}_1 = \bar{b}_2$.

Para probar el recíproco, sea $B : [0,c]^n \to \mathbb{R}^+$ una función tal que se cumplen las afirmaciones (1) y (2).

Sea $(X_i, p_i)_{i=1}^n$ una familia de espacios métricos parciales c-acotados. Consideremos $\bar{x}, \bar{y} \in X = \prod_{i=1}^n X_i$ tales que

$$P_B(\bar{x}, \bar{y}) = P_B(\bar{x}, \bar{x}) = P_B(\bar{y}, \bar{y}),$$

es decir,

$$B\left((p_i(x_i, y_i))_{i=1}^n\right) = B\left((p_i(x_i, x_i))_{i=1}^n\right) = B\left((p_i(y_i, y_i))_{i=1}^n\right).$$

El hecho de que, por un lado, se verifique que $p_i(x_i, x_i) \leq p_i(x_i, y_i)$ y $p_i(y_i, y_i) \leq p_i(x_i, y_i)$ para todo $i \in \{1, 2, \ldots, n\}$ y que, por otro, se verifique la condición (2) implica que $p_i(x_i, y_i) = p_i(x_i, x_i) = p_i(y_i, y_i)$ para todo $i \in \{1, 2, \ldots, n\}$. Dado que cada (X_i, p_i) es un espacio métrico parcial, tenemos que $x_i = y_i$ para todo i. Por lo tanto, $\bar{x} = \bar{y}$. Luego P_B verifica la condición (P1) de la definición de métrica parcial.

Tomando $\bar{b} = \bar{f} = \bar{0}$ en la afirmación (2), obtenemos que B es monótona. Ahora, para $\bar{x}, \bar{y} \in X$, tenemos que $p_i(x_i, x_i) \leq p_i(x_i, y_i)$ para todo $i \in \{1, 2, \ldots, n\}$. Por lo tanto, $B\left((p_i(x_i, x_i)_{i=1}^n)\right) \leq B\left((p_i(x_i, y_i))_{i=1}^n\right)$, lo que implica que $P_B(\bar{x}, \bar{x}) \leq P_B(\bar{x}, \bar{y})$. Con lo cual P_B verifica la condición (P2) de la definición de métrica parcial.

Claramente, dado que $p_i(x_i, y_i) = p_i(y_i, x_i)$, se tiene que $P_B(\bar{x}, \bar{y}) = P_B(\bar{y}, \bar{x})$. Por lo tanto, P_B verifica la condición (P3) de la definición de métrica parcial.

Finalmente, para $\bar{x}, \bar{y}, \bar{z} \in X$, tenemos que

$$p_i(x_i, y_i) + p_i(z_i, z_i) \leq p_i(x_i, z_i) + p_i(y_i, z_i)$$

para todo $i \in \{1, 2, \ldots, n\}$. La condición (1) entonces implica que

$$B\left((p_i(x_i, y_i))_{i=1}^n\right) + B((p_i(z_i, z_i))_{i=1}^n)$$
$$\leq B((p_i(x_i, z_i))_{i=1}^n) + B((p_i(z_i, y_i))_{i=1}^n)$$

Esto es equivalente a

$$P_B(\bar{x}, \bar{y}) + P_B(\bar{z}, \bar{z}) \leq P_B(\bar{x}, \bar{z}) + P_B(\bar{y}, \bar{z})$$

Por lo tanto, (X, P_B) es un espacio métrico parcial. Nos falta verificar que es acotado. Sin embargo, observemos que de la condición (1), tomando $\bar{b} = \bar{f} = \bar{0}$ y $\bar{e} = (c, \ldots, c)$, se deduce que $B(\bar{a}) \leq B(c, \ldots, c)$. Luego B es una función acotada. Por tanto, se tiene que

$$P_B(\bar{x}, \bar{y}) = B((p_i(x_i, y_i))_{i=1}^n) \leq B(c, \ldots, c)$$

para todo $\bar{x}, \bar{y} \in X$. De donde concluimos que (X, P_B) es un espacio métrico parcial acotado. Se sigue entonces que B es una FAEMPA. □

El siguiente ejemplo proporciona una instancia de FAEMPA.

Ejemplo 2. *Considérese la función $B : [0,1]^2 \to \mathbb{R}^+$ definida por $B(\bar{a}) = \frac{a_1 + a_2}{4} + \frac{1}{2}$ para todo $\bar{a} = (a_1, a_2) \in [0,1]^2$. No resulta difícil comprobar que B satisface las condiciones (1) y (2) del Teorema 1. Por tanto, B es una FAEMPA. Obsérvese*

que dicha función induce un espacio métrico parcial acotado cuya constante de acotación no coincide con la constante de acotación de los espacios métricos parciales que agrega. En efecto, consideremos el espacio métrico parcial acotado $([0,1], \text{máx})$. *Sean* $([0,1], p_1)$ *y* $([0,1], p_2)$ *los espacios métricos parciales acotados tales que* $p_1(x,y) = p_2(x,y) = \frac{1}{2}\,\text{máx}(x,y)$ *para todo* $x, y \in [0,1]$. *Claramente* $p_1(x,y) \leq \frac{1}{2}$ *y* $p_2(x,y) \leq \frac{1}{2}$ *par todo* $x, y \in [0,1]$. *Sin embargo, el espacio métrico parcial inducido a través de* B *al agregar* $([0,1]^2, P_B)$ *es* $\frac{3}{4}$*-acotado. Con lo cual, la función de agregación no preserva la constante de acotación.*

IV. Conclusiones

Motivados por el creciente interés en las métricas parciales acotadas y la importancia del problema de agregación de distancias en problemas aplicados, en este trabajo hemos planteado el problema de fusionar una colección de espacios métricos parciales acotados dando como resultado un nuevo espacio métrico parcial acotado definido en el producto. Para ello, hemos introducido la noción de función de agregación de espacios métricas parciales acotados y se ha dado una caracterización de este tipo de funciones. Como consecuencia se ha demostrado que deben ser monótonas no decrecientes, acotadas y satisfacer una propiedad especial de separación. Sin embargo, aún quedan algunas preguntas abiertas.

En términos generales la función de agregación induce un espacio métrico parcial acotado cuya constante de acotación no coincide con la constante de acotación de los espacios métricos parciales agregados. Parece interesante analizar bajo qué condiciones la función de agregación preserva la constante de acotación.

Dado que toda métrica acotada es una métrica parcial acotada, planeamos analizar la relación existente entre las funciones que agregan espacios métricos parciales acotados y aquellas que agregan espacios métricos acotados. Además, prestaremos especial atención a la cuestión sobre cuándo las primeras preservan los espacios métricos, es decir, cuándo al agregar una colección de espacios métricos acotados induce como resultado concretamente un espacio métrico acotado sobre el producto.

Se explorará la posibilidad de extender el problema abordado considerando una colección de espacios métricos parciales acotados donde cada uno de ellos admite una constante de acotación diferente, y se analizará cuál es la relación entre las constantes de acotación originales y la obtenida mediante la agregación.

Finalmente, exploraremos la posibilidad de aplicar la teoría desarrollada al procesamiento de imágenes.

Referencias

[1] M. Grabisch, J.-L. Marichal, lu Radko Mesiar, and E. Pap, “”aggregation functions”, cambridge university press,” *2008 6th International Symposium on Intelligent Systems and Informatics*, pp. 1–7, 2008.

[2] Borsík and Doboš, “On a product of metric spaces,” *Mathematica Slovaca*, no. 2, pp. 193–205.

[3] J. Borsík, *Metric Preserving Functions*. Košice: Štroffek, 1998.

[4] A. Pradera and E. Trillas, “A note on pseudometrics aggregation,” *International Journal of General Systems*, vol. 31, no. 1, pp. 41–51, 2002.

[5] M. Deza and E. Deza, *Encyclopedia of Distances*. Springer-Verlag, 2009.

[6] G. Mayor and O. Valero, “Aggregation of asymmetric distances in computer science,” *Information Sciences*, vol. 180, no. 6, pp. 803–812, 2010.

[7] S. Massanet and O. Valero, “Is the aggregation of extended quasi-metrics an ex- tended quasi-metric?” in *Proceedings of the Workshop in Applied Topology (WiAT’12)*, 2012, pp. 1–8.

[8] ——, “On aggregation of metric structures: the extended quasi-metric case,” *International Journal of Computational Intelligence Systems*, vol. 6, no. 1, pp. 115–126, 2013.

[9] ——, “When is the aggregation of weightable metrics a weightable metric?” in *Proceedings of the XVIII Spanish Conference on Fuzzy Tecnologies and Fuzzy Logic (Estylf 2014)*, 2014, pp. 283–288.

[10] J.-J. Miñana and O. Valero, “On matthews’ relationship between quasi-metrics and partial metrics: An aggregation perspective,” *Results in Mathematics*, vol. 75, 2020.

[11] M. Lichman, P. Nowakowski, and F. Turoboś, “On functions preserving products of certain classes of semimetric spaces,” *Tatra Mountains Mathematical Publications*, vol. 78, pp. 175–198, 2021.

[12] S. G. Matthews, “Partial metric topology,” *Annals of the New York Academy of Sciences*, vol. 728, no. 1, pp. 183–197, 1994.

[13] ——, “An extensional treatment of lazy data flow dedlock,” *Theoretical Computer Science*, vol. 151, no. 1, pp. 195–205, 1995.

[14] A. Stojmirović and Y. Yu, “Geometric aspects of biological sequence comparison,” *Journal of Computational Biology*, vol. 16, no. 4, pp. 579–610, 2009.

[15] A. Stojmirović, “Quasi-metrics, similarities and searches: aspects of geometry of protein datasets,” Ph.D. dissertation, Victoria University of Wellington, 2005.

[16] P. Hitzler and S. A.K, “Generalized distance functions in the theory of computation,” *The Computer Journal*, vol. 53, no. 4, pp. 443–464, 2010.

[17] P. Hitzler and S. A.K., *Mathematical Aspects of Logic Programming Semantics*. Boca Raton, FL: CRC Press, 2011.

[18] M. Alghamdi, S. Naseer, and O. Valero, “Projective contractions, generalized metrics, and fixed points,” *Fixed Point Theory and Applications*, vol. 2015:81, 2015.

[19] S. Massanet and O. Valero, “New results on metric aggregation,” in *Proceedings of the 17th Spanish Conference on Fuzzy Technology and Fuzzy Logic (Estylf 2012)*, 2012, pp. 558–563.

Similitudes basadas en medidas de incrustación

Agustina Bouchet
Dept. Estadística, I.O. & D.M.
Unvidersidad de Oviedo
Oviedo, España
bouchetagustina@uniovi.es

Susana Díaz-Vázquez
Dept. Estadística, I.O. & D.M.
Unvidersidad de Oviedo
Oviedo, España
diazsusana@uniovi.es

Irene Díaz
Dept. de Informática
Universidad de Oviedo
Oviedo, España
sirene@uniovi.es

Susana Montes
Dept. Estadística, I.O. & D.M.
Unvidersidad de Oviedo
Oviedo, España
montes@uniovi.es

***Abstract*—En este trabajo demostramos la conexión entre los conceptos de incrustación y similitud. Proporcionamos una nueva definición de medida de similitud entre intervalos basada en cuán incrustados están cada uno de los intervalos considerados en el otro. Medidas de similitud conocidas en la literatura proporcionan una similitud cero a cualquier par de intervalos que tengan intersección vacía. La novedad de nuestro enfoque radica en el hecho de que definimos una medida de similitud que puede distinguir si dos intervalos están cerca uno del otro a pesar de no superponerse, ya que no proporciona un valor cero cuando los intervalos tienen intersección vacía.**

***Index Terms*—medida de similitud, función de agregación, medida de incrustación, intervalos**

I. Introducción

El concepto de similitud es fundamental y ampliamente utilizado en muchas de las disciplinas científicas. Algunas de estas disciplinas incluyen la toma de decisiones, el reconocimiento de patrones, métodos de agrupamiento, entre otras [4], [6], [7], [10], [11], [13], [18], [20]–[23]. Estas numerosas aplicaciones han hecho que el concepto matemático de similitud entre diferentes elementos haya sido ampliamente estudiado en la literatura durante los últimos años (véanse [5], [7], [8], [15], [20], entre muchos otros).

Se podría decir que todas las medidas de distancia en la recta real generan, mediante una transformación decreciente, una medida de similitud entre números. No obstante, no es tan evidente como cuantificar la similitud o cercanía entre una extensión de números, como pueden ser los intervalos. Sin embargo, hoy en día los datos que vienen descritos mediante intervalos aparecen en muchos campos, fruto de la obtención de los mismos, siendo especialmente significativo su uso si los datos son subjetivos y proceden de opiniones de personas. Es evidente que los números reales no siempre son la solución idónea en este campo, pudiendo ser demasiado inflexibles o estrictos si se quiere reflejar el pensamiento humano de manera fiel. Así, el uso de intervalos está siendo una alternativa exitosa para expresar formalmente opiniones, grados de preferencia, etc. Esta alternativa para representar algunos aspectos de la vida real requiere todo un conjunto de nuevas herramientas matemáticas que permitan trabajar con ellos de manera formal y, en particular, se precisa de medidas de similitud entre intervalos. A pesar de que la similitud entre intervalos no es un tema nuevo y se pueden encontrar propuestas previas en la literatura, hemos observado un inconveniente importante en algunas de las definiciones de similitud más comúnmente utilizadas. Generalmente, proporcionan resultados satisfactorios cuando los intervalos tienen intersección no vacía, es decir, dan resultados lógicos o cercanos a la intuición. Sin embargo, son poco informativos cuando los intervalos son disjuntos, no siendo capaces de distinguir si los intervalos están o no cercanos entre sí. Este trabajo surge con el objetivo de dar medidas de similitud que sí tengan en cuenta esta circunstancia. En concreto, nuestra propuesta se basará en el concepto de medida de incrustación que presentamos en el artículo [3].

Una función de incrustación es una función de valor real que proporciona el grado en que un intervalo está contenido en otro. Es bien sabido que el concepto de subconjuntos y las medidas de similitud son ideas conectadas (véase, por ejemplo, [16]). Presentamos una nueva manera de medir la similitud basada en el concepto de incrustación. En una primera propuesta presentamos una definición de similitud basada en la agregación de medidas de incrustación. Sin embargo, un análisis detallado de la nueva definición muestra que no captura la idea de proximidad entre los intervalos cuando son disjuntos. Abordamos este problema con una segunda propuesta. Mostramos que esta segunda definición permite medir la proximidad entre los intervalos tanto cuando tienen intersección no vacía, como cuando son disjuntos.

El trabajo está organizado de la siguiente manera: en la sección II introducimos algunos conceptos necesarios y establecemos la notación utilizada en las siguientes secciones. El núcleo del manuscrito se encuentra en la sección III. Contiene nuestras propuestas para medir la similitud entre intervalos y una discusión sobre la idoneidad de cada definición. En la sección IV presentamos algunas conclusiones y algunos de los problemas abiertos que se derivan del contenido de este trabajo.

II. Preliminares

En esta sección, introducimos las notaciones y principales conceptos que se utilizarán a lo largo de este trabajo.

Dado que la información que proporciona una persona se formaliza mejor, en términos de capturar la imprecisión asociada al pensamiento humano, mediante un intervalo, vamos a trabajar con dichos conjuntos. Solo consideramos, por ser un

Los autores han sido apoyados por el Proyecto del Ministerio de Ciencia e Innovación de España PID2022-139886NB-l00.

planteamiento habitual en este área, intervalos contenidos en $[0,1]$. Por lo tanto, llamamos $L([0,1])$ a la familia de intervalos cerrados incluidos en el intervalo unitario $[0,1]$:

$$L([0,1]) = \{[\underline{a}, \overline{a}] \mid \underline{a}, \overline{a} \in [0,1] \text{ y } \underline{a} \leq \overline{a}\} \quad (1)$$

Por otro lado, trabajamos con medidas de incrustación. En [3], la incrustación para intervalos se presentó de la siguiente manera:

***Definición** 2.1:* [3] La función $E : L([0,1])^2 \to [0,1]$ es una medida de incrustación (*embedding* en inglés) en $L([0,1])$ si para cualesquiera $a, b, c \in L([0,1])$ se satisfacen las siguientes propiedades:

A.1 $E(a,b) = 1$ si y solo si $a \subseteq b$
A.2 Si $a \cap b = \emptyset$, entonces $E(a,b) = E(b,a) = 0$
A.3 Si $b \subseteq c$, entonces $E(a,b) \leq E(a,c)$
A.4 Si $a \subseteq b \subseteq c$, entonces $E(c,a) \leq E(b,a)$

Un ejemplo de una medida de incrustación es aquella basada en el ancho de los intervalos [3]:

$$E_w(a,b) = \begin{cases} 1 & \text{si } w(a) = 0 \text{ y } a \cap b \neq \emptyset \\ 0 & \text{si } w(a) = 0 \text{ y } a \cap b = \emptyset \\ \frac{w(a \cap b)}{w(a)} & \text{si } w(a) \neq 0 \end{cases} \quad (2)$$

donde w denota la amplitud del intervalo, es decir,

$$w(a) = w([\underline{a}, \overline{a}]) = \overline{a} - \underline{a}. \quad (3)$$

Como se puede deducir de la definición, estos operadores permiten medir la inclusión de un intervalo en otro. Son una forma de medir el grado de contenido entre intervalos. Estas funciones alcanzan el valor máximo de 1 cuando el primer intervalo está completamente incluido en el segundo y solo en este caso. Proporcionan un valor en el rango $(0,1)$, que representa el grado de inclusión, si la intersección entre ambos intervalos es no vacía. Y, por último, da como resultado 0 si ambos intervalos no tienen elementos en común.

Otro concepto importante en este trabajo son las normas triangulares:

***Definición** 2.2:* [12] Una función $T : [0,1]^2 \to [0,1]$ es una norma triangular si satisface las siguientes propiedades:

- es conmutativa,
- es asociativa,
- es creciente en cada componente y
- $T(1,x) = x$ para todo $x \in [0,1]$.

Las normas triangulares, también denominadas t-normas, surgieron en el entorno de los espacios métricos probabilísticos, pero han sido ampliamente utilizadas para extender el operador lógico "y" en el contexto de la lógica difusa. Como hemos visto anteriormente, originalmente se definieron para solo dos argumentos, pero pueden operar sobre cualquier tupla de n argumentos con $n \in \mathbb{N}$ ya que son asociativas, es decir,

$$T(x_1, \ldots, x_n) = T(\ldots(T(T(x_1, x_2), x_3), \ldots, x_n) \quad (4)$$

Un miembro importante de esta familia de operadores es el mínimo, que denotaremos en lo que sigue como min y viene dado por $min(x,y) = \min\{x,y\}$. Otras t-normas importantes son el operador de Łukasiewicz:

$$T_L(x,y) = max\{x + y - 1, 0\} \quad (5)$$

y el producto:

$$T_P(x,y) = x \cdot y. \quad (6)$$

Las t-normas son una familia particular de funciones de agregación. Recordamos la definición de esta familia más general de operadores:

***Definición** 2.3:* [1], [2], [14] Una función de agregación $\mathcal{M} : [0,1]^n \to [0,1]$ es una función que satisface las siguientes propiedades:

- es creciente en cada argumento y
- $\mathcal{M}(0, \ldots, 0) = 0$ y $\mathcal{M}(1, \ldots, 1) = 1$.

Es evidente que estos operadores son más generales que las t-normas. No requieren conmutatividad ni asociatividad. Si son conmutativos, se denominan simétricos:

***Definición** 2.4:* [1] Una función de agregación $\mathcal{M}$ es simétrica, si su resultado no depende de la permutación de los argumentos, es decir,

$$\mathcal{M}(x_1, x_2, \ldots, x_n) = \mathcal{M}(x_{P(1)}, x_{P(2)}, \ldots, x_{P(n)}) \quad (7)$$

para todo $(x_1, x_2, \ldots, x_n) \in [0,1]^n$ y para toda permutación $P = (P(1), P(2), \ldots, P(n))$ de $(1, 2, \ldots, n)$.

***Definición** 2.5:* [19] Una función de agregación $\mathcal{M}$ se denomina 1-estricta cuando:

$$\mathcal{M}(x_1, x_2, \ldots, x_n) = 1 \quad (8)$$

solo si $x_i = 1$ para todo $i = 1, \ldots, n$.

Es claro que las t-normas son 1-estrictas. Esto se debe a que $T(x_1, x_2, \ldots, x_n) \leq min(x_1, x_2, \ldots, x_n)$ para cada $(x_1, x_2, \ldots, x_n) \in [0,1]^n$, entonces si $T(x_1, x_2, \ldots, x_n) = 1$ se tiene que $(x_1, x_2, \ldots, x_n) = (1, \ldots, 1)$, siendo la otra implicación trivial. Por lo tanto, son una familia particular de operadores de agregación simétricos (conmutativos) 1-estrictos, teniendo además la propiedad de asociatividad.

El último concepto que necesitamos establecer antes de comenzar con los nuevos resultados es el de medida de similitud. La idea habitual que subyace es que es una función de valor real que proporciona un valor que cuantifica cómo de similares son los dos elementos bajo comparación. En nuestro caso, los elementos a comparar son intervalos contenidos en el intervalo $[0,1]$ y, como consecuencia de esto, el grado de similitud está acotado en el intervalo unitario. Cuanto más similares sean los dos intervalos, mayor será el resultado de esta medida, alcanzándose el límite superior (1) solo si ambos intervalos son idénticos. Se han considerado diferentes axiomas en la definición de una medida de similitud. La

definición más básica se puede encontrar en [17] y la más restrictiva en [11], donde se requieren hasta nueve axiomas. Definimos la similitud considerando solo los cuatro axiomas identificados en [11] como "las cuatro propiedades comunes". Cabe notar que aquí solo enumeramos explícitamente tres de esos axiomas porque el cuarto considerado en [11] es que la similitud debe estar acotada entre 0 y 1 y eso ya se deduce de forma implícita en nuestra definición:

***Definición** 2.6:* Una medida de similitud entre intervalos es una función $S : L([0,1])^2 \to [0,1]$ tal que satisfice las siguientes propiedades para cualesquiera $a, b, c \in L([0,1])$:

$S1$: $S(a,b) = S(b,a)$
$S2$: $S(a,b) = 1$ si y solo si $a = b$
$S3$: Si $a \subseteq b \subseteq c$, entonces $S(a,c) \leq S(a,b)$ y $S(a,c) \leq S(b,c)$

Un ejemplo muy conocido de medida de similitud es el índice de Jaccard [9], definido como:

$$J(a,b) = \frac{w(a \cap b)}{w(a \cup b)} \tag{9}$$

Es fácil demostrar que este operador definido sobre $L([0,1])$ es una medida de similitud ya que siempre toma valores en $[0,1]$ y satisface los axiomas $S1 - S3$.

Otro ejemplo muy conocido de medida de similitud es el coeficiente de Sorensen, también llamado índice de Sorensen-Dice o índice de Dice [5], que se define como:

$$D(a,b) = \frac{2\, w(a \cap b)}{w(a) + w(b)} \tag{10}$$

También es fácil demostrar que es una medida de similitud en el sentido de la definición 2.6.

III. Similitudes basadas en medidas de incrustación

La idea principal de esta sección es introducir similitudes basadas en medidas de incrustación. Es claro que, en el contexto nítido, es absolutamente intuitivo pensar que dos conjuntos (intervalos) son iguales si el primero está incluido en el segundo y viceversa:

$$a = b \iff a \subseteq b \text{ y } b \subseteq a \tag{11}$$

Pero esta definición es estricta en el sentido de que solo podemos determinar si son iguales o no, pero no podemos determinar los grados de igualdad o similitud que estamos buscando. Si queremos permitir grados de similitud, podemos reemplazar $a \subseteq b$ por una medida de incrustación $E(a,b)$ y la conjunción clásica por el operador min obteniendo:

$$min(E(a,b), E(b,a)) \tag{12}$$

Se puede demostrar que es una función de similitud de acuerdo con la definición 2.6. La generalización natural de esta definición es reemplazar la función min por cualquier t-norma:

***Definición** 3.1:* Sea E una medida de incrustación y T una t-norma. La función $S_T^E : L([0,1])^2 \to [0,1]$ dada por:

$$S_T^E(a,b) = T(E(a,b), E(b,a)) \tag{13}$$

se llama similitud basada en t-norma según la incrustación.

Es conocido que $T_D \leq T \leq T_{min}$ siendo T_D la norma drástica, T cualquier norma y T_{min} la norma del mínimo. Por lo tanto, $S_{T_D}^E \leq S_T^E \leq S_{T_{min}}^E$.

La definición 3.1 generaliza la medida de similitud basada en subconjuntos bidireccionales introducida en [11]. La hemos llamado "similitud" porque cumple con las propiedades de la definición 2.6, como mencionamos en la siguiente proposición:

***Proposición** 3.1:* Sea E una medida de incrustación y T una t-norma. La similitud basada en t-norma según la incrustación S_T^E introducida en la definición 3.1 es una medida de similitud para intervalos.

Sin embargo, la asociatividad no es necesaria y la condición de frontera que satisfacen las t-normas ($T(x,1) = x$) tampoco es importante. Solo se utiliza la propiedad 1-estricto (junto con la monotonía). Por lo tanto, podemos extender la definición anterior a cualquier función de agregación 1-estricta de la siguiente manera:

***Definición** 3.2:* Sea E una medida de incrustación para intervalos y $\mathcal{M}$ una función de agregación 1-estricta y simétrica. La función $S_{\mathcal{M}}^E : L([0,1])^2 \to [0,1]$ dada por:

$$S_{\mathcal{M}}^E(a,b) = \mathcal{M}(E(a,b), E(b,a)) \tag{14}$$

se denomina similitud basada en la función de agregación según la incrustación.

La siguiente proposición muestra que el nombre considerado es el adecuado, puesto que esta función es realmente una medida de similitud.

***Proposición** 3.2:* Sea E una medida de incrustación para intervalos y $\mathcal{M}$ una función de agregación simétrica y 1-estricta. La similitud basada en la función de agregación según la incrustación dada en la definición 3.2 y denotada como $S_{\mathcal{M}}^E$ es una medida de similitud para intervalos.

El resultado podría extenderse aún más a operadores que no necesariamente cumplen con la definición de función de agregación, ya que la propiedad de frontera $\mathcal{M}(0,\ldots,0) = 0$ no se utiliza. No obstante, nos vamos a centrar en el caso de medidas generadas a partir de funciones de agregación, dada la habitual importancia de las mismas en la literatura.

Una función particularmente importante en la familia de operadores de agregación simétricos y 1-estrictos es la media aritmética, la cual, por tanto, nos permite definir una medida de similitud:

***Corolario** 3.1:* Sea E una medida de incrustación. La función $S^E_{AM} : L([0,1])^2 \to [0,1]$ dada por:

$$S^E_{AM}(a,b) = \frac{E(a,b) + E(b,a)}{2} \tag{15}$$

es una similitud para intervalos.

La función de similitud introducida en la proposición 3.2 es una forma fácil y natural de medir cuánto se parecen dos intervalos cuando tienen algunos elementos en común. El siguiente ejemplo muestra este comportamiento.

***Ejemplo** 3.1:* Examinemos la similitud entre el intervalo $x_1 = [0.2, 0.5]$ y los intervalos $x_2 = [0.4, 0.8]$, $x_3 = [0.1, 0.8]$ y $x_4 = [0.6, 0.9]$. En la figura 1 se observa la representación gráfica de cada uno de ellos.

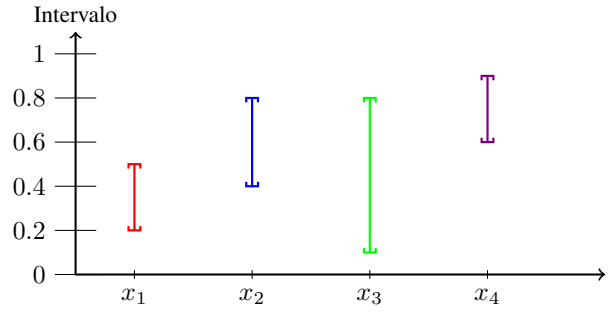


Fig. 1. Representación gráfica de los intervalos x_1, x_2, x_3 y x_4.

Podemos visualizar la proximidad entre x_1 y cada uno de los otros intervalos. Nuestra intuición dice que la similitud entre x_1 y x_3 debería ser mayor que la similitud entre x_1 y x_2, y esta mayor que la similitud entre x_1 y x_4.

Si consideramos la medida de incrustación basada en el ancho de los intervalos dada en la ecuación 2, la tabla I muestra la intersección entre los intervalos y la tabla II muestra los valores de la medida de incrustación asociada.

Entonces, considerando como funciones de agregación el mínimo y la media aritmética, obtenemos los valores de similitud mostrados en la tabla III. Por lo tanto, los valores obtenidos para las medidas de similitud son los esperados: en ambos casos (mínimo y media aritmética), el mayor valor de similitud está asociado a x_1 y x_3.

TABLE I
INTERSECCIÓN ENTRE LOS INTERVALOS

$\cap$	$x_2 = [0.4, 0.8]$	$x_3 = [0.1, 0.8]$	$x_4 = [0.6, 0.9]$
$x_1 = [0.2, 0.5]$	$[0.4, 0.5]$	$[0.2, 0.5]$	$\emptyset$

Tal y como dijimos, esta definición de similitud es una forma fácil y natural de medir cuánto se parecen dos intervalos cuando tienen algunos elementos en común. Pero falla en capturar la idea de similitud cuando no tienen puntos en común. En el ejemplo anterior obtuvimos $S^{E_w}_{\min}(x_1, x_4) = 0$ y $S^{E_w}_{AM}(x_1, x_4) = 0$. Habríamos obtenido el mismo resultado si

TABLE II
VALORES DE LA MEDIDA DE INCRUSTACIÓN ENTRE LOS INTERVALOS

$E_w(x_1, x_2) = 1/3$	$E_w(x_2, x_1) = 1/4$
$E_w(x_1, x_3) = 1$	$E_w(x_3, x_1) = 3/7$
$E_w(x_1, x_4) = 0$	$E_w(x_4, x_1) = 0$

TABLE III
VALORES DE LA SIMILITUD BASADA EN EL MÍNIMO Y EN LA MEDIA ARITMÉTICA SEGÚN LA INCRUSTACIÓN

Intervals	$S^{E_w}_{\min}$	$S^{E_w}_{AM}$
(x_1, x_2)	$1/4$	$7/24$
(x_1, x_3)	$3/7$	$5/7$
(x_1, x_4)	0	0

$x_4 = [0.5+\epsilon, 0.9]$ para cualquier $0 < \epsilon < 0.1$. Es decir, incluso considerando intervalos que estén tan cerca como se desee, la medida de similitud es 0 siempre que no tengan intersección. En una situación extrema, la similitud entre $[0.5 - \epsilon, 0.5 - \frac{\epsilon}{2}]$ y $[0.5 + \frac{\epsilon}{2}, 0.5 + \epsilon]$ es la misma que la similitud entre $[0, \epsilon]$ y $[1 - \epsilon, 1]$, sin importar lo pequeño que sea ϵ. La figura 2 muestra gráficamente un ejemplo de la situación mencionada.

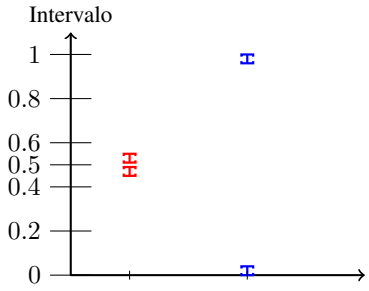


Fig. 2. Representación de los intervalos $[0.5-\epsilon, 0.5-\frac{\epsilon}{2}]$ y $[0.5+\frac{\epsilon}{2}, 0.5+\epsilon]$ en rojo y $[0, \epsilon]$ y $[1-\epsilon, 1]$ en azul.

A pesar de que esta desventaja también es compartida por otras medidas de similitud importantes y ampliamente utilizadas, como las similitudes de Jaccard o Dice, consideramos que es una situación contraintuitiva y creemos que debería ser corregida.

Nuestro objetivo es proponer una medida de similitud que pueda abordar la desventaja mencionada anteriormente. Nuestra propuesta es involucrar algún tipo de unión de los intervalos que estamos comparando para capturar su proximidad en caso de que no tengan intersección. Sin embargo, la unión clásica presenta una desventaja importante: no es un intervalo cuando los intervalos de partida son disjuntos. Por esta razón, no consideramos la unión clásica de intervalos, sino la siguiente:

***Definición** 3.3:* Dados dos intervalos $a, b \in L([0,1])$, la operación $a \sqcup b$ se define como el menor elemento en $L([0,1])$

tal que $a \subseteq a \sqcup b$ y $b \subseteq a \sqcup b$. Es decir, si $a = [\underline{a}, \overline{a}]$ y $b = [\underline{b}, \overline{b}]$, se tiene que:

$$a \sqcup b = \left[min\left\{\underline{a}, \underline{b}\right\}, max\{\overline{a}, \overline{b}\}\right] \quad (16)$$

Esta operación se convierte en la unión cuando los intervalos no son disjuntos. Sin embargo, no puede considerarse una unión de manera clásica, ya que incluye elementos que no están en ninguno de los intervalos originales si son disjuntos. Dados los intervalos $a = [0.2, 0.3]$ y $b = [0.4, 0.5]$, el intervalo $a \sqcup b$ es $[0.2, 0.5]$. Incluye, por ejemplo, el valor 0.35 que no está en ninguno de los conjuntos originales. A pesar de que no es una unión clásica, sirve para nuestros propósitos. Permite proporcionar una medida alternativa de proximidad entre intervalos que satisface la definición de similitud.

***Definición** 3.4:* Sea E una medida de incrustación para intervalos y $\mathcal{M}$ una función de agregación 1-estricta y simétrica. La función $S^E_{\mathcal{M}-\sqcup} : L([0,1])^2 \to [0,1]$, definida como:

$$S^E_{\mathcal{M}-\sqcup}(a, b) = \mathcal{M}(E(a \sqcup b, a), E(a \sqcup b, b)) \quad (17)$$

se denomina similitud basada en la función de agregación según la $\sqcup$-incrustación.

La siguiente proposición muestra que es una medida de similitud en el sentido de la definición 2.6.

***Proposición** 3.3:* Sea E una medida de incrustación para intervalos y $\mathcal{M}$ una función de agregación simétrica y 1-estricta. El operador $S^E_{\mathcal{M}-\sqcup}$ dado en la definición 3.4 es una medida de similitud para intervalos.

A continuación, mostramos que si consideramos la incrustación basada en el ancho del intervalo dada por la ecuación 2, la similitud introducida en la definición 3.4 supera el problema descrito anteriormente, ya que es monótona con respecto a la distancia entre los intervalos: cuanto más cercanos sean los intervalos considerados, mayor será el valor de la medida de similitud.

***Proposición** 3.4:* Sea E_w la medida de incrustación basada en el ancho del intervalo dada por la ecuación 2 y $\mathcal{M}$ una función de agregación simétrica y 1-estricta. El operador $S^{E_w}_{\mathcal{M}-\sqcup}$ dado en la definición 3.4 es monótono con respecto a la distancia entre los intervalos de igual ancho.

Por último, proporcionamos un ejemplo del resultado anterior:

***Ejemplo** 3.2:* Consideremos la medida de incrustación basada en el ancho del intervalo dada en la ecuación 2 y $\mathcal{M} = min$ (el resultado será el mismo para cualquier función de agregación simétrica y 1-estricta que sea idempotente).

Sean $a_\alpha = [\alpha, \alpha + 0.1]$ y $b_\alpha = [0.9 - \alpha, 1 - \alpha]$ donde $\alpha = \{0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35\}$.

La tabla IV muestra los valores de la similitud $S^{E_w}_{\mathcal{M}-\sqcup}$ entre a_α y b_α para los diferentes valores de α. La figura 3 ilustra gráficamente el comportamiento de la medida de similitud $S^E_{\mathcal{M}-\sqcup}$ cuando la distancia entre los intervalos varía.

TABLE IV
VALORES DE LA SIMILITUD ENTRE a_α Y b_α

α	a_α	b_α	$S(a_\alpha, b_\alpha)$
0	$[0, 0.1]$	$[0.9, 1]$	0.1
0.05	$[0.05, 0.15]$	$[0.85, 0.95]$	0.11
0.1	$[0.1, 0.2]$	$[0.8, 0.9]$	0.12
0.15	$[0.15, 0.25]$	$[0.75, 0.85]$	0.14
0.2	$[0.2, 0.3]$	$[0.7, 0.8]$	0.17
0.25	$[0.25, 0.35]$	$[0.65, 0.75]$	0.2
0.3	$[0.2, 0.4]$	$[0.6, 0.7]$	0.25
0.35	$[0.35, 0.45]$	$[0.55, 0.65]$	0.33

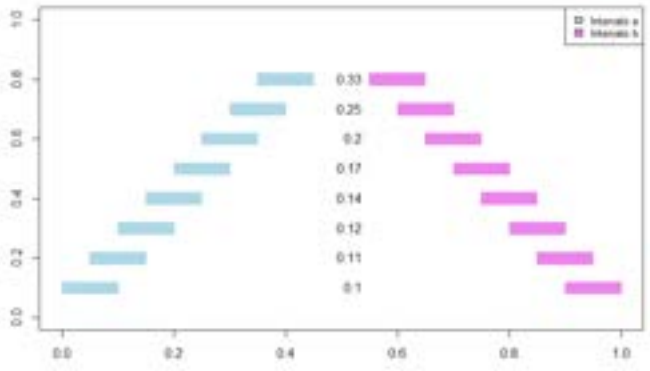


Fig. 3. Similitud entre los intervalos según su distancia

IV. CONCLUSIONES

En este artículo se han presentado similitudes basadas en medidas de incrustación. Se identificaron dos tipos distintos de similitudes. La primera, aunque cumple con las condiciones para ser considerada como tal, presenta una limitación significativa: no logra distinguir adecuadamente la proximidad entre dos intervalos cuando su intersección es vacía. Esta limitación se manifiesta en el hecho de que la similitud siempre da como resultado cero cuando los intervalos tienen una intersección vacía. Este inconveniente también lo presentan similitudes ampliamente conocidas y utilizadas en la literatura, como Jaccard y Dice. Sin embargo, consideramos que es una desventaja importante en el sentido de que no capturan la proximidad real entre los intervalos.

En respuesta a esta limitación, hemos propuesto una segunda caracterización para las similitudes basadas en medidas de incrustación que aborda esta deficiencia. Estas nuevas similitudes son capaces de distinguir con mayor precisión la proximidad entre intervalos, incluso cuando no tienen intersección en común. Al asignar valores no nulos que aumentan a medida que los intervalos se acercan, estas similitudes

proporcionan una medida más robusta y precisa de la relación entre intervalos.

Como trabajo futuro, proponemos explorar las propiedades que esta nueva definición satisface. Especialmente aquellas relacionadas con su comportamiento cuando se modifica la distancia entre los intervalos.

REFERENCIAS

[1] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Springer, 2007.

[2] G. Beliakov, H. Bustince Sola, and T. Calvo. *A Practical Guide to Averaging Functions*. Springer, 2016.

[3] A. Bouchet, M. Sesma-Sara, G. Ochoa, H. Bustince, S. Montes, and I. Díaz. Measures of embedding for interval-valued fuzzy sets. *Fuzzy Sets and Systems*, 467:108505, 2023.

[4] H. Bustince, C. Marco-Detchart, J. Fernandez, C. Wagner, J.M. Garibaldi, and Z. Takáč. Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders. *Fuzzy Sets and Systems*, 390:23–47, 2020.

[5] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[6] B. Gohain, R. Chutia, and P. Dutta. A distance measure for optimistic viewpoint of the information in interval-valued intuitionistic fuzzy sets and its applications. *Engineering Applications of Artificial Intelligence*, 119:105747, 2023.

[7] D.S. Guru, B.B. Kiranagi, and P. Nagabhushan. Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns. *Pattern Recognition Letters*, 25(10):1203–1213, 2004.

[8] P. Jaccard. étude comparative de la distribuition florale dans une portion des alpes et des jura. *Bull. Soc. Vaud. Sci. Nat.*, 7:547–579, 1901.

[9] P. Jaccard. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.*, 44:223–270, 1908.

[10] S. Kabir, C. Wagner, T. Havens, D. Anderson, and U. Aickelin. Novel similarity measure for interval-valued data based on overlapping ratio. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017.

[11] S. Kabir, C. Wagner, T. C. Havens, and D. T. Anderson. A similarity measure based on bidirectional subsethood for intervals. *IEEE Transactions on Fuzzy Systems*, 28(11):2890–2904, 2020.

[12] E.P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Springer, Dordrecht, Netherland, 2000.

[13] Y.S. Lin, J.Y. Jiang, and S.J. Lee. A similarity measure for text classification and clustering. *IEEE transactions on knowledge and data engineering*, 26(7):1575–1590, 2013.

[14] G. Mayor and E. Trillas. On the representation of some aggregation functions. In *Proceedings of IEEE International Symposium on Multiple-Valued Logic*, pages 111–114. IEEE, 1986.

[15] Gashirai K Mbizvo, Kyle H Bennett, Colin R Simpson, Susan E Duncan, Richard FM Chin, and Andrew J Larner. Using critical success index or gilbert skill score as composite measures of positive predictive value and sensitivity in diagnostic accuracy studies: Weather forecasting informing epilepsy research. *Epilepsia*, 64(6):1466–1468, 2023.

[16] Hung Nguyen and Vladik Kreinovich. Computing degrees of subsethood and similarity for interval-valued fuzzy sets: Fast algorithms. 01 2008.

[17] S. Ontañón. An overview of distance and similarity functions for structured data. *Artificial Intelligence Review*, 53:5309–5351, 2020.

[18] B. Pekala, K. Dyczkowski, P. Grzegorzewski, and U. Bentkowska. Inclusion and similarity measures for interval-valued fuzzy sets based on aggregation and uncertainty assessment. *Information Sciences*, 547:1182–1200, 2021.

[19] A. Pradera, G. Beliakov, H. Bustince, and B. De Baets. A review of the relationships between implication, negation and aggregation functions from the point of view of material implication. *Information Sciences*, 329:357–380, 2016.

[20] N. Rico, P. Huidobro, A. Bouchet, and I. Díaz. Similarity measures for interval-valued fuzzy sets based on average embeddings and its application to hierarchical clustering. *Information Sciences*, 615:794–812, 2022.

[21] M. Saqlain, H. Garg, P. Kumam, and W. Kumam. Uncertainty and decision-making with multi-polar interval-valued neutrosophic hypersoft set: A distance, similarity measure and machine learning approach. *Alexandria Engineering Journal*, 84:323–332, 2023.

[22] S. Vilar, R. Harpaz, H. S. Chase, S. Costanzi, R. Rabadan, and C. Friedman. Facilitating adverse drug event detection in pharmacovigilance databases using molecular structure similarity: application to rhabdomyolysis. *Journal of the American Medical Informatics Association*, 18(Supplement_1):i73–i80, 2011.

[23] G.W. Wei. Some similarity measures for picture fuzzy sets and their applications. *Iranian Journal of Fuzzy Systems*, 15(1):77–89, 2018.

Algunas propiedades de promedios para datos intervalares

Sergio F. Alonso
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
fernandezalosergio@uniovi.es

Susana Diaz-Vazquez
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
diazsusana@uniovi.es

Susana Montes
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—La necesidad de herramientas que permitan resumir grandes cantidades de información existe desde muy antiguo. No obstante, los ordenadores y su capacidad para almacenar vastas cantidades de datos han hecho que esa necesidad sea aún más apremiante en los últimos años. Por otro lado, cuando se trabaja con información imprecisa, por ejemplo si se manejan opiniones o preferencias de personas, se suelen usar intervalos en lugar de números, ya que permiten formalizar el concepto de imprecisión de una forma más fiel. Este trabajo se centra en el estudio de las funciones de agregación de intervalos. Más concretamente, en el estudio de algunas funciones que generalizan a la (clásica) media aritmética. En la primera parte de la contribución se discute la definición de promedio y en la segunda, se demuestra que, dependiendo del orden fijado, el comportamiento de las generalizaciones de la media consideradas es diferente.**

***Index Terms*—Datos intervalares, orden para intervalos, agregación, promedio.**

I. Introducción

Cuando se tienen múltiples anotaciones/medidas/opiniones sobre un mismo elemento, es necesario disponer de algún procedimiento que nos permita sintetizar toda esa información y proporcionar un único resultado final que sea representativo del conjunto de datos del que se ha obtenido. Este procedimiento son las funciones de agregación [Grabisch et al., 2009], [Beliakov et al., 2007], [Beliakov et al., 2016]. Las funciones de agregación son fundamentales a la hora de resumir grandes cantidades de datos. Su utilidad ha tenido especial relevancia en las últimas décadas en las que la informatización de los datos ha permitido recopilar y almacenar ingentes cantidades de información sobre todo tipo de ámbitos.

El uso de intervalos en lugar de números reales para recoger cierta información es un método ampliamente usado ya que, sobre todo en el caso de modelar pensamiento humano, la incertidumbre asociada a esa información puede no quedar bien reflejada mediante un único valor. En muchas ocasiones también son intervalos los α-cortes que caracterizan a los números borrosos. De esta manera, la existencia de herramientas para operar con intervalos, permitiría definir operaciones entre números borrosos. Cuando se dispone de un volumen elevado de información, son necesarias técnicas que nos permitan resumir dicha información. También, si la información procede de diferentes fuentes, como por ejemplo cuando diferentes expertos expresan su opinión sobre un mismo producto, es necesario obtener una valoración promedio, de consenso o representativa de las distintas opiniones recibidas. Un axioma fundamental a la hora de definir una medida de agregación es la monotonía: si se agregan valores mayores, el resultado debe ser mayor. El concepto subyacente a la idea de monotonía es el de orden. Si no existe un orden entre los elementos a agregar, no se puede definir la monotonía. En el caso de los números reales existe un orden total (el orden natural entre números reales) que sirve de base para hablar de monotonía. Sin embargo, este no es el caso para intervalos. El principal escollo al que debemos enfrentarnos al definir función monótona para intervalos es que no disponemos de una única relación de orden total que se use de manera estandarizada. La relación de orden más habitual es el orden producto o reticular, que se define comparando los extremos del intervalo y que no es un orden total. Existen varios órdenes totales, pero ninguno que se use de manera generalizada. Por ello, en este trabajo vamos a considerar distintos tipos de órdenes totales y proponer definiciones de crecimiento y función promedio dependientes de la relación de orden elegida. Además, estudiaremos varios operadores intervalares respecto al orden reticular [Fishburn, 1987] y a los órdenes $\preceq_{\alpha+}$ y $\preceq_{\alpha-}$ definidos en [Bustince et al., 2013].

Los autores querrían agradecer la ayuda recibida del Ministerio de Ciencia e Innovación a través del proyecto PID2022-139886NB- I00.

II. Preliminares

En esta sección se fijan algunos conceptos y la notación que se usarán en las secciones siguientes.

Denotaremos por $L(\mathbb{R})$ al conjunto de los intervalos cerrados de números reales, esto es:

$$L(\mathbb{R}) = \{[a, b] \mid a, b \in \mathbb{R} \text{ y } a \leq b\}$$

y por $K(\mathbb{R})$ al conjunto de los intervalos abiertos de números reales. La notación $L(\mathbb{R}^+)$ hará referencia al conjunto de los intervalos cerrados de valores no negativos, esto es,

$$L(\mathbb{R}^+) = \{[a, b] \in L(\mathbb{R}) \mid 0 \leq a\}.$$

Salvo que se indique lo contrario, $[a, b]$ denotará cualquier intervalo cerrado de $\mathbb{R}$, es decir, a, b serán dos números reales cualesquiera satisfaciendo únicamente que $a \leq b$.

Comenzaremos recordando nociones relacionadas con la agregación de números reales y posteriormente, pasaremos a datos intervalares. En esta sección las funciones no estarán definidas en toda la recta real, sino en intervalos para trabajar con las condiciones de frontera impuestas en [Grabisch et al., 2009].

La necesidad de agregar o resumir información es muy antigua y la existencia de funciones como la media, también. En 1821, Cauchy [Cauchy, 1821] formalizó la idea de función de promedio. Simplemente exigía que fuese interna.

Definición 1. *Se dice que una función* $f : [a,b]^n \longrightarrow \mathbb{R}$ *es interna si cumple que* $Min\{x_1,\ldots,x_n\} \leq f(x_1,\ldots,x_n) \leq Max\{x_1,\ldots,x_n\}$ *para cualquier tupla* $(x_1,\ldots,x_n) \in [a,b]^n$.

Cauchy no exigía monotonía: la función $f : [0,1]^2 \to [0,1]$ definida como

$$f(x,y) = \begin{cases} Max(x,y) & \text{si } (x,y) \in [0,0.5]\times[0,0.5] \\ & \text{o } (x,y) \in [0.5,1]\times[0.5,1] \\ Min(x,y) & \text{en otro caso.} \end{cases}$$

es una función interna y por tanto, según Cauchy, sería un promedio.

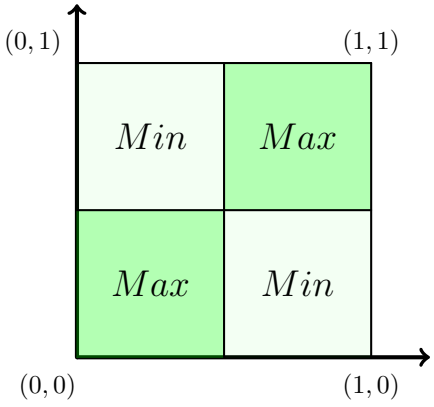


Fig. 1. Ejemplo de función de Cauchy.

Sin embargo, tal como se puede observar en su representación gráfica (ver figura 1), la monotonía no parece ser verificada. De hecho, el siguiente contraejemplo prueba que esta función no es creciente en la segunda componente: $f(0.2,0.45) = 0.45 > 0.2 = f(0.2,0.6)$.

No obstante, parece lógico requerir que cuanto mayores son los valores que agregamos, mayor sea el resultado de dicha agregación. Recordemos la definición formal de monotonía creciente.

Definición 2. *Se dice que una función* $f : [a,b]^n \longrightarrow \mathbb{R}$ *es monótona (creciente) si cumple que*

$$f(x_1,\ldots x_i,\ldots,x_n) \leq f(x_1,\ldots x_i',\ldots,x_n)$$

para todo $x_i \leq x_i'$ *con* i *cualquier valor en* $\{1,\ldots,n\}$.

A lo largo de esta contribución no trabajaremos con funciones decrecientes, por lo que hablaremos simplemente de función monótona/monotonía para referirnos a las funciones monótonas crecientes.

Puesto que, como hemos comentado, la monotonía de la función es una condición lógica a la hora de agregar la información, la definición que surgió posteriormente de función de agregación y que está ampliamente aceptada en la comunidad científica, impone como axioma este requerimiento (véase, por ejemplo, [Beliakov et al., 2007]):

Definición 3. *Se dice que una función* $f : [a,b]^n \to \mathbb{R}$ *es una función de agregación si:*

- *Monotonía: para todo* $x_i \leq x_i'$ *con* i *cualquier valor en* $\{1,\ldots,n\}$ *se tiene que* $f(x_1,\ldots x_i,\ldots,x_n) \leq f(x_1,\ldots x_i',\ldots,x_n)$.
- *Condiciones frontera:* $f(a,\ldots,a) = a$ *y* $f(b,\ldots,b) = b$.

Es evidente que las condiciones anteriores siguen siendo muy generales y, por tanto, verificadas por un amplio números de familias de funciones. En concreto, las t-normas o las t-conormas son funciones de agregación, pero no verifican otra condición lógica que es la que impuso Cauchy: no necesariamente son internas.

En su libro, [Grabisch et al., 2009] definen función promedio a partir de las ideas de función idempotenciable y de sección diagonal, que se recogen a continuación:

Definición 4. *Se llama sección diagonal de una función* $f : [a,b]^n \longrightarrow \mathbb{R}$ *a la función* $\delta_f : [a,b] \longrightarrow \mathbb{R}$ *dada por* $\delta_f(x) = f(x,\ldots,x),\ \forall x \in [a,b]$.

Definición 5. *Una función* $f : [a,b]^n \longrightarrow \overline{\mathbb{R}}$ *es idempotenciable si su sección diagonal* δ_f *es estrictamente creciente y satisface que* $ran(\delta_f) = ran(f)$*, donde ran se refiere a la imagen de la función.*

La idea de función idempotenciable se obtiene de debilitar la noción de función idempotente, que es central en el estudio de las funciones promedio. Recordamos primero la definición de función idempotente y luego vemos que existen funciones idempotenciables que no son idempotentes.

Definición 6. *Una función* $f : [a,b]^n \longrightarrow [a,b]$ *se dice idempotente si verifica que* $f(x,\ldots,x) = x,\ \forall x \in [a,b]$.

La definición anterior puede considerarse de igual modo para funciones definidas de $\mathbb{R}^n$ en $\mathbb{R}$.

Veamos ahora que no toda función idempotenciable es idempotente. La función $f : [0,1]^n \to [0,1]$ definida como

$$f(x_1,\ldots,x_n) = \sqrt{Min_i\{x_i\}}$$

no es idempotente, pero sí idempotenciable porque su sección diagonal es $\delta_f(x) = \sqrt{x}$. Esta función es estrictamente creciente y $ran(\delta_f) = ran(f) = [0,1]$.

Podemos presentar, entonces, la definición de función promedio de [Grabisch et al., 2009].

Definición 7. *Una función* $M : [a,b]^n \longrightarrow [a,b]$ *se llama promedio en* $[a,b]^n$ *si existe una función creciente e idempotenciable* $f : [a,b]^n \longrightarrow \mathbb{R}$ *tal que* $f = \delta_f \circ M$*. Se dice, entonces, que M es un promedio asociado con f en* $[a,b]^n$.

Aunque las definiciones de [Grabisch et al., 2009] y Cauchy parezcan alejadas, el siguiente resultado muestra la conexión.

Proposición 1. *Sea* $f : [a,b]^n \longrightarrow [a,b]$. *Las siguientes condiciones son equivalentes:*

i. *f es idempotente y creciente.*
ii. *f es interna y creciente.*
iii. *f es una función promedio.*

En la proposición anterior se puede ver que la monotonía juega un papel central en la formalización de la idea de promedio. A su vez, se pone de manifiesto que un concepto básico en la formalización de la monotonía es el de orden.

Definición 8. *Dado un conjunto* X, *una relación* $\leq$ *en* X *es un orden si es reflexiva, antisimétrica y transitiva.*

Es evidente que pueden existir para una relación de orden elementos incomparables. En ese caso se dice que el orden es parcial. En caso contrario que es total. Más en concreto,

Definición 9. *Un orden* $\leq$ *en* X *se dice total si para todo par de elementos* $x, y \in X$ *se tiene que* $x \leq y$ *o* $y \leq x$.

El orden natural entre números reales es un orden total.

Cuando se trata con intervalos, surgen problemas relativos al orden. Para empezar, hay diferentes propuestas de orden ya que ninguno está universalmente admitido como "natural" o "definitivo" y para continuar, el que se considera más intuitivo y por tanto el que quizá se emplea más habitualmente en la literatura, no es total, es decir, no permite comparar todos los pares de intervalos. Algunos de los órdenes más habituales para intervalos son:

Definición 10. *Dados* $[a_1,b_1],[a_2,b_2] \in L(\mathbb{R})$*:*

- *El orden reticular [Fishburn, 1987] también llamado orden producto [Neggers and Kim, 1998], orden por coordenadas [Davey and Priestley, 2002] o incluso orden por componentes [Taylor, 1999] en* $L(\mathbb{R})$ *se define como:*

$$[a_1,b_1] \preceq_{LO} [a_2,b_2] \Leftrightarrow a_1 \leq a_2 \ y \ b_1 \leq b_2.$$

  Las letras LO hacen referencia al nombre en inglés: "Lattice Order".
- *El orden lexicográfico de tipo 1 compara los extremos inferiores de los intervalos y si estos son iguales, compara los extremos superiores. Formalmente se define como:*

$$[a_1,b_1] \preceq_{Lex1} [a_2,b_2] \Leftrightarrow a_1 < a_2 \ o \ (a_1 = a_2 \ y \ b_1 \leq b_2).$$

- *El orden lexicográfico de tipo 2 invierte el orden con respecto al orden lexicográfico de tipo 1: compara primero los extremos superiores y en caso de igualdad, compara los extremos inferiores. Formalmente:*

$$[a_1,b_1] \preceq_{Lex2} [a_2,b_2] \Leftrightarrow b_1 < b_2 \ o \ (b_1 = b_2 \ y \ a_1 \leq a_2)$$

- *El orden de Xu y Yager [Xu and Yager, 2006] en* $L(\mathbb{R})$ *compara primero los puntos medios de los intervalos y si estos son iguales, compara las amplitudes de los intervalos. Formalmente se define como:*

$$[a_1,b_1] \preceq_{XY} [a_2,b_2]$$
$$\Updownarrow$$
$$\frac{a_1+b_1}{2} < \frac{a_2+b_2}{2}$$
$$o$$
$$\frac{a_1+b_1}{2} = \frac{a_2+b_2}{2} \quad y \quad \frac{b_1-a_1}{2} \leq \frac{b_2-a_2}{2}$$

El orden más extendido a la hora de comparar intervalos es el orden reticular. Sin embargo, este orden presenta un importante inconveniente y es que no es total, es decir, no permite comparar todos los intervalos. Si consideramos por ejemplo, los intervalo $[2,5]$ y $[3,4]$, se tiene que $[2,5] \not\preceq_{LO} [3,4]$ y $[3,4] \not\preceq_{LO} [2,5]$. Dicho de otro modo, los intervalos $[2,5]$ y $[3,4]$ son incomparables de acuerdo con el orden reticular. Sin embargo, se trata de un orden intuitivo. Por este motivo, cuando hace falta que el orden con el que trabajamos sea total, se consideran otros órdenes, pero que respetan la forma de comparar del orden reticular.

Definición 11. *[Bustince et al., 2013] Un orden admisible,* $\preceq_{ao}$*, es un orden total que generaliza al orden reticular. Esto es, si* $[a_1,b_1] \preceq_{LO} [a_2,b_2]$*, entonces se da también que* $[a_1,b_1] \preceq_{ao} [a_2,b_2]$.

Los órdenes lexicográficos de tipo 1 y 2 y el orden de Xu y Yager son órdenes admisibles.

En el mismo trabajo en el que se da la definición anterior, se propone una forma de generar órdenes admisibles, a los que llaman órdenes admisibles generados, a partir de funciones de agregación, del siguiente modo:

Definición 12. *[Bustince et al., 2013] Se dice que un orden admisible* $\preceq_{ao}$ *es un orden admisible generado si existen dos funciones* $f, g : K(\mathbb{R}) \longrightarrow \mathbb{R}$ *tales que para todo par de intervalos* $[a_1,b_1],\ [a_2,b_2] \in L(\mathbb{R})$ *se cumple que*

$$[a_1,b_1] \preceq_{ao} [a_2,b_2]$$
$$\Updownarrow$$
$$[f(a_1,b_1), g(a_1,b_1)] \preceq_{Lex1} [f(a_2,b_2), g(a_2,b_2)]$$

Es evidente que, tomando f como la proyección en la primera componente ($f(a,b) = a$) y g como la proyección en la segunda componente ($g(a,b) = b$), se obtiene el orden lexicográfico de tipo 1, es decir el orden lexicográfico de tipo 1 es un orden admisible generado. También lo son el orden lexicográfico de tipo 2 y el de Xu y Yager. Para obtener el orden lexicográfico de tipo 2 basta considerar f como la proyección en la segunda componente y g la proyección en la primera componente. Tomando f como la media, $f(a,b) = \frac{a+b}{2}$ y $g(a,b) = b - a$, se obtiene el orden de Xu y Yager.

Las funciones que permiten escribir los órdenes lexicográficos y el orden de Xu y Yager como generados son casos particulares de una familia más general, las funciones $\{K_\alpha\}_{\alpha \in [0,1]}$ definidas por Atanassov en 1983 [Atanassov,

1983], como $K_\alpha(a,b) = (1-\alpha)a + \alpha b$. Si tomamos K_α y K_β con $\alpha \neq \beta$, se obtiene el orden admisible generado definido, para dos intervalos $I_1, I_2 \in L(\mathbb{R})$, por

$$\begin{array}{c} I_1 \preceq_{\alpha,\beta} I_2 \\ \Updownarrow \\ k_\alpha(I_1) < k_\alpha(I_2) \\ \text{o} \\ k_\alpha(I_1) = k_\alpha(I_2) \quad \text{y} \quad k_\beta(I_1) < k_\beta(I_2). \end{array}$$

En el mismo trabajo [Bustince et al., 2013] tambien prueban que para cualquier $\alpha \in [0,1)$, todos los órdenes $\preceq_{\alpha\beta}$ con $\beta > \alpha$ coinciden y reciben el nombre de $\preceq_{\alpha+}$.

Análogamente, para cualquier $\alpha \in (0,1]$, todos los órdenes $\preceq_{\alpha\beta}$ con $\beta < \alpha$ coinciden y se denotan por $\preceq_{\alpha-}$.

III. Funciones promedio para datos intervalares

Como se comentó brevemente en la sección anterior, la idea de promedio de números reales se aproximó en la literatura a través de conceptos que a primera vista pueden parecer dispares, pero que revisados con calma acaban siendo equivalentes. En este trabajo consideraremos la siguiente definición:

Definición 13. *Una función $f : [a,b]^n \to \mathbb{R}$ es una función promedio si es creciente e idempotente.*

A. Función $\preceq$-creciente y $\preceq$-promedio. Definiciones.

El primer objetivo de esta contribución es dar una definición de función promedio para datos intervalares. Atendiendo a los números reales, se ve que la noción de función promedio está íntimamente ligada a la idea de crecimiento, por lo que necesitamos tener una definición de crecimiento para funciones definidas en $L(\mathbb{R})$. Además, la definición de crecimiento está asociada a un orden. En el caso de los números reales es el orden natural. En el caso de los intervalos, como ya se comentó en la sección previa, dado que no existe una única relación de orden comúnmente utilizada para $L(\mathbb{R})$, la definición de monotonía tampoco es única, depende del orden considerado.

Definición 14. *Dada una relación de orden $\preceq$ en $L(\mathbb{R})$. Se dice que una función $f : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es $\preceq$-creciente cuando cumple que si $[a_i,b_i] \preceq [a'_i,b'_i]$, entonces*

$$f([a_1,b_1],\ldots,[a_i,b_i],\ldots,[a_n,b_n]) \preceq$$
$$f([a_1,b_1],\ldots,[a'_i,b'_i],\ldots,[a_n,b_n])$$

para cualquier $i \in \{1,\ldots,n\}$.

Una vez fijada una noción de crecimiento, estableciendo un paralelismo con las definiciones sobre números reales, generalizamos la definición de función idempotente de manera directa como sigue:

Definición 15. *Una función $f : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es idempotente cuando cumple que*

$$f([a,b],\ldots,[a,b]) = [a,b],$$

para todo $[a,b] \in L(\mathbb{R})$.

Una vez generalizados ambos conceptos, la definición de promedio se extiende fácilmente, teniendo en cuenta, eso sí, que no será única, ya que hereda del concepto de monotonía la dependencia del orden considerado.

Definición 16. *Decimos que una función $f : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es una función $\preceq$-promedio, o simplemente un $\preceq$-promedio, si es*

i) *idempotente y*
ii) *$\preceq$-creciente.*

En el caso de agregar números reales, hemos visto que se hicieron propuestas (a priori) diferentes y que luego se demostró (proposición 1) que son equivalentes. Una de las primeras dudas que nos surge es esa: ¿se puede extender la proposición 1 al caso de intervalos? Para responder a esta pregunta deberíamos tener una definición de función interna para intervalos, pero esa definición depende de los conceptos de mínimo y máximo de un conjunto de intervalos y bajo estos, vuelve a subyacer la idea de orden (entre intervalos).
En un primer momento se puede pensar que tendremos una definición diferente de mínimo y máximo para cada orden considerado, como sucedía con la definición de función monótona. Lamentablemente, en este caso no se puede trabajar con cualquier orden, este debe ser total.

Definición 17. *Dado un orden total $\preceq$, definimos $Min_\preceq : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ como*

$$Min_\preceq([a_1,b_1],\ldots,[a_n,b_n]) = [a_j,b_j],$$

con $j \in \{1,\ldots,n\}$ tal que $[a_j,b_j] \preceq [a_i,b_i]$, $\forall i \in \{1,\ldots,n\}$.

Definición 18. *Dado un orden total $\preceq$, definimos $Max_\preceq : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ como*

$$Max_\preceq([a_1,b_1],\ldots,[a_n,b_n]) = [a_j,b_j],$$

con $j \in \{1,\ldots,n\}$ tal que $[a_i,b_i] \preceq [a_j,b_j]$, $\forall i \in \{1,\ldots,n\}$.

Si el orden no es total, no se garantiza la existencia del mínimo ni del máximo de un conjunto de intervalos como se ve, por ejemplo, en el orden reticular tomando dos conjuntos $\preceq_{LO}$-incomparables, $[2,5]$ y $[3,4]$. No tiene sentido hablar de mínimo (máximo) de esos dos intervalos según $\preceq_{LO}$ cuando son incomparables según el mismo.

Si nos restringimos a órdenes totales, una vez definidos el mínimo y el máximo de un conjunto de intervalos, la noción de función interna se sigue de forma directa.

Definición 19. *Dado un orden total $\preceq$ decimos que $f : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es $\preceq$-interna siempre que verifique que*

$$Min_\preceq([a_1,b_1],\ldots,[a_n,b_n]) \preceq$$
$$f([a_1,b_1],\ldots,[a_n,b_n]) \preceq$$
$$Max_\preceq([a_1,b_1],\ldots,[a_n,b_n]).$$

La noción de $\preceq$-promedio es más general que la de función $\preceq$-interna porque se puede dar para cualquier orden, no necesariamente total. Sin embargo, cuando nos restringimos

a órdenes totales, sí se tiene la equivalencia entre función idempotente e interna.

Proposición 2. *Sean $\preceq$ un orden total y $f : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ una función $\preceq$-creciente, entonces son equivalentes:*

i) *f es $\preceq$-interna.*
ii) *f es idempotente.*

Además de considerar los operadores $Min_{\preceq}$ y $Max_{\preceq}$ para los cuales es necesario que el orden subyacente sea total, podrían considerarse los operadores $Inf_{\preceq}$ y $Sup_{\preceq}$ para los que bastaría que el orden $\preceq$ induzca una estructura de retículo sobre $L(\mathbb{R})$ y se definirían del modo siguiente.

Definición 20. *Si $\preceq$ es un orden que induce una estructura reticular en $L(\mathbb{R})$, definimos $Inf_{\preceq} : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ como $Inf_{\preceq}([a_1, b_1], \ldots, [a_n, b_n]) = [a_0, b_0]$, tal que $[a_0, b_0] \preceq [a_i, b_i]$ $\forall i \in \{1, \ldots, n\}$ y que si $[a, b] \preceq [a_i, b_i]$, $\forall i \in \{1, \ldots, n\}$, entonces también $[a, b] \preceq [a_0, b_0]$.*

Definición 21. *Si $\preceq$ es un orden que induce una estructura reticular en $L(\mathbb{R})$, definimos $Sup_{\preceq} : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ como $Sup_{\preceq}([a_1, b_1], \ldots, [a_n, b_n]) = [a_0, b_0]$, tal que $[a_0, b_0] \succeq [a_i, b_i]$ $\forall i \in \{1, \ldots, n\}$ y que si $[a, b] \succeq [a_i, b_i]$, $\forall i \in \{1, \ldots, n\}$, entonces también $[a, b] \succeq [a_0, b_0]$.*

Obviamente, cuando el orden considerado es total, el ínfimo de un conjunto de intervalos coincide con el mínimo y el supremo, con el máximo.

Si se considera el orden reticular, hemos obtenido expresiones explícitas para el ínfimo y el supremo:

Proposición 3. *Dado el orden reticular $\preceq_{LO}$, el operador Inf_{LO} puede expresarse como*

$$Inf_{LO}([a_1, b_1], \ldots, [a_n, b_n]) = [Min_i\{a_i\}, Min_i\{b_i\}].$$

Nótese que el intervalo resultante no necesariamente es un elemento del conjunto de intervalos de partida. Así,

$$Inf_{LO}([2, 5], [3, 4]) = [2, 4],$$

que no es ninguno de los intervalos a agregar.

Análogamente se puede probar que

Proposición 4. *Dado el orden reticular, el operador Sup_{LO} puede expresarse como*

$$Sup_{LO}([a_1, b_1], \ldots, [a_n, b_n]) = [Max_i\{a_i\}, Max_i\{b_i\}].$$

Tampoco en este caso el resultado es, necesariamente, un elemento del conjunto de intervalos a agregar:

$$Sup_{LO}([2, 5], [3, 4]) = [3, 5].$$

B. Funciones intervalo valoradas, crecimiento y promedio respecto a un orden

Una vez formalizado el concepto de promedio para intervalos, estudiamos qué operadores de agregación para intervalos verifican la definición vista.

En primer lugar, podemos comprobar que el operador Inf_{LO}, caracterizado en la proposición 3 es un $\preceq_{LO}$-promedio, lo cual es muy natural.

Proposición 5. *El operador $Inf_{LO} : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es un $\preceq_{LO}$-promedio.*

Sin embargo, no es un promedio cuando se consideran algunos de los órdenes totales más habituales. El problema radica en que no se tiene garantizada la monotonía.

Contraejemplo 6. *El operador Inf_{LO} no es $\preceq_{\alpha+}$-creciente para ningún $\alpha \in [0, 1)$ ni $\preceq_{\alpha-}$-creciente para ningún $\alpha \in (0, 1]$.*

Este resultado indica que Inf_{LO} no es monótona para los órdenes lexicográficos ni para el de Xu y Yager. Recordemos que en la familia $\preceq_{\alpha+}$ están el orden lexicográfico de tipo 1 ($\alpha = 0$) y el orden de Xu y Yager ($\alpha = 0.5$). La familia $\preceq_{\alpha-}$ incluye al orden lexicográfico de tipo 2 ($\alpha = 1$).

Ambos resultados se pueden trasladar al supremo. Es un promedio para el orden reticular.

Proposición 7. *El operador $Sup_{LO} : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es un $\preceq_{LO}$-promedio.*

Pero no lo es para los órdenes del tipo $\preceq_{\alpha+}$ o $\preceq_{\alpha-}$.

Contraejemplo 8. *El operador Sup_{LO} no es $\preceq_{\alpha+}$-creciente para ningún $\alpha \in [0, 1)$ ni $\preceq_{\alpha-}$-creciente para ningún $\alpha \in (0, 1]$.*

Aunque otros operadores podrían tener más interés en el ámbito de los números borrosos, el operador de agregación por antonomasia es la media aritmética. Por eso en este trabajo preliminar hemos comenzado por ella. Vamos a formalizar la idea de media de un conjunto de intervalos y después estudiar si es un $\preceq$-promedio para los órdenes más habituales.

Definición 22. *Dada una colección de intervalos $[a_1, b_1], \ldots, [a_n, b_n]$, su media se define como*

$$Mean([a_1, b_1], \ldots, [a_n, b_n]) = \left[\tfrac{a_1 + \cdots + a_n}{n}, \tfrac{b_1 + \cdots + b_n}{n}\right].$$

Como cabía esperar, esta media presenta un buen comportamiento con respecto a los órdenes más habituales. Es un promedio respecto al orden reticular.

Proposición 9. *El operador $Mean : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es un $\preceq_{LO}$-promedio.*

Y también es un promedio respecto a los órdenes admisibles generados más comunes:

Proposición 10. *El operador $Mean : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ es:*

- *un $\preceq_{\alpha+}$-promedio para cualquier $\alpha \in [0, 1)$,*
- *un $\preceq_{\alpha-}$-promedio para cualquier $\alpha \in (0, 1]$.*

De modo que en particular, es un promedio con respecto a los órdenes lexicográfico de tipo 1 y 2 y al de Xu y Yager.

De igual modo que, sobre números reales se pueden definir medias generalizadas, en este trabajo consideramos la idea de media generalizada para intervalos. Para formalizar este concepto necesitamos definir previamente las potencias de un intervalo.

Definición 23. *Dado* $[a,b] \in L(\mathbb{R})$*, y* $n \in \mathbb{N}$*, definimos la enésima potencia del intervalo* $[a,b]$ *como*

$$[a,b]^n = \begin{cases} [0, \max\{a^n, b^n\}] & \text{si } 0 \in [a,b] \\ [\min\{a^n, b^n\}, \max\{a^n, b^n\}] & \text{en otro caso} \end{cases}$$

si n *es un número par, y como* $[a,b]^n = [a^n, b^n]$*, si* n *es un número impar.*

También seguimos la misma idea que en números reales para definir la raíz de intervalos.

Definición 24. *Dado un intervalo* $[a,b]$ *con* $a \geq 0$ *y* $n \geq 2$ *y par, se define la raíz enésima de* $[a,b]$ *como* $\sqrt[n]{[a,b]} = [\sqrt[n]{a}, \sqrt[n]{b}]$.

Sin embargo, si el índice es impar, la raíz se puede definir para cualquier intervalo, tal como se muestra a continuación.

Definición 25. *Dado un intervalo* $[a,b] \in L(\mathbb{R})$ *y* $n \geq 3$ *e impar, se define la raíz enésima de* $[a,b]$ *como* $\sqrt[n]{[a,b]} = [\sqrt[n]{a}, \sqrt[n]{b}]$.

Formalizados los conceptos de potencia y raíz de un intervalo, ya podemos definir la media generalizada de orden m en el contexto de datos intervalares.

Definición 26. *Llamamos media generalizada de orden* m*, y lo denotamos por* mPM *a la función* $mPM : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ *definida por*

$$mPM(I_1, \ldots, I_n) = \frac{\sqrt[m]{I_1^m + \cdots + I_n^m}}{n}.$$

La definición anterior permite una caracterización en términos de medias generalizadas de orden m de ciertos elementos del intervalo.

Proposición 11. *Dada una colección de intervalos* $\{I_i = [a_i, b_i]\}_{i=1}^n \subset L(\mathbb{R})$*, su media generalizada de orden* m *puede escribirse como:*

$$mPM(I_1, \ldots, I_n) = \left[\sqrt[m]{\frac{\sum_{i=1}^n Min_{x_i \in [a_i,b_i]}\{x_i^m\}}{n}}, \sqrt[m]{\frac{\sum_{i=1}^n Max\{a_i^m, b_i^m\}}{n}} \right].$$

Cabe notar que esta caracterización requiere de la expresión anterior para el caso de órdenes pares. En el caso de impares, se obtiene una expresión más sencilla e intuitiva: para m impar los extremos inferior y superior de la media generalizada de orden m son (respectivamente) las medias generalizadas de orden m de los extremos inferiores y superiores de los intervalos a agregar. Esto simplifica la demostración del siguiente resultado.

Proposición 12. *El operador* $mPM : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ *es un* $\preceq_{LO}$*-promedio para todo* $m \geq 3$ *impar.*

Sorprendentemente, el resultado no se cumple para ningún m par ya que en este caso no se puede garantizar la monotonía.

Contraejemplo 13. *El operador* $mPM : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ *no es* $\preceq_{LO}$*-creciente para ningún* $m \geq 2$ *par.*

Este mal comportamiento se soluciona si nos restringimos a intervalos de números no negativos.

Proposición 14. *El operador* $mPM : L(\mathbb{R}^+)^n \longrightarrow L(\mathbb{R})$ *es un* $\preceq_{LO}$*-promedio para todo* $m \geq 2$.

En el conjunto de los intervalos positivos, el buen comportamiento se extiende a los órdenes lexicográficos 1 y 2.

Proposición 15. *El operador* $mPM : L(\mathbb{R}^+)^n \longrightarrow L(\mathbb{R})$ *es*

- *un* $\preceq_{Lex1}$*-promedio para todo* $m \geq 2$*,*
- *un* $\preceq_{Lex2}$*-promedio para todo* $m \geq 2$.

Sin embargo, no es un promedio para ningún otro orden de la familia, incluso restringiendo el dominio a los intervalos positivos.

Contraejemplo 16. *El operador* $mPM : L(\mathbb{R})^n \longrightarrow L(\mathbb{R})$ *no es* $\preceq_{\alpha+}$*-creciente ni* $\preceq_{\alpha-}$*-creciente para ningún* $\alpha \in (0,1)$ *para* $m \geq 2$.

IV. Conclusiones

Este es un trabajo sobre el estudio de funciones promedio para datos intervalares. En él hemos formalizado la idea de función promedio y hemos visto que esta debe depender del orden (entre intervalos) considerado. También hemos estudiado el comportamiento de algunas de las funciones de agregación más habituales: las medias aritméticas generalizadas. Los resultados obtenidos no siempre han sido satisfactorios, ya que estos operadores son promedios para los órdenes más clásicos, pero no para muchos órdenes admisibles generados.

Se trata de un trabajo preliminar. Queremos estudiar más funciones que permitan promediar intervalos, ver sus propiedades y caracterizaciones, así como explorar para qué órdenes verifican la definición aquí introducida de promedio.

Referencias

[Atanassov, 1983] Atanassov, K. (1983). Intuitionistic fuzzy sets. In *VIIth ITKR Session*, pages 1684–1697, Sofia,Bulgaria. Deposited in the Central Science and Technology Library of the Bulgarian Academy of Sciences.

[Beliakov et al., 2007] Beliakov, G., Pradera, A., Calvo, T., et al. (2007). *Aggregation functions: A guide for practitioners*, volume 221. Springer.

[Beliakov et al., 2016] Beliakov, G., Sola, H. B., and Sánchez, T. C. (2016). *A practical guide to averaging functions*, volume 329. Springer.

[Bustince et al., 2013] Bustince, H., Fernandez, J., Kolesárová, A., and Mesiar, R. (2013). Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems 86*, pages 69–77.

[Cauchy, 1821] Cauchy, A. L. B. (1821). *Cours d'analyse de l'École Royale Polytechnique*, volume 1. Imprimerie royale.

[Davey and Priestley, 2002] Davey, B. and Priestley, H. (2002). *Introduction to Lattices and Order*. Cambridge mathematical textbooks. Cambridge University Press.

[Fishburn, 1987] Fishburn, P. (1987). Interval ordenings. *Wiley, New York*.

[Grabisch et al., 2009] Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation Functions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press.

[Neggers and Kim, 1998] Neggers, J. and Kim, H. (1998). *Basic Posets*. G - Reference,Information and Interdisciplinary Subjects Series. World Scientific.

[Taylor, 1999] Taylor, P. (1999). *Practical Foundations of Mathematics*. Number v. 59 in Cambridge Studies in Advanced Mathematics. Cambridge University Press.

[Xu and Yager, 2006] Xu, Z. and Yager, R. R. (2006). Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems*, 35(4):417–433.

Using Principal Component Analysis for Generating Admissible Orders of Interval-Valued Fuzzy Sets

Emilio Torres-Manzanera
Dpt. of Statistics
University of Oviedo
Oviedo, Spain
torres@uniovi.es

Susana Díaz-Vázquez
Dpt. of Statistics
University of Oviedo
Oviedo, Spain
diazsusana@uniovi.es

Susana Montes
Dpt. of Statistics
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—Ordering interval-valued fuzzy sets is an open question. Admissible orders are a clever method to order them. Previous research has established the relationship between the most used admissible orders, linear transformations and aggregation functions. In this context, we provide a method to generate the most common admissible orders using Principal Component Analysis with random variables.**

***Index Terms*—Interval-valued fuzzy sets, admissible orders, Principal Component Analysis**

I. Introduction

The use of intervals is very common in the literature. The main disadvantage compared to using numbers is the absence of a universally accepted order, which would allow us to compare any pair of intervals. The typical order considered, known as lattice order, contains incomparabilities, and various approaches have been proposed to address this issue. In particular, Bustince et al. proposed a linear order for intervals in [4], which preserves the lattice order and is referred to as an admissible order. In fact, they introduced a method to derive admissible orders using aggregation functions. As it is stated in [7], there are several reasons to study such type of order, and in particular, by the fact that they find applications in many different areas as, for instance, image processing [3], classification [8] or decision making [5], for instance.

If we consider the field of image processing, independently of the previous comments, a very well-known technique is Principal Component Analysis (PCA) [6], which reduces the linear dimensionality of data. Thus, the main purpose of this paper is to link the PCA method with the most used admissible orders. It is organized as follows: Section II provides the theoretical background of both admissible orders and Principal Component Analysis. Additionally, it details some random variables. The next section discusses the relationship between the Lexicographical order, the Xu-Yager order, and random variables using PCA. Finally, the last part provides some concluding remarks and a guideline of future work.

This work was supported by the Spanish Ministry of Science and Innovation (Project PID2022-139886NB-I00).

II. Preliminaries

A. Admissible orders

Let Z denotes the universe of discourse. An interval-valued fuzzy set of Z is a mapping $A : Z \to L[0,1]$ such that $A(x) = [\underline{A}(x), \overline{A}(x)]$, where $L[0,1] = \{[r,s] \mid 0 \leq r \leq s \leq 1\}$, that is, it is the set of all closed intervals included in the interval $[0,1]$ [2]. An order $\preceq$ on the set $L([0,1])$ is a binary relation on $L([0,1])$ which is reflexive, antisymmetric and transitive, i.e.,

- Reflexivity: $a \preceq a$, $\forall a \in L([0,1])$.
- Antisymmetric: If $a \preceq b$ and $b \preceq a$, then $a = b$, $\forall a, b \in L([0,1])$.
- Transitivity: If $a \preceq b$ and $b \preceq c$, then $a \preceq c$, $\forall a, b, c \in L([0,1])$.

The most usual order in $L([0,1])$ is the Lattice order, which is defined as follows: $a \leq_2 b$ if and only if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$ with $a = [\underline{a}, \overline{a}]$, $b = [\underline{b}, \overline{b}]$. It is clear that it is induced by the typical order in $\mathbb{R}^2$.

A linear order is an order defined for any pair of elements of $L([0,1])$, that is, $a \preceq b$ or $b \preceq a$ for any $a, b \in L([0,1])$. Otherwise, it is said to be partial.

A very important family of linear orders are the admissible orders. Thus, it is said that $\preceq$ is an admissible order on $L([0,1])$ if it is a linear order that preserves the Lattice order, i.e., $a \leq_2 b$ implies $a \preceq b$.

The following relations are very typical examples of admissible orders:

- Lexicographical order: $[\underline{a}, \overline{a}] \preceq_{\text{Lex1}} [\underline{b}, \overline{b}] (\equiv)_{\text{def.}} (\underline{a} < \underline{b}) \vee [(\underline{a} = \underline{b}) \wedge (\overline{a} \leq \overline{b})]$. We denote the lexicographical order in $\mathbb{R}^2$ as $\leq_{\text{Lex1}}$: $(r_1, r_2) \leq_{\text{Lex1}} (s_1, s_2)$ when $(r_1 < s_1) \vee [(r_1 = s_1) \wedge (r_2 \leq s_2)]$.
- Antilexicographical order: $[\underline{a}, \overline{a}] \preceq_{\text{Lex2}} [\underline{b}, \overline{b}] (\equiv)_{\text{def.}} (\overline{a} < \overline{b}) \vee [(\overline{a} = \overline{b}) \wedge (\underline{a} \leq \underline{b})]$.
- Xu and Yager's order [9]: $[\underline{a}, \overline{a}] \preceq_{\text{xu-yager}} [\underline{b}, \overline{b}]$ if and only if

$$(\underline{a}+\overline{a} < \underline{b}+\overline{b}) \text{ or } [(\underline{a}+\overline{a} = \underline{b}+\overline{b}) \wedge (\overline{a}-\underline{a} \leq \overline{b}-\underline{b})]. \quad (1)$$

Concerning the Xu-Yager's order, if we consider the upper triangle of the unit square, $K([0,1]) = \{(x,y) \in [0,1]^2 \mid x \leq y\}$, this order projects the points onto the diagonal. In

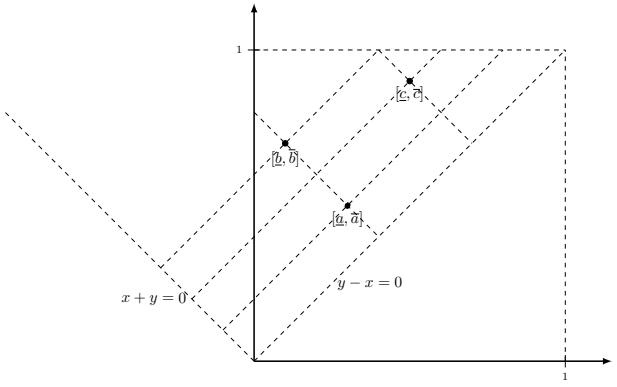


Fig. 1. Projections of the Xu and Yager's order into the diagonal or orthogonal projections.

case of equal projections over the diagonal, it compares the orthogonal projection (Fig. 1).

The following theorem characterises the generation of admissible orders using linear transformations.

Proposition 1. *[1] Let W be a 2×2 matrix. The binary relation*

$$[\underline{a}, \overline{a}] \preceq [\underline{b}, \overline{b}] \text{ if and only if } (\underline{a}, \overline{a}) \cdot W \leq_{Lex1} (\underline{b}, \overline{b}) \cdot W \quad (2)$$

is an admissible order if and only if W is full rank and for every row of W the first non null element is positive.

Example 1. The Lexicographic order can be represented by

$$W_{\text{Lex1}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (3)$$

and the Xu-Yager's order as

$$W_{\text{xu-yager}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \quad (4)$$

B. Random Variables

A random variable is a mathematical variable used in Probability Theory and Statistics to represent numerical outcomes of random phenomena. It's essentially a variable whose possible values are outcomes of a random event. One of the most well-known random variables is the uniform distribution. This is a random variable with a constant probability density or mass function that assigns equal probability o density to any value within the domain.

Example 2. If the region of the domain is $K([0,\ 1])$, then the density function of a uniform variable (X, Y) is

$$p(x, y) = \begin{cases} 2 & 0 \leq x \leq y \leq 1, \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

with matrix of covariances

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{18} & \frac{1}{36} \\ \frac{1}{36} & \frac{1}{18} \end{pmatrix}, \quad (6)$$

as $E(X) = \int\int_{K([0,1])} xp(x, y)\, dx\, dy$, $E(Y) = \int\int_{K([0,1])} yp(x, y)\, dx\, dy$, $\sigma_X^2 = \int\int_{K([0,1])} (x - E(X))^2 p(x, y)\, dx\, dy$, $\sigma_Y^2 = \int\int_{K([0,1])} (y - E(Y))^2 p(x, y)\, dx\, dy$ and $\sigma_{XY} = \int\int_{K([0,1])} (x - E(X))(y - E(Y))p(x, y)\, dx\, dy$.

Of course, this is just an example of random variable defined in $K([0,\ 1])$. Some other examples are shown below.

Example 3. The following random variable

$$p(x, y) = \begin{cases} \frac{1}{1-x} & 0 \leq x \leq y \leq 1, \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

has the covariance matrix

$$\Sigma = \begin{pmatrix} \frac{1}{12} & \frac{1}{24} \\ \frac{1}{24} & \frac{7}{144} \end{pmatrix}. \quad (8)$$

Example 4. The discrete variable

$$\Pr(X = x, Y = y) = \begin{cases} \frac{1}{4} & x \in \{0, \frac{1}{2}\}, y \in \{\frac{1}{2}, 1\}, \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

shows the covariance matrix

$$\Sigma = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}. \quad (10)$$

C. Principal Component Analysis

Principal Component Analysis (PCA) is a statistical method used for analysing and simplifying high-dimensional data while retaining its essential features [6]. It descomposes the covariance matrix obtaining its eigenvectors and eigenvalues. Eigenvectors represent the directions of maximum variance in the data, and eigenvalues represent the magnitude of variance along those directions. The first principal component, which corresponds to the eigenvector with greater eingenvalue, explain the most variance followed by the other eigenvectors. They are orthogonal.

Suppose that $\mathbf{X}$ is a vector of p random variables with a known covariance matrix Σ, $\mathbf{v}_k$ is an eigenvector of Σ corresponding to the kth largest eigenvalue λ_k for $k = 1, 2, \ldots, p$. If Σ is the matrix of variances and covariances, then eigenvalues λ are the solution of

$$\det(\Sigma - \lambda I) = 0 \quad (11)$$

with I the $p \times p$ identity matrix, and $\Sigma \cdot \mathbf{v}_k = \lambda_k \mathbf{v}_k$. The full principal components decomposition of $\mathbf{X}$ can therefore be given as $\mathbf{Z} = \mathbf{XV}$ where $\mathbf{V}$ is a $p \times p$ orthogonal matrix whose k column corresponds to the k eigenvector, $z_k = x_k \mathbf{v}_k$ and $\text{Var}(x_k \mathbf{v}_k) = \lambda_k$.

Example 5. In particular, the eigenvectors of a matrix of covariances

$$\Sigma = \begin{pmatrix} r & s \\ s & t \end{pmatrix} \quad (12)$$

with $r > 0$, $t > 0$ and $s \neq 0$ are

$$\mathbf{v}_1 = \left(-\frac{-r + t - \sqrt{r^2 + 4s^2 - 2rt + t^2}}{2s}, 1 \right), \quad (13)$$

$$\mathbf{v}_2 = \left(-\frac{-r + t + \sqrt{r^2 + 4s^2 - 2rt + t^2}}{2s}, 1 \right), \quad (14)$$

with eigenvalues

$$\lambda_1 = \frac{1}{2}\left(\sqrt{r^2 - 2rt + 4s^2 + t^2} + r + t \right), \quad (15)$$

and

$$\lambda_2 = \frac{1}{2}\left(-\sqrt{r^2 - 2rt + 4s^2 + t^2} + r + t \right) \quad (16)$$

respectively.

Example 6. When the random variables are uncorrelated, the covariance matrix becomes

$$\Sigma = \begin{pmatrix} r & 0 \\ 0 & t \end{pmatrix} \quad (17)$$

with $r > 0$ and $t > 0$. The eigenvalues are $\lambda_1 = r$ and $\lambda_2 = t$ with eigenvectors

$$\mathbf{v}_1 = (1, 0) \text{ and } \mathbf{v}_2 = (0, 1), \quad (18)$$

respectively.

III. Generating some admissible orders using PCA

PCA method provides a linear transformation of data. In the case of a uniform distribution over $K([0, 1])$, the eigenvectors (13) and (14) of the covariance matrix (6) form the columns of a square matrix

$$W = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \quad (19)$$

On the other hand, the Xu-Yager's order can be represented through a linear transformation using the lexicographical order: $[\underline{a}, \overline{a}] \preceq_{\mathrm{g}} [\underline{b}, \overline{b}]$ if and only if $(\underline{a}, \overline{a}) \cdot W \leq_{\mathrm{Lex1}} (\underline{b}, \overline{b}) \cdot W$ with

$$(\underline{a}, \overline{a}) \cdot W = (\underline{a}, \overline{a}) \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = (\underline{a} + \overline{a}, \overline{a} - \underline{a}). \quad (20)$$

Indeed, the process of deriving an admissible order from a random variable using the covariance matrix enables us to encapsulate the maximum variance inherent in the data. The subsequent inquiry pertains to identifying the inherent characteristics of random variables that universally yield an admissible order.

Let W be a linear transformation of PCA that generates the Xu-Yager order,

$$W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \quad (21)$$

What prerequisites must be satisfied by the covariance matrix to engender this admissible order?

Proposition 2. *A random variable defined in* $K[0, 1] = \{(x, y) \mid 0 \leq x \leq y \leq 1\}$ *generates the Xu-Yager order through the Principal Component Analysis method* (20) *if and only if the covariance matrix is*

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{pmatrix} \quad (22)$$

with $\sigma_X^2 = \sigma_Y^2$ *and* $\sigma_{X,Y} \neq 0$.

Proof. Suppose that Σ is a matrix of the type

$$\Sigma = \begin{pmatrix} r & s \\ s & r \end{pmatrix} \quad (23)$$

with $r > 0$ and $s > 0$. Then its orthogonal PCA transformations (13) and (14) become (19). And this is the linear transformation of the Xu-Yager method employing the lexicographical order (20).

Consider that the vectors $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (-1, 1)$ are the eigenvectors of a covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}. \quad (24)$$

By the properties of the eigenvectors, it holds that

$$\begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (25)$$

$$\begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad (26)$$

$$tr(\Sigma) = \sigma_X^2 + \sigma_Y^2 = \lambda_1 + \lambda_2 \quad (27)$$

After some algebraic steps, we obtain that, in the case that X and Y are non degenerate random varibles (i.e. constant variables), $\lambda_1 \neq 0$, $\lambda_1 + \lambda_2 > 0$ and

$$\sigma_X^2 = \sigma_Y^2 = \frac{\lambda_1 + \lambda_2}{2}, \quad \sigma_{XY} = \frac{\lambda_1 - \lambda_2}{2}. \quad (28)$$

If $\sigma_{XY} = 0$, Σ becomes (17) with eigenvectors (18), different to $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (-1, 1)$. Therefore, it results that $\sigma_{XY} \neq 0$.

□

Example 7. The uniform distribution of Example 2 generates the Xu-Yager order.

In the previous case, we have characterised the Xu-Yager order as a PCA transformation of correlated random variables. In particular, the uniform distribution over $K[0, 1]$ generates this order. Let us check out the case when the random variables are uncorrelated.

Proposition 3. *A random variable defined in* $K[0, 1] = \{(x, y) \mid 0 \leq x \leq y \leq 1\}$ *generates the lexicographical order through the Principal Component Analysis method* (20) *if and only if the covariance matrix is*

$$\Sigma = \begin{pmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{pmatrix}. \quad (29)$$

Proof. The linear transformation associated to the lexicographical order is

$$W = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (30)$$

Taking into account (17) and (18), the proof is trivial.

□

So, when the variables are uncorrelated, they generate the lexicographical order.

Example 8. The discrete random variable of Example 4 generates the lexicographical order.

In general, we can characterise the linear transformations of PCA that become admissible orders using the following result.

Theorem 1. *A random variable (X, Y) defined in $K[0,1] = \{(x, y) \mid 0 \leq x \leq y \leq 1\}$ generates an admissible order through the Principal Component Analysis method if and only if they are uncorrelated ($\sigma_{XY} = 0$) or when $\sigma_{XY} \neq 0$, we have that*

$$-\frac{-\sigma_X^2 + \sigma_Y^2 - \sqrt{\sigma_X^4 + 4\sigma_{XY}^2 - 2\sigma_X^2\sigma_Y^2 + \sigma_Y^4}}{2\sigma_{XY}} > 0, \quad (31)$$

or

$$\begin{aligned} -\frac{-\sigma_X^2 + \sigma_Y^2 - \sqrt{\sigma_X^4 + 4\sigma_{XY}^2 - 2\sigma_X^2\sigma_Y^2 + \sigma_Y^4}}{2\sigma_{XY}} &= 0, \\ -\frac{-\sigma_X^2 + \sigma_Y^2 + \sqrt{\sigma_X^4 + 4\sigma_{XY}^2 - 2\sigma_X^2\sigma_Y^2 + \sigma_Y^4}}{2\sigma_{XY}} &> 0. \end{aligned} \quad (32)$$

Proof. When X and Y are uncorrelated, it is the Proposition 3. When the correlation is not zero, then Proposition 1 obtains the desired result due to that the eigenvectors (13) and (14) are the columns of the linear transformation matrix W. □

IV. Conclusions

Image processing paves the way to apply the concept of admissible order of interval-valued fuzzy sets in this field. Through the Principal Component Analysis, we have established the relationship among random variables and the most used admissible order.

This is a first approach to this problem. Thus, several open problems need to be studied. In particular, the next step is to apply such findings in real cases of image processing. This process entails extrapolating the findings to three-dimensional data. Apart from that, we are interested on using this methodology in the most general environment for admissible orders.

Acknowledgment

The authors have been supported by the Spanish Ministry of Science and Innovation Project PID2022-139886NB-l00.

References

[1] Juan Baz, Maria Martinez, Susana Diaz-Vazquez, and Susana Montes. Ordering tuples. An application to risk matrices under uncertainty. Unpublished.

[2] Agustina Bouchet, Mikel Sesma-Sara, Gustavo Ochoa, Humberto Bustince, Susana Montes, and Irene Díaz. Measures of embedding for interval-valued fuzzy sets. *Fuzzy Sets and Systems*, 467:108505, 2023.

[3] 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. Theme: Information Processing and Applications.

[4] H. Bustince, J. Fernández, A. Kolesárová, and R. Mesiar. Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems*, 220:69–77, 2013.

[5] L. De Miguel, H. Bustince, J. Fernandez, E. Induráin, A. Kolesárová, and R. Mesiar. Construction of admissible linear orders for interval-valued atanassov intuitionistic fuzzy sets with an application to decision making. *Information Fusion*, 27:189–197, 2016.

[6] Ian T. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[7] Fagner Santana, Benjamin Bedregal, Petrucio Viana, and Humberto Bustince. On admissible orders over closed subintervals of [0,1]. *Fuzzy Sets and Systems*, 399:44–54, 2020. Fuzzy Intervals.

[8] J. Sanz, A. Fernández, H. Bustince, and 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):751–766, 2011.

[9] Zeshui Xu and Ronald Yager. Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems*, 35(4):417–433, 2006.

A Bi-criteria Approach to address the Interval Shortest Path Problem through TOPSIS

Pelayo S. Dosantos
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
suarezdpelayo@uniovi.es

Agustina Bouchet
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
bouchetagustina@uniovi.es

Irene Mariñas-Collado
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
marinasirene@uniovi.es

Susana Montes
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—This work addresses the Shortest Path Problem with interval costs through a bi-criteria perspective. Instead of optimizing the sum of the interval costs in the paths, it is suggested to optimize simultaneously both the lower and upper bounds of those intervals. This transforms the original problem into a bi-criteria Shortest Path Problem, allowing to solve it using the TOPSIS optimization method. Interpreting the problem as a bi-criteria problem leads to the emergence of different Pareto optimal solutions in practice. The use of TOPSIS as the optimization technique, ranks every Pareto optimal solution enabling the Decision Maker to find the shortest path, or best optimal solution in a network with intervals as costs. This methodology introduces a new perspective to handle uncertainty in the Interval Shortest Path Problem context that is both competitive and effective. The methodology is illustrated through a case study to demonstrate its performance and simplicity.**

***Index Terms*—Interval Shortest Path Problem, TOPSIS, Bi-criteria, Decision making**

I. Introduction

The Shortest Path Problem (SPP) is one of the most studied optimization problems due to its versatility and applications across various fields, including transportation, logistics and telecommunications [1], [2]. The SPP seeks to identify the most efficient route between two points in a network, minimizing a characteristic such as travel time, distance, or cost.

There are multiple variants of the problem tailored to very specific applications. However, a more general version of these problems, which is often overlooked, arises when imprecision is introduced into the data. This imprecision can occur in various ways, such as in the presence or absence of nodes or arcs in a network, or in the actual cost of traversing an arc. This work focuses on addressing the latter type of imprecision, i.e., when the cost of passing through an arc in a network, is not an exact real value.

To model such imprecision, there are also different options, including considering the costs as uncertain random variables [3] or using fuzzy set theory [4], [5]. However, for ease of interpretation by non-experts, it has been decided to express the imprecision of arc costs in the form of intervals. This implies that the true value of the cost falls within the range defined by the minimum and maximum of the interval. This approach leads to the Interval Shortest Path Problem (ISPP).

The problem that naturally arises when attempting to solve the ISPP is the absence of a generic total order in the set of intervals. This is why there is no general optimal solution for any given problem, and a methodology or an optimal criterion must be established to provide a solution. Regarding this, in [6] a robust shortest path solution is selected as the optimal solution. The approach involves finding the path that incurs the least loss when, once chosen, it is subjected to the worst-case scenario, where the traversed arcs have maximum cost and the remaining arcs have minimum cost. The cost of the path is compared to the cost of the shortest path in this scenario. In [7] a method for comparing intervals based on fuzzy theory is developed and they used their approach to solve an ISPP. In [8] a similar strategy is carried out. An acceptability index to state how true the sentence *an interval is preferred over another* is developed, and base on the index, a minimum in a set of intervals can be established, and thus, solve an ISPP. In [9], they find the shortest path that simultaneously minimize the total cost of the path and the risk of considering that path. The risk for each path is defined in terms of the distance from a point in the interval to the upper bound and the width of it.

In this work, a different approach has been contemplated to address the problem. The main idea is to consider each interval as a cost in a two dimensional space and solve the problem as if it were a bi-criteria shortest path problem. For that task, the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [10] is used, as it is widely used technique that provides good solutions in different contexts [11], [12]. There are many works where a multicriteria shortest path problem is solved in different ways (see for example [13]), but as far as is known, an ISPP has never been solved using any of these techniques. One of the main advantages of this approach is that, since the problem has already been addressed considering various criteria, it is sufficient to correctly make analogies with the interval problem. Additionally, the problems that arise due to the low correlation between criteria diminish, since in this case the endpoints of the interval serve to express the same magnitude.

The rest of the paper is organized as follows: Section II introduces the preliminary concepts needed for the rest of the

Authors would like to express their gratitude to the Spanish Ministry of Science and Innovation (fund reference: PREP2022-000355. Project: MCINN-23-PID2022-139886NB-I00).

work, Section III presents the proposed methodology to solve the ISPP, in Section IV a problem example is solved, Section V draws some conclusions about the work and mentions some research lines to be considered in the future.

II. Preliminaries

This section covers the essential concepts and notations required to understand the rest of the paper.

In the ISPP, the cost of traversing an edge in a graph is represented in terms of a closed interval $[\underline{a}, \overline{a}]$, where $0 \leq \underline{a} \leq \overline{a} < \infty$, meaning that the true cost is some value between its bounds. In that sense, the wider the interval, the more imprecision is represented. In the most common formulation, the objective is to find the path from a source node to a target node so that the sum of the costs of the edges in that path is minimized. The sum of two intervals $[\underline{a_1}, \overline{a_1}], [\underline{a_2}, \overline{a_2}]$ is defined as follows:

$$[\underline{a_1}, \overline{a_1}] \oplus [\underline{a_2}, \overline{a_2}] = [\underline{a_1} + \underline{a_2}, \overline{a_1} + \overline{a_2}] \tag{1}$$

Note that there exists a natural bijection between the intervals and the plane region:

$$\begin{aligned} i : \mathcal{L}([\underline{a}, \overline{a}]) &\to \mathcal{K}([\underline{a}, \overline{a}]) \\ i([\underline{a^*}, \overline{a^*}]) &= (\underline{a^*}, \overline{a^*}) \end{aligned} \tag{2}$$

where $\mathcal{L}([\underline{a}, \overline{a}]) = \{[\underline{a^*}, \overline{a^*}] : \underline{a} \leq \underline{a^*} \leq \overline{a^*} \leq \overline{a}\}$ and $\mathcal{K}([\underline{a}, \overline{a}]) = \{(\underline{a^*}, \overline{a^*}) : \underline{a} \leq \underline{a^*} \leq \overline{a^*} \leq \overline{a}\}$.

With this bijection, it is possible to interpret the ISPP as a bi-objective problem where the objective functions are, on one hand, minimize the sum of the lower bounds of the edges in a path from the source to the target node, and on the other, the corresponding sum of the upper bounds. Mathematically, the problem can be written as:

$$\begin{aligned} \min \quad & F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x})) \\ \text{subject to:} \quad & \mathbf{x} \in U, \end{aligned} \tag{3}$$

where $\mathbf{x} \in U$ means that $\mathbf{x}$ is a valid path and f_1 and f_2 are the mentioned sums.

The main issue to deal with in multi-objective problems, is that there is not guarantee of having uniqueness of optimal solution. In general, there exist many solutions that are neither better nor worse that others. This are often referred as Pareto optimal solutions [14]. More specifically, a solution $\mathbf{x}$ is said to be Pareto optimal, if there is not another solution $\mathbf{y}$ such that $\mathbf{y}$ is equally preferred or preferred in all the objectives and there is, at least, one objective in which $\mathbf{y}$ is strictly preferred over $\mathbf{x}$. If such $\mathbf{y}$ exists, it is said that ***y*** *dominates* ***x***. The set containing all the Pareto optimal solutions is called the *Pareto set*. To refer to the Pareto set of a problem defined by the objective function F and feasible region U, the notation $PS_{F,U}$ will be used.

Two interesting points in multi-objective problems are the Positive Ideal Solution (PIS) $F(\mathbf{x})^+$ and the Negative Ideal Solution (NIS) $F(\mathbf{x})^-$ [15]. These are the combination of the best and worst values of each considered solution (usually the Pareto optimal ones). In the bi-objective problem defined by (3), the PIS and NIS can be written as:

$$F(\mathbf{x})^+ = \left(\min_{\mathbf{x} \in PS_{F,U}} f_1(\mathbf{x}), \min_{\mathbf{x} \in PS_{F,U}} f_2(\mathbf{x}) \right) \tag{4}$$

$$F(\mathbf{x})^- = \left(\max_{\mathbf{x} \in PS_{F,U}} f_1(\mathbf{x}), \max_{\mathbf{x} \in PS_{F,U}} f_2(\mathbf{x}) \right) \tag{5}$$

Note that if there exists $\mathbf{x} \in U$ such that it minimizes $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$, then $F(\mathbf{x}) = F(\mathbf{x})^+$ and $\mathbf{x}$ is the optimal solution, although it could happen that more than one $\mathbf{x}$ satisfy this condition. If so, any of them give the same results and it does not matter the final election. However, this is not the case in most of the real life problems. The utility that the PIS and the NIS have is that they can act as reference point to evaluate the solutions in the feasible region.

III. TOPSIS applied to the ISPP

In this section, the proposed method for solving the ISPP through TOPSIS is introduced, along with the advantages of using this method in the context of this problem.

Recall that the ISPP is defined as follows: the pair $G(V, E)$ defines a graph where $V = \{v_1, \ldots, v_n\}$ is the set of nodes, and $E \subseteq V \times V$ is the set of edges in the graph. Each edge $e_{ij} = (v_i, v_j) \in E$ has an associated interval cost expressed by $c_{i,j} = [\underline{c_{i,j}}, \overline{c_{i,j}}]$. Given a source node v_s and a target node v_t, the objective is to find the path P from v_s to v_t that minimizes the sum of the costs along the path.

The issue here is the lack of a general order in the set $\mathcal{L}([\underline{a}, \overline{a}])$. Therefore, there is no clear minimum when comparing different path costs. A possible approach to address this problem is the use of admissible orders [16]. These are total orders that refine the usual partial order in $\mathcal{L}([\underline{a}, \overline{a}])$, denote $\precsim$, defined by $[\underline{a_1}, \overline{a_1}] \precsim [\underline{a_2}, \overline{a_2}]$ if and only if $\underline{a_1} \leq \underline{a_2} \wedge \overline{a_1} \leq \overline{a_2}$. In [16] it is established how to generate admissible orders through aggregation functions applied to the bounds of the intervals. However, when performing comparisons in these terms, the choice of the aggregation functions may give more weight to one of the bounds of the interval when making comparisons. To avoid this, the use of the bijection i introduced in Section II is suggested to consider the ISPP as a bi-criteria problem that can be solved using TOPSIS. Now the aim is to simultaneously minimize the following objective functions

$$f_1(P) = \sum_{e_{ij} \in P} \underline{c_{i,j}}, \quad f_2(P) = \sum_{e_{ij} \in P} \overline{c_{i,j}} \tag{6}$$

With the problem defined in this manner, we can now apply the TOPSIS method. Nevertheless, it is convenient to make some brief notes on how to use the method in this case.

A. TOPSIS procedure

The following describes how TOPSIS is applied to solve the bi-criteria problem equivalent to ISPP. From now on, it is assumed that the Pareto set $PS = \{P_1, \ldots, P_k\}$ of optimal paths has already been found. Each path $P_i \in PS$ has a pair of costs $c_i = (\underline{c_i}, \overline{c_i})$.

When dealing with multi-objective problems, it is common for the involved functions to measure different aspects of the problem, being necessary a standardization to compare the performances of solutions in the objective space. The classical TOPSIS method also includes this step. However, in this case, it can be skipped, due to the the fact that the lower and upper bounds of the intervals measure, in the same scale, the same magnitude in the problem.

The next step in TOPSIS is the selection of weights ω_1, ω_2 to reflect the importance of the objective functions f_1 and f_2 respectively. Nevertheless, as it is supposed that there is not additional information apart from that the true value of the costs of the edges is between their bounds, equal weights are considered and this step can also be skipped.

After that, the PIS and NIS need to be identified. To do so, let $PS_0 = \{P_{(1)}, \dots, P_{(k)}\}$ be the Pareto set of optimal paths PS, ordered according to the first component of their costs. Each path $P_{(i)} \in PS_0$ has a pair of costs $c_{(i)} = (\underline{c_{(i)}}, \overline{c_{(i)}})$ and $\underline{c_{(i)}} < \underline{c_{(j)}}$ if $i<j$, $i, j = \{1, \dots, k\}$. The PIS and NIS are characterized by the following proposition.

***Proposition** 3.1:* Let $PS_0 = \{P_{(1)}, \dots, P_{(k)}\}$ be the Pareto set of a bi-criteria problem defined by an ISPP such as $\underline{c_{(1)}} < \underline{c_{(2)}} < \dots < \underline{c_{(k)}}$. Then, $\overline{c_k} < \overline{c_{(k-1)}} < \dots < \overline{c_{(1)}}$.

Proof: Lets see that $\min\limits_{i \in \{1,\dots,k\}} \{\overline{c_{(i)}}\} = \overline{c_{(k)}}$.

Suppose that there exists $\overline{c_j}$, $j \neq k$ such as $\overline{c_{(j)}} < \overline{c_{(k)}}$. Then,

$$f_1(P_{(j)}) = \underline{c_{(j)}} < \underline{c_{(k)}} = f_1(P_{(k)})$$

and

$$f_2(P_{(j)}) = \overline{c_{(j)}} < \overline{c_{(k)}} = f_2(P_{(k)}).$$

Therefore, $P_{(j)}$ dominates $P_{(k)}$ and $P_{(k)} \notin PS_0$. That is a contradiction, so

$$\min_{i \in \{1,\dots,k\}} \{\overline{c_{(i)}}\} = \overline{c_{(k)}}.$$

To proof the rest of the inequalities, it is sufficient to apply an iterative reasoning considering the set $PS_0 \setminus \{P_{(k)}\}$.

Note that strict inequalities are not restrictive at all in bi-criteria problems. If $c_{(i)} = c_{(j)}$ for $i \neq j$, then, two cases can occur: if $\overline{c_{(i)}} \leq \overline{c_{(j)}}, \overline{P_{(j)}} \notin \overline{PS_0}$ and if $\overline{c_{(i)}} = \overline{c_{(j)}}$, either $P_{(j)}$ or $P_{(i)} \notin PS_0$.

In view of Proposition 3.1, it is possible to easily identify the PIS and NIS of the problem:

$$F(P)^+ = (\underline{c_{(1)}}, \overline{c_{(k)}}), \quad F(P)^- = (\underline{c_{(k)}}, \overline{c_{(1)}}). \tag{7}$$

Once the PIS and NIS are identified, the distance in the objective space from each solution P_i in PS to $F(P)^+$ and $F(P)^-$ is computed. This distances are denoted D_i^+ and D_i^- respectively. Any distance can be used at this stage. In this case, the euclidean distance has been chosen.

Finally, each solution $P_i \in PS$ gets a score R_i defined by

$$R_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{8}$$

and all are ranked in decreasing order of score. This score measures how far is each Pareto optimal solution from the NIS and how near they are to the PIS.

IV. Case study

In this section a brief example to show how the proposed approach can help to solve the ISPP is presented. A directed graph with just 9 nodes has been considered to provide clarity to the methodology. The aim is to find the shortest path from v_s to v_t. The graph can be seen in Fig. 1.

The first step is to interpret every cost as a pair in $\mathbb{R}^2$. After that, the Pareto set needs to be found. To do it, there are many strategies that one can follow. If it is desired to find the whole Pareto set, the multicriteria Dijkstra's algorithm presented in [17] can be used. If just an approximation of the Pareto set is needed, any genetic or evolutionary algorithm is more than sufficient to find it. As in this case, the number of nodes is 9 and the number of edges is 28, there is a reasonable number of paths and an evolutionary algorithm has many chances to find the entire Pareto set in a very short time. For that reason, the code provided by Xavier Ma in the GitHub repository[1], that follows the approaches in [18] and [19], has been used. In Table I, the Pareto optimal paths found for the problem are displayed.

As can be seen, despite working with a small network of only 9 nodes, a total of 5 optimal paths have been found. This fact highlights the need for a technique to handle potential solutions when dealing with a larger and more realistic problem. The Pareto front of the problem along with the PIS and NIS is shown in Fig. 2.

Finally, in order to rank the five alternatives, the score ratio (8) is used. The scores obtained by each path can be seen in Table II, where they have also been arranged in order of preference from left to right.

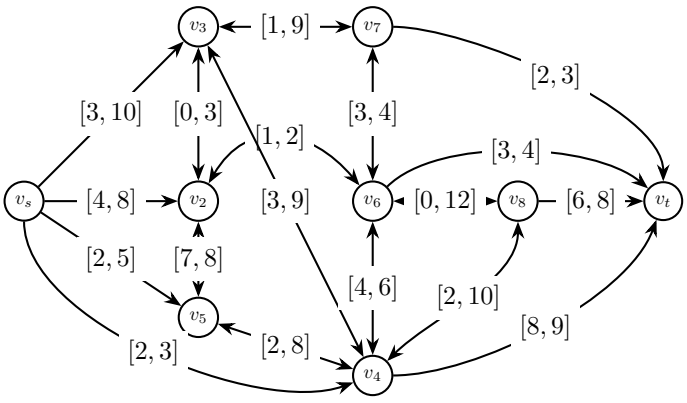


Fig. 1. Graph representing the ISPP.

[1]https://github.com/Xavier-MaYiMing/The-MOEAD-for-the-multi-objective-shortest-path-problem

TABLE I
PARETO OPTIMAL PATHS OF THE ISPP DEFINED IN FIG. 1.

Label	Shortest path	Interval costs
P_1	$v_s - v_3 - v_7 - v_t$	[6, 22]
P_2	$v_s - v_2 - v_6 - v_t$	[8, 14]
P_3	$v_s - v_3 - v_2 - v_6 - v_t$	[7, 19]
P_4	$v_s - v_4 - v_6 - v_t$	[9,13]
P_5	$v_s - v_4 - v_t$	[10, 12]

In view of Table II, the minimum path chosen by this approach to the ISPP is P_2, the path located between the two most extreme solutions in terms of interval width. Additionally, it is worth noting that the path with the widest interval, P_1, and thus the greatest uncertainty, has turned out to be the worst option for the method. However, this is not necessarily general, as no pattern seems to be observed that relates the path scores to the width of the cost interval, although it is an interesting topic to consider for future research.

As a final note, it is worth mentioning that TOPSIS is a widely used technique in multicriteria optimization problems, which is not only employed to select the best alternative from a set but also to narrow down a manageable set of alternatives for a human. Since the method provides a preference ranking for all options, a desired number of optimal solutions can be preset, allowing the technique to provide the Decision Maker (DM) with that number of alternatives from which they can make the final decision. In this case, if just three paths are desired, the chosen ones will be P_2, P_4 and P_5, and an expert DM can then decide based on its own perspective.

V. CONCLUSION

In this work, the use of the direct relationship of the space of positive intervals with a subset of $\mathbb{R}^{+2}$, to treat the ISPP as a bi-criteria optimization problem has been proposed. To solve this problem, TOPSIS has been employed, as it is a widely used technique and has shown to perform well in a variety of contexts. Due to the nature of the problem, it is sufficient to use a simplified version of TOPSIS, without the need

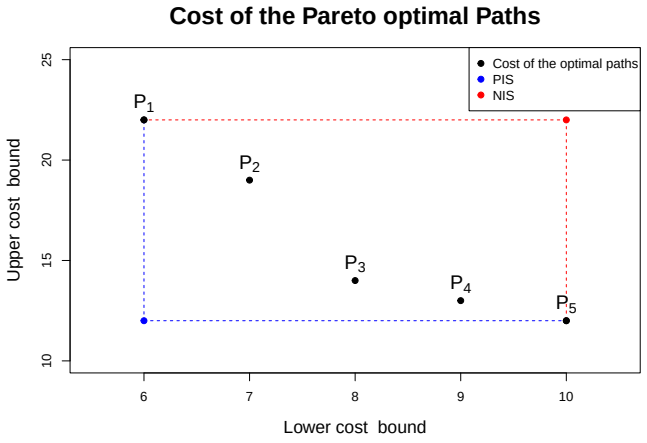


Fig. 2. Costs of the Pareto optimal paths for the ISPP defined in Fig 1.

TABLE II
SCORES OF THE PARETO OPTIMAL PATHS ACCORDING TO TOPSIS CRITERION.

Path (P_i)	P_2	P_4	P_5	P_3	P_1
Score (R_i)	0.7446	0.7412	0.7143	0.3750	0.2857

of several steps such as the standardization of the objectives or assigning weights to them.

To find the Pareto set a MOEA/D algorithm has been used. It is important understand that although there is not guarantee that this kind of algorithms will find the entire Pareto set of a multicriteria problem, for large scale problems, the approximations of the set of optimal solutions that they provide are more than enough. In addition, it is also relevant to keep always in mind that by definition, the Pareto set is form by all the solutions such as there are not others better than them, so even in the case of an incomplete Pareto, the final solution will be adequate.

The proposed approach has been applied to a toy example and has shown to choose a sensible solution. In addition, the use of TOPSIS to create a reduced Pareto set has also been mentioned and seem to be a good use of this technique.

In summary, the idea of solving the ISPP using a technique designed to address multicriteria problems such as TOPSIS has benefited from the advantages of these well-studied techniques and has yielded promising and satisfactory results.

As future research lines, it is planned to investigate whether there is any relationship between the solutions provided by the method and the properties of the intervals involved in the problems, such as the midpoint, width, and endpoints. Discovering such relationships, if they exist, could be highly beneficial for relaxing the computational costs of the method by focusing the search on solutions that satisfy specific criteria. It also seems natural to consider using other bi-criteria optimization techniques and comparing the solutions provided with those of TOPSIS, in addition to studying the same properties for the intervals.

REFERENCES

[1] Z. Zhang, W. Jigang, and X. Duan, "Practical algorithm for shortest path on transportation network," in International Conference on Computer and Information Application, 2010 pp. 48-51.
[2] C. Gloaguen, F. Fleischer, H. Schmidt, and V. Schmidt, "Analysis of shortest paths and subscriber line lengths in telecommunication access networks," Networks and Spatial Economics, vol.10, pp. 15-47. March 2010.
[3] Y. Sheng, and Y. Gao, "Shortest path problem of uncertain random network," Computers & Industrial Engineering, vol.99, pp. 97-105. July 2016.
[4] C. M. Klein, "Fuzzy shortest paths," Fuzzy sets and systems, vol.39, 1, pp. 27-41, January 1991.
[5] S. Okada and T. Soper, "A shortest path problem on a network with fuzzy arc lengths," Fuzzy sets and systems, vol.109, 1, pp. 129-140, January 2000.
[6] R. Montemanni and L. M. Gambardella, "The robust shortest path problem with interval data via Benders decomposition," 4or, vol.3, pp. 315-328, December 2005.
[7] A. Sengupta, and T. K. Pal, "Solving the shortest path problem with interval arcs," Fuzzy Optimization and Decision Making, vol.5, pp. 71-89, January 2006.

[8] A. Ebrahimnejad, “An acceptability index based approach for solving shortest path problem on a network with interval weights,” RAIRO-Operations Research, vol.55, pp. S1767-S1787, March 2021.
[9] X. Chen, J. Hu, and X. Hu, “A new model for path planning with interval data,” Computers & operations research, vol.36, 6, pp.1893-1899, June 2009.
[10] C. L. Hwang and K. Yoon, “Methods for multiple attribute decision making,” Multiple attribute decision making: methods and applications a state-of-the-art survey, LNE, vol.186, pp. 58-191, 1981.
[11] M. Behzadian, S. K. Otaghsara, M. Yazdani, and J. Ignatius, “A state-of the-art survey of TOPSIS applications,” Expert Systems with applications, vol.39, 17, pp. 13051-13069. December 2012.
[12] K. Palczewski, and W. Sałabun, ”The fuzzy TOPSIS applications in the last decade,” Procedia Computer Science, vol.159, pp. 2294-2303, January 2019.
[13] Z. Tarapata, “Selected multicriteria shortest path problems: An analysis of complexity, models and adaptation of standard algorithms,” International Journal of Applied Mathematics and Computer Science, vol.17, 2, pp. 269-287, June 2007.
[14] J. Lin, “Multiple-objective problems: Pareto-optimal solutions by method of proper equality constraints,” IEEE Transactions on Automatic Control, vol.21, 5, pp. 641-650. October 1976.
[15] Y. J. Lai, T. Y. Liu, and C. L. Hwang, “Topsis for MODM,” European journal of operational research, vol.76, 3, pp. 486-500. August 1994.
[16] H. Bustince, J. Fernandez, A. Kolesarova, and R. Mesiar “Generation of linear orders for intervals by means of aggregation functions,” Fuzzy Sets and Systems, vol.220, pp. 69-77, June 2013.
[17] N. Hanusse, D. Ilcinkas, A. and Lentz, “Framing algorithms for approximate multicriteria shortest paths,” In 20th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2020). Schloss-Dagstuhl-Leibniz Zentrum für Informatik, Vol. 85, pp. 11:1-11:19, November 2020.
[18] Q. Zhang, and H. Li, “MOEA/D: A multiobjective evolutionary algorithm based on decomposition,” IEEE Transactions on evolutionary computation, vol.11, 6, pp. 712-731, December 2007.
[19] C. W. Ahn, and R. S. Ramakrishna, “A genetic algorithm for shortest path routing problem and the sizing of populations,” IEEE transactions on evolutionary computation, vol.6, 6, pp. 566-579, December 2002.

Penalty functions for intervals

Pedro Huidobro
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
huidobropedro@uniovi.es

Agustina Bouchet
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
bouchetagustina@uniovi.es

Susana Díaz-Vázquez
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
diazsusana@uniovi.es

Susana Montes
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—Aggregation functions simplify complex problems by deriving a single value from multiple inputs, often used in decision-making and image processing. Penalty functions complement aggregation by quantifying the cost of deviating from a consensus value, gaining attention for their versatility. Interval-valued functions are also rising in importance, especially in uncertain scenarios. This paper extends penalty functions to intervals demonstrating their applicability in handling interval-valued data and providing a structured approach to decision-making in uncertain environments.**

***Index Terms*—Penalty function, interval, admissible order, aggregation function**

I. Introduction

Nowadays, numerous situations involve multidimensional problems. To simplify complex problems, aggregation functions prove invaluable, as they enable the derivation of a singular value from a multitude of input values [2], [11]. This characteristic has proven particularly beneficial for researchers dealing with challenges in decision-making and image processing [8], [16].

Penalty functions play a crucial role in this process by providing a mechanism to quantify the cost or penalty associated with deviating from a consensus value. Imagine a scenario where you have several inputs representing potential values, but you need to choose a single representative value for further analysis or decision-making. You need a tool that helps you to decide which value should be chosen and penalty functions can overcome this task. Since they were introduced, the interest in penalty functions has grown among researchers due to their ability to assign a cost to the divergence of each value from the selected consensus value (see for instance [4], [6]).

Simultaneously, the significance of interval-valued functions is on the rise, especially in real-world scenarios where obtaining an accurate value can be challenging, prompting a preference for using interval-valued data [3].

Considering the aforementioned points, our focus is on exploring the aggregation of interval-valued functions and analyzing the determination of the most suitable aggregation method. From a sample of interval-valued data we will choose the one with less penalty value associated.

This paper is organized as follows: Section II gives the basic concepts about intervals such as admissible orders. In Section III, we review some historical definitions of penalty function. Section IV presents a first approach to penalty functions with interval-valued data, along with some examples. Finally, conclusions and future work are presented in Section V.

Authors would like to thank for the support of the Spanish Ministry of Science and Innovation (project PID2022-139886NB-I00).

II. Preliminaries

We denote by $L(\mathbb{R})$ the family of closed intervals of the real line, i.e., $L(\mathbb{R}) = \{x = [\underline{x}, \overline{x}] : \underline{x}, \overline{x} \in \mathbb{R} \text{ and } \underline{x} \leq \overline{x}\}$.

In this work, it will be essential to arrange intervals in a specific order. Hence, establishing a total order between intervals becomes necessary. A common partial order is the lattice order:

***Definition** 2.1:* [10] The lattice order $\preceq_{Lo}$ in $L([0,1])$ is given by $x \preceq_{Lo} y$ iff $\underline{x} \leq \underline{y}$ and $\overline{x} \leq \overline{y}$.

Using this partial order, a total order between intervals can be defined as follows:

***Definition** 2.2:* [5] The order $\preceq$ on $L([0,1])$ is called an admissible order if:

- $\preceq$ is a total order on $L([0,1])$,
- $\preceq$ refines the lattice order on $L([0,1])$, i.e., for all $x, y \in L([0,1])$, $x \preceq y$ whenever $x \preceq_{Lo} y$.

***Example** 2.1:* In the literature some examples of admissible orders can be found:

- Lexicographical order type 1 [5]:

$$a \preceq_{Lex1} b \text{ if } \underline{a} < \underline{b} \text{ or } (\underline{a} = \underline{b} \text{ and } \overline{a} \leq \overline{b})$$

- Lexicographical order type 2 [5]:

$$a \preceq_{Lex2} b \text{ if } \overline{a} < \overline{b} \text{ or } (\overline{a} = \overline{b} \text{ and } \underline{a} \leq \underline{b})$$

- The Xu and Yager order [13]:

$$a \preceq_{XY} b \text{ if } \underline{a}+\overline{a} < \underline{b}+\overline{b} \text{ or } \underline{a}+\overline{a} = \underline{b}+\overline{b} \text{ and } \overline{a}-\underline{a} \leq \overline{b}-\underline{b})$$

Bustince et. al [5] also proposed a method to construct admissible orders using aggregation functions.

***Definition** 2.3:* [1], [12] Let $\mathcal{A} : [0,1]^n \to [0,1]$ such that

- $\mathcal{A}(0,0,\ldots,0) = 0, \mathcal{A}(1,1,\ldots,1) = 1$,
- $\mathcal{A}$ is increasing in each variable,

then $\mathcal{A}$ is an aggregation function.

There is a natural bijection between $L([0,1])$ and $K([0,1]) = \{(u,v) \in [0,1]^2 \,|\, u \leq v\}$ linking the interval $[\underline{a}, \overline{a}]$ to the point formed by its endpoints in $\mathbb{R}^2$. Aggregation methods allow us to combine information presented as an interval.

***Proposition** 2.1:* [5] Let $\mathcal{A}, \mathcal{B}$ be continuous aggregation functions, such that for all $(u,v),(u',v') \in K([0,1])$, the equalities $\mathcal{A}(u,v) = \mathcal{A}(u',v')$ and $\mathcal{B}(u,v) = \mathcal{B}(u',v')$ can only hold if $(u,v) = (u',v')$. Define the relation $\preceq_{\mathcal{A},\mathcal{B}}$ on $L([0,1])$ by $a \preceq_{\mathcal{A},\mathcal{B}} b$ if and only if

$$\mathcal{A}(\underline{a}, \overline{a}) < \mathcal{A}(\underline{b}, \overline{b})$$

or

$$\mathcal{A}(\underline{a}, \overline{a}) = \mathcal{A}(\underline{b}, \overline{b}) \text{ and } \mathcal{B}(\underline{a}, \overline{a}) \leq \mathcal{B}(\underline{b}, \overline{b})$$

Then $\preceq_{\mathcal{A},\mathcal{B}}$ is an admissible order on $L([0,1])$.

***Example** 2.2:* The lexicographical orders type 1 and type 2 or the Xu and Yager order can be constructed by aggregation functions:

- Lexicographical order type 1:
  - $\mathcal{A}_{\text{Lex1}}(u,v) = u$
  - $\mathcal{B}_{\text{Lex1}}(u,v) = v$
- Lexicographical order type 2:
  - $\mathcal{A}_{\text{Lex2}}(u,v) = v$
  - $\mathcal{B}_{\text{Lex2}}(u,v) = u$
- The Xu and Yager order:
  - $\mathcal{A}_{\text{XY}}(u,v) = u + v$
  - $\mathcal{B}_{\text{XY}}(u,v) = u - v$

In our case, we are dealing with intervals on the real line; thus, we need to extend this admissible order from $L([0,1])$ to $L(\mathbb{R})$. This extension is trivial and occurs naturally.

Finally, we introduce the concept of aggregation function for intervals that we will use in order to characterize the penalty functions. This is a tricky concept as the original definition of aggregation function is defined on the unit interval $[0,1]$. As we are considering intervals in $L(\mathbb{R})$, we also need to extend this notion. Based on the ideas of Grabisch [11], the natural extension would be something as follows:

A function $\mathcal{A} : L(\mathbb{R})^n \to L(\mathbb{R})$ is called an aggregation function for intervals if:

- $\mathcal{A}$ is non decreasing with respect to a prefixed order ($\preceq_o$), i.e., if $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y} \in L(\mathbb{R})$, and there exists $i \in \{1, \ldots, n\}$ such that $\mathbf{y} \preceq_o \mathbf{x}_i$, then

$$\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n) \preceq_o \mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$$

- $\mathcal{A}$ fulfills that

$$\inf_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in L(\mathbb{R})} \{\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} = \inf\{L(\mathbb{R})\}$$

and

$$\sup_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in L(\mathbb{R})} \{\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} = \sup\{L(\mathbb{R})\}$$

where $\inf$ and $\sup$ obviously would depend on the prefixed order. It is well known that $L(\mathbb{R})$ has neither infimum nor supremum for the classical orders. In this preliminary contribution, in order to overcome this situation, we will not work in the most general situation, $L(\mathbb{R})$ but we will restrict to an interval in $\mathbb{R}$: the greatest interval that includes all the intervals $x_i = [\underline{x_i}, \overline{x_i}]$ to be aggregated, i.e., $L([\underline{a}, \overline{a}])$ where:

$$\underline{a} = \min\{\underline{x_i}\}_{i=1}^n$$

and

$$\overline{a} = \max\{\overline{x_i}\}_{i=1}^n$$

In that way, we guarantee the existence of the infimum and supremum.

***Definition** 2.4:* Let $[a,b] \in L(\mathbb{R})$ and let $\preceq$ be an order such that $(L([a,b]), \preceq)$ is a lattice. A function $\mathcal{A} : L([a,b])^n \to L([a,b])$ is called an $\preceq$-aggregation function for intervals if:

- $\mathcal{A}$ is non decreasing with respect to $\preceq$, i.e., if $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y} \in L([a,b])$, and exist $i \in \{1, \ldots, n\}$ such that $\mathbf{y} \preceq \mathbf{x}_i$, then

$$\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{y}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n) \preceq \mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$$

- $\mathcal{A}$ fulfills that

$$\inf_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in L([a,b])} \{\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} = \inf\{L([a,b])\}$$

and

$$\sup_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in L([a,b])} \{\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} = \sup\{L([a,b])\}$$

where $\inf$ and $\sup$ are the infimum and supremum with respect to $\preceq$.

Observe that since we have asked $(L([a,b]), \preceq)$ to be a lattice, the existence of the infimum and supremum are guaranteed. The lattice order and therefore any admissible order are examples of orders that can be used to properly define an aggregation function for intervals.

III. Penalty Function

In this section, we briefly revisit some historical definitions. A more in-depth exploration can be found in [4].

In 1993, Yager [14] introduced the initial concepts of employing a penalty function to support aggregation functions. He proposed associating a penalty or cost with each data point x_i, payable if we disregard it when concluding the aggregation process with a conflicting value for y.

Subsequently, in 1997, Yager and Rybalov [15] advocated minimizing a penalty function as a means to obtain a fused value for n observations. Their approach is outlined as follows:

***Definition** 3.1:* [4] The function $LP : \mathbb{R}^2 \to \mathbb{R}^+$ is a local penalty function if, for any $x_i, x_j, y \in \mathbb{R}$ and $i, j = 1, ..., n$, it satisfies:

- $LP(x_i, y) = 0$, if $x_i = y$,
- $LP(x_i, y) > 0$ if $x_i \neq y$,
- $LP(x_i, y) \geq LP(x_j, y)$ if $|x_i - y| > |x_j - y|$,

where y is termed the fused value related to each observation in x.

They construct penalty functions from local penalty functions.

***Definition** 3.2:* [4] Let $LP : \mathbb{R}^2 \to \mathbb{R}^+$ be a local penalty function. A penalty function $P : \mathbb{R}^{n+1} \to \mathbb{R}^+$ is defined, for any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$ as:

$$P(\vec{x}, y) = \sum_{i=1}^{n} LP(x_i, y)$$

where y is called the fused value related to each observation in x.

Using penalty functions allows us to measure the total penalty incurred when y serves as the fused value of x. The optimal fused value y^* for x minimizes the penalty function, obtained as:

$$P(\vec{x}, y^*) = \min_{y} P(\vec{x}, y).$$

With this definition, the set of minimizers of $P(\vec{x}, \cdot)$ may not exist or be unique in general. Several authors have discussed this topic, as seen in [6], [7] and [9]. We would like to highlight the following definition of a penalty function given by Beliakov and James [9]:

***Definition** 3.3:* The function $P : [0, 1]^{n+1} \to [0, \infty]$ is a penalty function if and only if it satisfies:

- $P(x_i, y) = 0$, if $x_i = y$,
- $P(x_i, y) > 0$ if $x_i \neq y$,
- For every fixed $\vec{x}$, the set of minimizers of $P(\vec{x}, \cdot)$ is either a singleton or a connected set.

Later, Bustince et. al [4] proposed the following definition, giving a more direct characterization. In this definition, they consider I as an interval contained in the real line.

***Definition** 3.4:* The function $P : I^n \times L([0, 1]) \to \mathbb{R}^+$ is a penalty function if and only if there exists $c \in \mathbb{R}$ such that:

- $P(\vec{x}, y) \geq c$, for all $\vec{x} \in I^n, y \in L([0, 1])$.
- $P(\vec{x}, y) = c$ if and only if $x_i = y$ for all $i = 1, .., n$.
- P is quasi-convex lower semi-continuous in y for each $\vec{x} \in I^n$.

Some examples of penalty function can be found in [4].

IV. Penalty Functions for Intervals

As we previously commented, interval-valued functions are experiencing increasing importance, particularly in practical situations where acquiring precise values is challenging. In this section, we will introduce a first approach to the definition of penalty function for intervals.

To introduce the concept of penalty function for intervals, as an initial step, we will consider a finite number of interval-valued data and an admissible order $\preceq$. Therefore, we denote by $\vec{\mathbf{x}}$ an interval-valued vector, i.e.:

$$\vec{\mathbf{x}} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$$

where $\mathbf{x}_i \in L(\mathbb{R})$, for all $i = 1, \ldots, n$.

As mentioned previously, we will consider the largest interval $a = [\underline{a}, \overline{a}]$ where $\underline{a}$ and $\overline{a}$ represent, respectively, the lowest lower bound and the greatest upper bound over all the intervals $\mathbf{x}_i$.

***Definition** 4.1:* A function $P : L([\underline{a}, \overline{a}])^{n+1} \to \mathbb{R}^+$ is called penalty function for intervals in $L([\underline{a}, \overline{a}])$ if the following properties are fulfill:

P1: $P(\vec{\mathbf{x}}, \mathbf{y}) \geq 0$ for all $\mathbf{x} \in L([\underline{a}, \overline{a}])^n$, $\mathbf{y} \in L([\underline{a}, \overline{a}])$

P2: $P(\vec{\mathbf{x}}, \mathbf{y}) = 0$ if $\mathbf{x}_i = \mathbf{y}$ for all $i \in \{1, \ldots, n\}$

It is important to note that in this case, it is not necessary to introduce the conditions of quasi-convexity and lower semi-continuity. In the case of working in $\mathbb{R}$ instead of $L([\underline{a}, \overline{a}])$, Bustince et al. [4] asked for those two conditions in order to guarantee the existence of a new $y \in \mathbb{R}$ that minimizes the penalty function considered. In our case, we can skip those two conditions because we are working with a finite number of intervals, ensuring the existence of a minimum and a maximum. Also, as we are in $L([\underline{a}, \overline{a}])$, an admissible order ensures the existence of an infimum and a supremum, in fact, even for the lattice order (and therefore for any admissible order) the existence of a minimum (the interval $[\underline{a}, \underline{a}]$) and a maximum (the interval $[\overline{a}, \overline{a}]$) is guaranteed.

The idea is that $P(\vec{\mathbf{x}}, \mathbf{y}) = f_{\vec{\mathbf{x}}}(\mathbf{y})$ because we aim to find the fused value $\mathbf{y} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ that best summarizes $\vec{\mathbf{x}} \in L([\underline{a}, \overline{a}])^n$.

***Proposition** 4.1:* Let $\epsilon > 0$ and $\vec{\mathbf{x}} \in L([\underline{a}, \overline{a}])^n$. The function $f_{\vec{\mathbf{x}}} : L([\underline{a}, \overline{a}])) \to [0, 1]$ given by:

$$P(\vec{\mathbf{x}}, \mathbf{y}) = f_{\vec{\mathbf{x}}}(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}_i = \mathbf{y}, \forall i = 1, \ldots, n \\ \epsilon & \text{otherwise} \end{cases}$$

is a penalty function for intervals in $L([\underline{a}, \overline{a}])$.

Proof:

- As $\epsilon > 0$, we have that $P(\vec{\mathbf{x}}, \mathbf{y}) \geq 0$ for all $\mathbf{x} \in L([\underline{a}, \overline{a}])^n$, $\mathbf{y} \in L([\underline{a}, \overline{a}])$
- If $\mathbf{x}_i = \mathbf{y}$ for all $i \in \{1, \ldots, n\}$, then $P(\vec{\mathbf{x}}, \mathbf{y}) = 0$.

***Proposition** 4.2:* Fixed $\vec{\mathbf{x}} \in L([\underline{a}, \overline{a}])^n$. Let $\preceq$ be an admissible order, $\mathcal{A} : L(\mathbb{R})^n \to L(\mathbb{R})$ an aggregation function

for intervals and $g : \mathbb{R}^2 \to (0, \overline{a}]$ increasing and continuous. The function $f_{\vec{\mathbf{x}}} : L(\mathbb{R}) \to [0, 1]$ given by:

$$P(\vec{\mathbf{x}}, \mathbf{y}) = f_{\vec{\mathbf{x}}}(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}_i = \mathbf{y}, \forall i = \{1, \ldots, n\} \\ g\left(\underline{y}, \overline{y}\right) & \text{if } \mathbf{y} \succ \mathcal{A}(\mathbf{x}_1, ..., \mathbf{x}_n) \\ \overline{a} & \text{otherwise} \end{cases}$$

is a penalty function for intervals in $L([\underline{a}, \overline{a}])$.

Proof: We have to check whether the axioms of Definition 4.1 are fulfilled:

P1: It is clear by definition that $P(\vec{\mathbf{x}}, \mathbf{y}) \geq 0$ for all $\mathbf{x} \in L([\underline{a}, \overline{a}])^n$, $\mathbf{y} \in L([\underline{a}, \overline{a}])$

P2: If $\mathbf{x}_i = \mathbf{y}$ for all $i \in \{1, \ldots, n\}$, then $P(\vec{\mathbf{x}}, \mathbf{y}) = 0$.

***Example** 4.1:* Let us suppose that we have these ten intervals:

$$\mathbf{x}_1 = [1, 8] \quad \mathbf{x}_2 = [3, 5] \quad \mathbf{x}_3 = [2, 7]$$

$$\mathbf{x}_4 = [4, 10] \quad \mathbf{x}_5 = [6, 12] \quad \mathbf{x}_6 = [9, 15]$$

$$\mathbf{x}_7 = [8, 14] \quad \mathbf{x}_8 = [5, 12] \quad \mathbf{x}_9 = [7, 13]$$

$$\mathbf{x}_{10} = [10, 20]$$

Since $\underline{a} = \min\{\underline{x_i}\}_{i=1}^{10} = 1$ and $\overline{a} = \max\{\overline{x_i}\}_{i=1}^{10} = 20$, the space we are working on is $L([1, 20])$.

If we consider the penalty function $f_{\vec{\mathbf{x}}} : L([1, 20]) \to [0, 1]$ given by:

$$f_{\vec{\mathbf{x}}}(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}_i = \mathbf{y}, \forall i = \{1, \ldots, 10\} \\ g\left(\underline{y}, \overline{y}\right) & \text{if } \mathbf{y} \succ \mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) \\ 20 & \text{otherwise} \end{cases}$$

where

$$\mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = \left[\frac{\sum_{i=1}^{10} \underline{\mathbf{x}_i}}{10}, \frac{\sum_{i=1}^{10} \overline{\mathbf{x}_i}}{10}\right]$$

and

$$g(\underline{y}, \overline{y}) = \frac{\underline{y} + \overline{y}}{2}.$$

It is easy to see that $\mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = [5.5, 11.6]$.

Table I shows the penalty values using lexicographical order type 1, lexicographical order type 2 and Xu and Yager order, respectively. In each example, the intervals with the least penalization are $[6, 12]$, $[5, 12]$ and $[6, 12]$, respectively.

***Example** 4.2:* Let us suppose we have the ten previous intervals and the penalty function $f_{\mathbf{x}} : L([1, 20]) \to [0, 1]$ given by:

$$f_{\vec{\mathbf{x}}}(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}_i = \mathbf{y}, \forall i = \{1, \ldots, 10\} \\ g\left(\underline{y}, \overline{y}\right) & \text{if } \mathbf{y} \succ \mathcal{A}_{min}(\mathbf{x}_1, ..., \mathbf{x}_{10}) \\ 20 & \text{otherwise} \end{cases}$$

TABLE I
VALUES OF THE PENALTY FUNCTION USING $\mathcal{A}_{mean}$

	i	1	2	3	4	5
	$\mathbf{x}_i$	$[1, 8]$	$[3, 5]$	$[2, 7]$	$[4, 10]$	**[6, 12]**
Lex1	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	20	20	20	20	**9**
Lex2	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	20	20	20	20	9
XY	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	20	20	20	20	**9**

	i	6	7	8	9	10
	$\mathbf{x}_i$	$[9, 15]$	$[8, 14]$	$[5, 12]$	$[7, 13]$	$[10, 20]$
Lex1	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	13	11	20	10	15
Lex2	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	13	11	**8.5**	10	15
XY	$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	13	11	20	10	15

where

$$\mathcal{A}_{min}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = \left[\min_{i=\{1,...,10\}}\{\underline{\mathbf{x}_i}\}, \min_{i=\{1,...,10\}}\{\overline{\mathbf{x}_i}\}\right]$$

and

$$g(\underline{y}, \overline{y}) = \frac{\underline{y} + \overline{y}}{2}.$$

It is easy to see that $\mathcal{A}_{min}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = [1, 6]$.

As all the intervals are greater than $[1, 6]$, we obtain the same values for the three orders, which are displayed in Table II

TABLE II
VALUES OF THE PENALTY FUNCTION USING $\mathcal{A}_{min}$

i	1	2	3	4	5
$\mathbf{x}_i$	$[1, 8]$	**[3, 5]**	$[2, 7]$	$[4, 10]$	$[6, 12]$
$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	4.5	**4**	4.5	7	9

i	6	7	8	9	10
$\mathbf{x}_i$	$[9, 15]$	$[8, 14]$	$[5, 12]$	$[7, 13]$	$[10, 20]$
$f_{\vec{\mathbf{x}}}(\mathbf{x}_i)$	13	11	8.5	10	15

The case where there is not just one minimizer can appear as we will see in the next example.

***Example** 4.3:* Let us change the fifth interval of the previous examples:

$$\mathbf{x}_1 = [1, 8] \quad \mathbf{x}_2 = [3, 5] \quad \mathbf{x}_3 = [2, 7]$$

$$\mathbf{x}_4 = [4, 10] \quad \mathbf{x}_5 = [6, 14] \quad \mathbf{x}_6 = [9, 15]$$

$$\mathbf{x}_7 = [8, 14] \quad \mathbf{x}_8 = [5, 12] \quad \mathbf{x}_9 = [7, 13]$$

$$\mathbf{x}_{10} = [10, 20]$$

We are working again in $L([1, 20])$.

Let us consider the penalty function $f_{\vec{\mathbf{x}}} : L([1, 20]) \to [0, 1]$ given by:

$$f_{\vec{\mathbf{x}}}(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}_i = y, \forall i = \{1, \ldots, 10\} \\ g\left(\underline{y}, \overline{y}\right) & \text{if } \mathbf{y} \succ \mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) \\ 20 & \text{otherwise} \end{cases}$$

where

$$\mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = \left[\frac{\sum_{i=1}^{10} \underline{\mathbf{x}_i}}{10}, \frac{\sum_{i=1}^{10} \overline{\mathbf{x}_i}}{10} \right]$$

and

$$g(\underline{y}, \overline{y}) = \frac{\underline{y} + \overline{y}}{2}.$$

It is easy to see that $\mathcal{A}_{mean}(\mathbf{x}_1, ..., \mathbf{x}_{10}) = [5.5, 11.8]$.

The penalty values using lexicographical order type 1 are shown in Table III, where the intervals with the least penalization are $[6, 14]$, and $[7, 13]$.

In the Examples 4.1 and 4.2 there is only one minimizer, so it is clear which interval is chosen. On the other hand, in Example 4.3 the minimum of the penalty function is not unique. In this case, the interval selected depends on the criterion of the researcher. However, we recognize the importance of further research and development in a more robust technique to choose just one minimizer.

TABLE III
VALUES OF THE PENALTY FUNCTION USING LEXICOGRAPHICAL ORDER TYPE 1 AND $\mathcal{A}_{mean}$

i	1	2	3	4	5
$\mathbf{x}_i$	$[1, 8]$	$[3, 5]$	$[2, 7]$	$[4, 10]$	**[6, 14]**
$f_{\overline{\mathbf{x}}}(\mathbf{x}_i)$	20	20	20	20	**10**

i	6	7	8	9	10
$\mathbf{x}_i$	$[9, 15]$	$[8, 14]$	$[5, 12]$	**[7, 13]**	$[10, 20]$
$f_{\overline{\mathbf{x}}}(\mathbf{x}_i)$	13	11	20	**10**	15

V. CONCLUSIONS

In this study, we explored the concept of penalty functions for interval-valued data, extending their application from vectors to vectors of intervals using admissible orders and aggregation functions. Through illustrative examples, we demonstrated the versatility and effectiveness of these functions in handling interval-valued data, particularly in scenarios characterized by imprecision and uncertainty. It is important to note that the definition proposed in this study specifically applies to intervals contained between all the intervals considered. This perspective could be useful in some context where we want to summarize a set with one value of the sets, such as when the median is used.

As future work, we aim to extend the definition of penalty functions for intervals to encompass any interval on the real number line, rather than confining it to the original intervals considered. This extension would broaden the applicability of penalty functions and provide a more comprehensive framework for handling interval-valued data in various domains. An interesting aproach could be extending the notions of lower semicontinuous and quasiconvexity to intervals, and then guarantee the existence of an interval that minimizes the penalty function considered.

REFERENCES

[1] G. Beliakov, H. Bustince, and T. Calvo. *A Practical Guide to Averaging Functions*. Springer, 2016.

[2] G. Beliakov, A. Pradera, T. Calvo, et al. *Aggregation functions: A guide for practitioners*, volume 221. Springer, 2007.

[3] H. Bustince. Interval-valued fuzzy sets in soft computing. *International Journal of Computational Intelligence Systems*, 3(2):215 – 222, 2010.

[4] H. Bustince, G. Beliakov, G. Pereira Dimuro, B. Bedregal, and R. Mesiar. On the definition of penalty functions in data aggregation. *Fuzzy Sets and Systems*, 323:1–18, 2017. Theme: Aggregation Functions.

[5] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar. Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems*, 220:69–77, 2013. Theme: Aggregation functions.

[6] T. Calvo and G. Beliakov. Aggregation functions based on penalties. *Fuzzy Sets and Systems*, 161(10):1420–1436, 2010. Theme: Aggregation functions.

[7] T. Calvo, R. Mesiar, and R.R. Yager. Quantitative weights and aggregation. *IEEE Transactions on Fuzzy Systems*, 12(1):62–69, 2004.

[8] J. Fumanal-Idocin, Y. Wang, C. Lin, J. Fernández, J. Sanz, and H. Bustince. Motor-imagery-based brain–computer interface using signal derivation and aggregation functions. *IEEE Transactions on Cybernetics*, 52(8):7944–7955, 2021.

[9] B. Gleb and J. Simon. A penalty-based aggregation operator for non-convex intervals. *Knowledge-Based Systems*, 70:335–344, 2014.

[10] J. A. Goguen. L-fuzzy sets. *Journal of mathematical analysis and applications*, 18(1):145–174, 1967.

[11] R. Mesiar M. Grabisch, J. Marichal and E. Pap. Aggregation functions. *Cambridge University Press*, 2009.

[12] R. Mesiar and M. Komorníková. Aggregation functions on bounded posets. In Chris Cornelis, Glad Deschrijver, Mike Nachtegael, Steven Schockaert, and Yun Shi, editors, *35 Years of Fuzzy Set Theory: Celebratory Volume Dedicated to the Retirement of Etienne E. Kerre*, pages 3 – 17, Berlin, Heidelberg, 2011. Springer.

[13] Z.S. Xu and R.R. Yager. Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems*, 35(4):417 – 433, 2006.

[14] R. R. Yager. Toward a general theory of information aggregation. *Information Sciences*, 68(3):191–206, 1993.

[15] R. R. Yager and A. Rybalov. Understanding the median as a fusion operator. *International Journal of General Systems*, 26(3):239–263, 1997.

[16] R. Yera, A. Alzahrani, and L. Martínez. A fuzzy content-based group recommender system with dynamic selection of the aggregation functions. *International Journal of Approximate Reasoning*, 150:273–296, 2022.

XXI Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)

Trabajos Destacados

Datil: Herramienta para el aprendizaje de tipos datos difusos en ontologías difusas

Ignacio Huitzil, Fernando Bobillo
Aragon Institute of Engineering Research (I3A)
University of Zaragoza, Zaragoza, Spain
{ihuitzil,fbobillo}@unizar.es

***Abstract*—Este artículo resume resultados recientes [1] sobre el aprendizaje de tipos de datos difusos en ontologías difusas. Concretamente, se describe brevemente la herramienta Datil y se discuten algunos ejemplos de aplicaciones reales.**

***Index Terms*—Ontologías difusas, aprendizaje, agrupamiento**

I. Motivación

Uno de los factores del reciente éxito de la Inteligencia Artificial es la posibilidad de usar conocimiento. El estándar *de facto* actual para la representación del conocimiento de un dominio de interés en aplicaciones inteligentes son las ontologías. A pesar de su éxito, las ontologías clásicas no son suficientes para manejar conocimiento impreciso, lo que es necesario en numerosas aplicaciones del mundo real. Para superar esta limitación, se han propuesto las *ontologías difusas* [2], que extienden las ontologías clásicas con elementos de la lógica difusa. Las ontologías clásicas describen un dominio mediante clases, instancias de dichas clases, propiedades que relacionan parejas de instancias o asocian una instancia a un valor de un tipo de dato y axiomas sobre los elementos anteriores (por ejemplo, inclusión de una subclase en una superclase). En las ontologías difusas, las clases se representan mediante conjuntos difusos, las propiedades mediante relaciones difusas, los axiomas pueden cumplirse parcialmente (en un cierto grado) y puede haber *tipos de datos difusos* (descritos mediante conjuntos difusos triangulares, trapezoidales, etc.).

Las ontologías difusas se han usado en varias aplicaciones y hay algunos recursos disponibles: lenguajes como Fuzzy OWL 2 [3], razonadores como fuzzyDL [4] y metodologías de desarrollo. Sin embargo, existen pocas ontologías difusas públicamente disponibles y uno de los posibles motivos puede ser la escasez de herramientas que faciliten su aprendizaje.

Usualmente, se asume que existe un experto humano encargado de definir los tipos de datos difusos. En este artículo nos centraremos en un caso diferente, donde no hay expertos disponibles pero sí ejemplos de datos numéricos que se pueden usar para el aprendizaje. Concretamente, presentaremos un resumen de [1], donde se propone una nueva estrategia para el aprendizaje de tipos de datos difusos en ontologías difusas.

II. Herramienta Datil

Datil (del inglés, "DATatypes with Imprecision Learner") es una herramienta software que implementa un algoritmo propio de aprendizaje de tipos de datos difusos para propiedades con un rango de valor numérico. La aplicación está públicamente disponible en http://webdiis.unizar.es/~ihvdis/Datil.

La idea principal consiste en aplicar un algoritmo de agrupamiento basado en centroides y usarlos para calcular funciones de pertenencia difusa que formen una partición del dominio. Concretamente, los centroides se pueden usar como los parámetros de una función izquierda (*left-shoulder*), varias funciones triangulares y una función derecha (*right-shoulder*). La Figura 1 ilustra el proceso: a partir de un vector que incluya todos los valores de una cierta propiedad P (de rango numérico) para todos los individuos de la ontología, se aplica un algoritmo de agrupamiento que produce un vector de centroides (en el ejemplo, $\{10, 20, 30\}$) y, a partir de ellos, se construyen las funciones de pertenencia. En el ejemplo: una función izquierda $\mathbf{left}(10, 20)$, que corresponde a *BajoP*, una función triangular $\mathbf{tri}(10, 20, 30)$, que corresponde a *NeutroP*, y una función derecha $\mathbf{right}(20, 30)$, que corresponde a *AltoP*.

Las principales contribuciones de Datil son las siguientes:

- Es independiente del dominio de aplicación y puede usarse en múltiples escenarios, siempre que los atributos sean numéricos y haya un gran volumen de datos.
- Admite diferentes formatos de archivos de entrada: CSV, FDL (el formato del razonador fuzzyDL) y OWL, que abarca OWL 2 (para extender ontologías clásicas al caso difuso) y Fuzzy OWL 2 (para enriquecer ontologías difusas). Como salida, admite Fuzzy OWL 2 y FDL.
- Soporta tres conocidos algoritmos de agrupamiento no supervisado: k-medias, c-medias difusas y mean shift. Podría ampliarse para admitir cualquier algoritmo de agrupación que devuelva un conjunto de centroides.
- Asigna automáticamente nombres legibles a las etiquetas de los tipos de datos difusos calculados, combinando modificadores como *Muy*, las etiquetas básicas *Bajo*, *Neutro* y *Alto*, y el nombre de la propiedad P.
- Permite evitar durante el aprendizaje datos atípicos (*outliers*) o con valor cero (pues a veces se usan en ficheros CSV para representar un valor desconocido), así como restringirse a un subconjunto (*segmento*) de los valores.
- Un razonador semántico (no difuso) recupera valores de las propiedades representados explícita e implícitamente.
- La herramienta tiene una interfaz gráfica de usuario intuitiva para ordenadores de escritorio pero también una versión móvil para dispositivos Android.
- Propone una versión incremental del algoritmo para per-

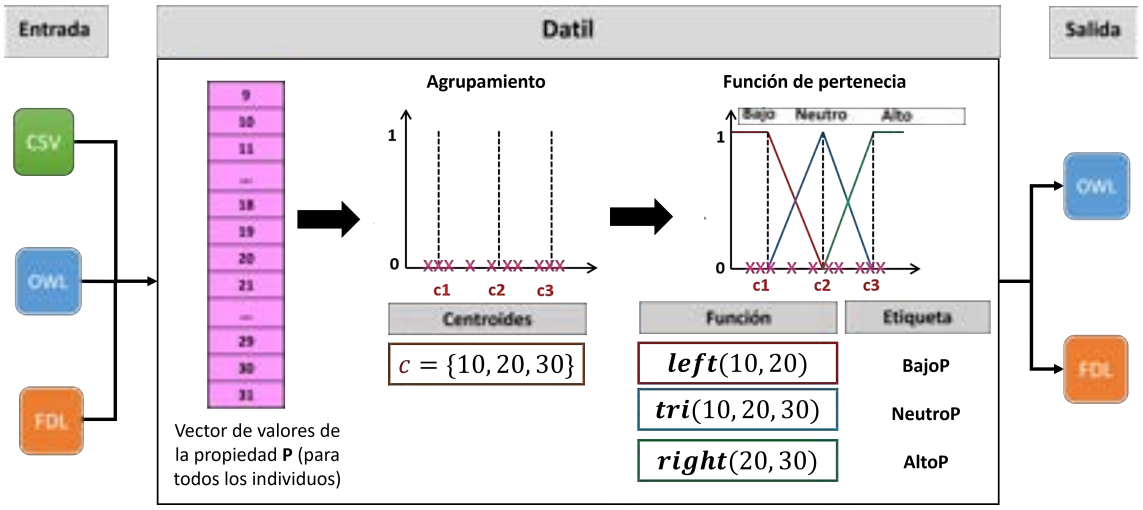


Fig. 1. Esquema general del aprendizaje de tipos de datos difusos

mitir aprendizaje dinámico, que implica cambios en la definición de los tipos de datos difusos de Fuzzy OWL 2.

III. Evaluación

Datil se ha usado con éxito en varias aplicaciones reales:

- Identificación de estilos de vida de los usuarios (como "TrabajadorMediterráneaneo") a partir de datos de sensores *wearable* (como el horario tardío y la duración larga del almuerzo) [5]. Datil calculó tipos de datos difusos, como "AlmuerzoLargo", que se usaron después para aprender definiciones de clases complejas basadas en ellos. Para validar la propuesta, se usaron datos reales de voluntarios, proporcionados por una empresa privada.
- Recomendación de cervezas en el sistema GimmeHop, para dispositivos Android y basado en ontologías difusas [6]. Datil aprendió a partir de datos reales etiquetas lingüísticas para la cantidad de alcohol o el amargor de las cervezas, aptas para las consultas al sistema.
- Reconocimiento de la marcha humana [7]. Dada una secuencia de datos de una persona caminando, se devuelve el individuo más similar conocido o se detecta que pertenece a un individuo desconocido. Datil permite mejorar la interpretabilidad del sistema, usando términos lingüísticos al explicar la decisión del sistema.
- Diagnóstico de la hepatitis C, donde investigadores independientes usaron Datil para construir tipos de datos difusos en un sistema de soporte a la toma de decisiones clínicas basado en una ontología difusa y una red bayesiana difusa [8].

También se han evaluado la calidad y el rendimiento:

- Para evaluar la calidad de los tipos de datos difusos aprendidos para la recomendación de cervezas, un grupo de aficionados clasificó el nivel de alcohol de algunas cervezas y se compararon con los resultados calculados por Datil, para diferentes algoritmos de agrupamiento y parámetros. Los mejores resultados se obtuvieron con mean-shift, usando un valor de umbral máximo para reducir datos atípicos.
- También se comparó el tiempo de aprendizaje en la versión móvil de Datil y en la versión de escritorio, para diferentes tamaños de ontologías. Aunque el tiempo es mayor en el teléfono de gama media-baja considerado, podría ser aceptable pues el aprendizaje típicamente se realiza una única vez. Además, el algoritmo de agrupamiento tiene un impacto: por ejemplo, mean-shift puede llegar a ser un 20% más lento que k-medias.

Agradecimientos

Hemos sido parcialmente financiados por los proyectos de I+D+i PID2020-113903RB-I00 (MCIN/AEI/10.13039/501100011033) y T42_23R (Gobierno de Aragón).

References

[1] I. Huitzil and F. Bobillo, "Fuzzy ontology datatype learning using Datil," *Expert Systems With Applications*, vol. 228, p. 120299, 2023.
[2] F. Zhang, J. Cheng, and Z. Ma, "A survey on fuzzy ontologies for the semantic web," *Knowledge Engineering Review*, vol. 31, no. 3, pp. 278–321, 2016.
[3] F. Bobillo and U. Straccia, "Fuzzy ontology representation using OWL 2," *International Journal of Approximate Reasoning*, vol. 52, no. 7, pp. 1073–1094, 2011.
[4] ——, "The fuzzy ontology reasoner fuzzyDL," *Knowledge-Based Systems*, vol. 95, pp. 12–34, 2016.
[5] N. Díaz-Rodríguez, A. Härmä, R. Helaoui, I. Huitzil, F. Bobillo, and U. Straccia, "Couch potato or gym addict? Semantic lifestyle profiling with wearables and knowledge graphs," in *Procs. of 6th NIPS Workshop on Automated Knowledge Base Construction (AKBC 2017)*, 2017.
[6] I. Huitzil, F. Alegre, and F. Bobillo, "GimmeHop: A recommender system for mobile devices using ontology reasoners and fuzzy logic," *Fuzzy Sets and Systems*, vol. 401, pp. 55–77, 2020.
[7] I. Huitzil, L. Dranca, J. Bernad, and F. Bobillo, "Gait recognition using fuzzy ontologies and Kinect sensor data," *International Journal of Approximate Reasoning*, vol. 113, pp. 354–371, 2019.
[8] I. Riali, M. Fareh, M. C. Ibnaissa, and M. Bellil, "A semantic-based approach for hepatitis C virus prediction and diagnosis using a fuzzy ontology and a fuzzy Bayesian network," *Journal of Intelligent & Fuzzy Systems*, vol. 4, no. 2, pp. 1–15, 2022.

Key work on *'Calculating the interaction index: a polynomial approach based on sampling'*

1 Inmaculada Gutiérrez
Faculty of Statistical Studies
Complutense University of Madrid
inmaguti@ucm.es
0000-0002-7103-393X

2 Javier Castro
Faculty of Statistical Studies
Complutense University of Madrid
jcastroc@estad.ucm.es
0000-0002-5345-8870

3 Daniel Gómez
Faculty of Statistical Studies
Complutense University of Madrid
dagomez@estad.ucm.es
0000-0001-9548-5781

4 Rosa Espínola
Faculty of Statistical Studies
Complutense University of Madrid
rosaev@estad.ucm.es
0000-0003-2336-3659

***Abstract*—This work is about the paper 'Calculating the interaction index: a polynomial approach based on sampling', published in the journal Fuzzy Sets and Systems in 2023. That paper explored the challenge of calculating fuzzy measures indices. Building upon the concept of fuzzy sets, Murofushi and Soneda introduced an interaction index to analyze relationships between two individuals. Grabisch later extended this index within a unified framework. Both indices are foundational in the field of fuzzy measures. However, the calculation process remains highly complex, lacking an approximate solution. To address this gap, we proposed an alternative characterization of the interaction index based on a representation of the Shapley value using orders. This alternative representation simplifies the handling of these indices, offering insights for both pairwise interactions and more complex scenarios involving arbitrary sets. We also proposed two polynomial methods based on sampling to estimate the interaction index, along with a method to approximate its generalized version. Computational experiments were conducted to assess the effectiveness of the proposed algorithms.**

***Index Terms*—Fuzzy Measures, Interaction Index, Interaction Representation, Shapley Value, Sampling Algorithm**

I. Introduction and state of art

The measure theory [1] has had a great deal of importance in the develop of several areas, as the fuzzy measures [2], the necessity and possibility measures [3], [4], the capacity measures [5], etc. Specifically, we focus on fuzzy measures, monotonic set functions which define a type of non-additive measures encompassing many measures such as belief, possibility, necessity or plausibility measures [6]. Fuzzy measures have been widely analyzed [7], [8] and applied in many fields. Unfortunately, the practical implementation of fuzzy measures is complex (2^n coefficients are needed to define a fuzzy measure over n individuals), and the related semantic is difficult to understand [9]. Many researchers have focused on the characterization, interpretation and representation of fuzzy measures, which may be understood as some specific cooperative games for which monotony holds.

Grant Plan Nacional de I+D+i, PID2020-116884GB-I00, PID2021-122905NB-C21.

To understand them, the analysis of the interactions between individuals turns essential. Regarding the importance index Shapley value [10], essential tool in game theory adapted to the field of fuzzy measures, that could be understood as an indication of the importance of each singleton. It has been deeply studied from a theoretical point of view [11], [12], but the calculation of its real value for general fuzzy measures is a NP-problem. Until not that long ago, there were only a few works to approximate it in some specific problems [13], [14]. In 2009, Castro et al. proposed a method based on sampling to approximate the Shapley value [15], generalized in [16] by a stratified random sampling with optimum allocation process.

Regarding interactions among several elements [17], Murofushi and Soneda defined the interactions index between two individuals [18], which quantifies its average contribution considering all the subsets it is part of. Then, Grabisch proposed an extension of it to deal with sets involving more than two elements [19]. As with the Shapley value, the approximation of the real value of these indices has not been studied in depth for general scenarios, due to its great complexity.

In our paper *'Calculating the interaction index of a fuzzy measure: a polynomial approach based on sampling'*, published in 2023 in the journal *Fuzzy Sets and Systems* [20], we proposed an alternative characterization based on orders of this interaction index, for both the simple case related to pairs of elements, and the corresponding extension to deal with any set. This new formulation of the interaction index (and of the representation index), provides a different and intuitive understanding of this it. Then, following the idea of sampling in [15], [16], we defined two methods to approach the value of the interaction index. Both are based on sampling, where the second one is a refinement of the very first, but also using stratified random sampling with optimum allocation.

II. Main Contributions

• **Proposal of an alternative characterization of the interaction index based on orders.** [18] introduced an interaction

index $I_{ij}(\mu)$ based on multi-attribute utility theory to model interactions between pairs of elements i, j regarding the fuzzy measure μ, i.e. to estimate the degree to which a pair of elements interact. $I_{ij}(\mu)$ can be seen as an average of the added value obtained by considering i and j in the same coalition, depending on μ. In fact, considering a singleton, the Shapley value coincides with the interaction index. We worked in an alternative characterization of the interaction index based on orders and permutations, which provides a simpler view of it. On the other hand, in [19], by analogy with the first- and second-order classes, Grabisch proposed a generalization of I_{ij}, $I_T(\mu)$ to represent interactions among the elements of any set T regarding the fuzzy measure μ. This representation index is useful to quantify the average contribution of a coalition when considering all the subsets it is part of. Again, for $T = \{i\}$, recovers $Sh_i(\mu)$, and I_{ij} for $T = \{i, j\}$. We also defined an alternative characterization of I_T based on orders. These characterizations gave us a new view of the interaction indices as an average of the added value obtained by considering them in the same coalition.

• **Development of a new methodology to approach the real value of the interaction index (and the representation index).** The complexity of Shapley value calculation is well-known: it is a $NP-$problem for general fuzzy measures. Although it has been deeply analyzed from a theoretical perspective, until not that long ago, there were only a few works about its real calculation until the proposals in [15], [16]. This lack of practical analysis also occurs with the interaction index. Its calculation can be done in polynomial time only for a few simple fuzzy measure, and there is not much research about feasible estimations of it. We proposed a new polynomial-time methodology to approach the interaction index, developed on the basis of its alternative characterization based on orders. We defined two polynomial-time algorithms to estimate the real value of the interaction index. The first one, named *ApproInteraction*, is based on simple random sampling. We detailed the step-by-step process and provide its pseudo-code for easy adaptation to any programming language. The obtained estimation has some desirable properties, specifically, it is unbiased and consistent in probability.

For situations with large variance caused by simple random sampling in which the *ApproInteraction* algorithm is not accurately enough, we proposed a refinement with the use of stratified random sampling named *StratifiedApproInteraction*, which improves the performance of the original method for cases with high variance. The key is to divide the whole population into subpopulations, each of which is internally homogeneous and as heterogeneous as possible with respect the others groups. We showed a detailed explanation of the *StratifiedApproInteraction* algorithm, as well as its pseudo-code. Again, the estimated value has desirable statistical properties, as being unbiased and consistent in probability.

Because of the lack of methods to estimate the interaction index, we could not compare our algorithms with others. Then, we calculated the real value of I for several simple fuzzy measures for which this calculation is feasible in polynomial-time (considering the alternative characterization based on orders), and we compared the real value with the estimations. Both methods were so accurate and the results were exact and reliable, so we could confirm the goodness of our methodology. Although we thoroughly explained and test the methods with respect to the interaction index, the stratified procedure could be generalized in much the same way to estimate other values, specifically for the representation index I_T. The process was summarized by its pseudo-code in the full text, named *Stratified ApproInteraction T* algorithm.

Although we focused on fuzzy measures, all the contributions may be adapted to a more general scenario related to cooperative game (as fuzzy measures can be understood as a subset of cooperative games with monotony).

REFERENCES

[1] P. Halmos, *Measure Theory*. Springer, 1950.

[2] M. Sugeno, "Fuzzy measures and fuzzy integrals: A survey," *Fuzzy Automata and Decision Processes*, vol. 78, pp. 89–102, 1977.

[3] L. A. Zadeh, "Similarity relations and fuzzy orderings," *Information Sciences*, vol. 3, pp. 177–200, 1971.

[4] L. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, pp. 3–28, 1978.

[5] G. Choquet, "Theory of capacities," *Annales de l'Institut Fourier*, vol. 5, pp. 131–295, 1954.

[6] M. Grabisch, T. Murofushi, and M. Sugeno, "Fuzzy measure of fuzzy events defined by fuzzy integrals," *Fuzzy Sets and Systems*, vol. 50, no. 3, pp. 293–313, 1992.

[7] M. Grabisch, "The representation of importance and interaction of features by fuzzy measures," *Patter Recognition Letters*, vol. 17, no. 6, pp. 567–575, 1996.

[8] G. Beliakov and J. Wu, "Learning k-maxitive fuzzy measures from data by mixed integer programming," *Fuzzy Sets and Systems*, vol. 412, pp. 41–52, 2021.

[9] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Elsevier, Ed., Amsterdam, The Netherlands, 1980.

[10] L. Shapley, "A value for n-person games," *Contribute. Theory Games*, vol. 2, no. 28, pp. 307–317, 1953.

[11] M. Ballasote, C. Hernández-Mancera, and A. Jiménez-Losada, "A new Shapley value for games with fuzzy coalitions," *Fuzzy Sets and Systems*, vol. 383, pp. 51–67, 2020.

[12] J. M. Gallardo and A. Jiménez-Losada, "A characterization of the Shapley value for cooperative games with fuzzy characteristic function," *Fuzzy Sets and Systems*, vol. 398, pp. 98–111, 2020.

[13] W. Cochran, *Sampling Techniques*. Wiley, 1977.

[14] H. Lohr, *Sampling: Design and Analysis*. Duxbury, 1999.

[15] J. Castro, D. Gómez, and J. Tejada, "Polynomial calculation of the Shapley value based on sampling," *Computers & Operations Research*, vol. 36, no. 5, pp. 1726–1730, 2009.

[16] J. Castro, D. Gómez, E. Molina, and J. Tejada, "Improving polynomial estimation of the Shapley value by stratified random sampling with optimum allocation," *Computers & Operations Research*, vol. 82, pp. 108–188, 2017.

[17] I. Kojadinovic, "A weight-based approach to the measurement of the interaction among criteria in the framework of aggregation by the bipolar Choquet integral," *European Journal of Operational Research*, vol. 179, no. 2, pp. 498–517, 2007.

[18] T. Murofushi and S. Soneda, "Techniques for reading fuzzy measures (III): interaction index," in *Proc. of 9^{th} Fuzzy System Symposium*. Sapporo, Japan: 9^{th} Fuzzy System Symposium, 1993, pp. 693–696.

[19] M. Grabisch, "k-order additive discrete fuzzy measures and their representation," *Fuzzy Sets and Systems*, vol. 92, no. 2, pp. 167–189, 1997.

[20] I. Gutiérrez, J. Castro, D. Gómez, and R. Espínola, "Calculating the interaction index of a fuzzy measure: A polynomial approach based on sampling," *Fuzzy Sets and Systems*, vol. 466, p. 108454, 2023.

A Supervised Approach to Community Detection Problem: How to Improve Louvain Algorithm by Considering Fuzzy Measures. A Review

María Barroso[1,3], Daniel Gómez[1], Inmaculada Gutiérrez[1]
[1]*Faculty of Statistics, Complutense University Puerta de Hierro, s/n, 28040 Madrid, Spain*
[3]E-mail: mbarro10@ucm.es

***Abstract*—Social Network Analysis has evolved significantly, driven by real-world data and advancements in analytics. The focus on addressing community detection problems has prompted to refinements in classical algorithms like Louvain, aimed at overcoming inherent challenges. This research introduced a novel supervised technique, leveraging the Louvain algorithm and flow extended fuzzy graphs. Our approach, a parametric and aggregation supervised method, surpasses limitations of local solutions, marking a substantial advancement beyond prior results. Operating within the machine learning paradigm and considering directed modularity, this study signified a noteworthy progression in community detection algorithms. Rigorous evaluation across diverse benchmark and real-world networks underscores the efficacy of our supervised technique, positioning it favorably among existing algorithms.**

***Index Terms*—Community Detection, Complex Networks, Flow Extended Fuzzy Graph, Flow Capacity Louvain, Directed Modularity**

I. Introduction

In this review, we summarise our research entitled "A supervised approach to community detection problem: How to improve Louvain algorithm by considering fuzzy measures", published in Springer in 2022 [1]. Where we introduced a supervised technique aimed at enhancing clustering methods and the quality measure of resulting partitions in Community Detection Problems (CDP) within complex networks. While the search for the 'best community structure technique' has predominantly focused on the structural information of a graph [2], [3], our work addresses the limitations of such local approaches. Notably, algorithms like the Louvain algorithm [4] have gained popularity due to their speed and effectiveness in large network, but they may encounter challenges in capturing global information.

Our emphasis lies in the enhancement of CDP techniques, particularly through the introduction of the *Flow Capacity Louvain* [5]. This method integrates the structural information of a graph with flow considerations, utilizing the concept of flow extended fuzzy graphs [5]. Given a crisp directed graph $G = (V, E)$ and the fuzzy measure defined as flow capacity measure $\mu^F : 2^V \longrightarrow [0, 1]$ defined over the set of nodes, the triplet $\widetilde{G} = (V, E, \mu^F)$ obtained from considering together the graph with the fuzzy measure which models the flow amongs nodes, is called flow extended fuzzy graph.

As a result, the *Flow Capacity Louvain* is a multi-phase heuristic algorithm based on the directed Louvain [6], incorporating an importance parameter $\alpha \in [0, 1]$ to balance the influence of flow in the partitioning process. The analysis of this aggregation process is meant to be addressed in this work. Therefore, the aggregation of additional information is facilitated by the directed interaction index [5], represented by the matrix I^D, being $\left(I^D_{ij}\right)_{i,j\in V} = \frac{f_{ij}}{\sum_{l,m\in V} f_{lm}}$, the elements of that matrix with manages the flow information among nodes (f_{ij}) modeled by the fuzzy measure μ^F. Then $\widetilde{G}$ is summarized by A^D, the directed adjacency matrix of the graph and I^D. The aggregation process is represented by an aggregation function $\Phi : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, where the matrix $M = \Phi(A^D, I^D)$ is obtained, and the parameter α gives weight to each matrix.

On the other hand, to evaluate the quality of partitions in CDP, it is extensively considered the directed modularity [7] $Q_d(A^D)$, adapted from the traditional modularity [8]. One widely employed technique for characterizing pattern structures is the Louvain algorithm [4] based on modularity maximization and local node movements. In the *Flow Capacity Louvain*, we extended this approach by incorporating the maximization of modularity, considering $\Delta Q^d_i(j)$ with $A^D_{ij} = M_{ij}$. Then, we provided a supervised approach in which we aim to find values of α that yield high modularity values in both undirected and directed networks [7], [8].

II. A supervised approach in Community Detection Problems

In this section, we outline the methodology of our supervised approach for addressing CDP through the *Flow Capacity Louvain* algorithm. Our primary objective involves acquiring essential components for formulating the objective function that maximizes modularity in the associated flow extended fuzzy graph. The algorithm relies on two critical matrices: the adjacency matrix of the crisp graph (A^D) and the directed interaction matrix derived from the flow capacity measure μ^F (I^D). To streamline the computation, we exploit the 2-additivity property of μ^F [5], which significantly reduces computational complexity [9]. The aggregate matrix M is then formed through parametric aggregations, where Φ approximates a linear combination using a weighted sum of both matrices: $M = \Phi(A^D, I^D) = \alpha A^D + (1 - \alpha)I^D$. This approach is specifically chosen for its computational efficiency, given its ability to provide flexibility in capturing community

structures. Additionally, in our exploration of diverse networks drawn from the literature, encompassing both directed and undirected types, we highlight the versatile nature of the *Flow Capacity Louvain* algorithm. Originally devised for directed networks, is adaptable to undirected graphs by treating non-directed edges as pairs of directed edges.

Thereafter, incorporating a machine learning supervised framework, our methodology unfolds in two distinct phases – training and testing. In the training phase, we systematically explore a range of α values, spanning from 0.01 to 0.01 within the interval $[0, 1)$. This exploration involves the execution of the *Flow Capacity Louvain*, generating a set of partitions with 100 α values, each labeled with its corresponding modularity Q_d [7]. The subsequent calculation of the average modularity $Q_d^{\ell} = \frac{\sum_{i=1}^{n} Q_d^{i\ell}}{n}$ for each α^{ℓ} $(\forall \ell = 1 \dots 100)$ allows us to pinpoint the optimal α^* that maximizes modularity.

Moving to the testing phase, the identified optimal α^* undergoes validation on selected networks. Following successful validation, hypothesis testing is employed to statistically affirm the robustness of the optimal α^* determined through our supervised technique. This meticulous and systematic approach guarantees the generalizability and effectiveness of the *Flow Capacity Louvain* algorithm across varied networks, leveraging an appropriate and validated α^*.

III. Evaluation and benchmarking

In this section, we present the computational results to assess the efficacy of the introduced supervised approach in CDP. The evaluation is carried out using the $Python$ software, implementing the two phases elucidated earlier. The outcome of this approach yields distinct community structures within the network contingent upon the value of α. These proposed sets are subsequently evaluated based on the directed modularity measure Q_d [7].

In Figure 1 the average modularity obtained with varying α are compared with the Louvain algorithm [4], specifically when $\alpha = 1$. The depicted values are restricted within the range of 0.5 to 1, where values exhibit a discernible increase in the average modularity across the training set.

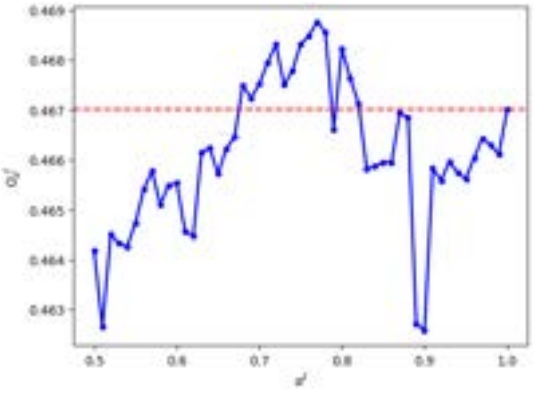


Fig. 1. Summarise the average modularity Q_d^{ℓ}, performing $\alpha^{\ell} \in [0.5, 1]$ $\ell = \{51 \dots 101\}$ for the training set

The optimal α^* value, determined as 0.77, is selected based on its ability to yield consistently high average modularity across diverse datasets. To validate the effectiveness of this selected α^*, the *Flow Capacity Louvain* and the Louvain algorithm are executed on a test dataset. Then hypothesis testing is employed to ascertain the statistical significance of the results, determining whether the average modularity under both algorithms is statistically equal. With the Wilcoxon signed rank test, significance levels were set at 1.25%.

Based on the outcomes of the test, we confidently propose an interval for α^* within $[0.75 - 0.82]$ at the specified significance level. This interval signifies a reliable range where improvements in modularity can be accepted. By incorporating this optimal α^* range, there is a high probability, specifically 98.75%, of enhancing the results obtained by the Louvain algorithm. This underscores the effectiveness and credibility of the proposed supervised technique in elevating community detection outcomes.

IV. Conclusions

Our study offers a comprehensive perspective by extending the supervised approach to other algorithms, contributing significantly to the field of CDP. In particular, [10] introduced a novel proposal aimed at enhancing the overall performance of algorithms in CDP, specifically in directed networks. So that, the incorporation of fuzzy measures and meticulous preprocessing considerations emerges as a promising area, providing valuable insights for optimizing algorithm efficacy.

Acknowledgment

This research has been partially supported by PGC2018096509-B-I00.

References

[1] M. Barroso, D. Gómez, I. Gutiérrez, A supervised approach to community detection problem: How to improve Louvain algorithm by considering fuzzy measures, in: International Conference on Intelligent and Fuzzy Systems, Springer, 2022, pp. 219–227.
[2] S. Fortunato, Community detection in graphs, Physics reports 486 (2010) 75–174.
[3] F. Malliaros, M. Vazirgiannis, Clustering and community detection in directed networks: A survey, Physics Reports 533 (4) (2013) 95–142.
[4] V. Blondel, J. Guillaume, R. Lambiotte, E. Lefevre, Fast unfolding of communities in large networks, Journal Of Statistical Mechanics-Theory And Experiment 10 (P10008) (2008).
[5] M. Barroso, I. Gutiérrez, D. Gómez, J. Castro, R. Espínola, Group definition based on flow in community detection, in: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 2020, pp. 524–538.
[6] L. Li, X. He, G. Yan, Improved Louvain method for directed networks, in: Intelligent Information Processing IX: 10th IFIP TC 12 International Conference, IIP 2018, Nanning, China, October 19-22, 2018, Proceedings 10, Springer, 2018, pp. 192–203.
[7] A. Arenas, J. Duch, A. Fernández, S. Gómez, Size reduction of complex networks preserving modularity, New Journal of Physics 9 (6) (2007) 176.
[8] M. Newman, M. Girvan, Finding and evaluating community structure in networks, Physical Reviews 69 (2004) 026113.
[9] M. Grabisch, K-order additive discrete fuzzy measures and their representation, Fuzzy sets and systems 92 (2) (1997) 167–189.
[10] I. Gutiérrez, M. Barroso, D. Gómez, J. Castro, Social networks analysis and fuzzy measures: a general approach to improve community detection in directed graphs, in: The 20th World Congress of the International Fuzzy Systems Association, 2023, pp. 246–253.

A Fuzzy-set based formulation for Minimum Cost Consensus Models

1st Diego García-Zamora
Dept. Informática
Universidad de Jaén
Jaén, España
dgzamora@ujaen.es

2nd Álvaro Labella
Dept. Informática
Universidad de Jaén
Jaén, España
alabella@ujaen.es

3rd Bapi Dutta
Dept. Informática
Universidad de Jaén
Jaén, España
bdutta@ujaen.es

4th Luis Martínez
Dept. Informática
Universidad de Jaén
Jaén, España
martin@ujaen.es

***Abstract*—This keywork encapsulates the primary points discussed in *A Fuzzy-set based formulation for Minimum Cost Consensus Models* [1]. While Minimum Cost Consensus (MCC) models have been widely utilized in Group Decision-Making (GDM) scenarios, prior extensions of these models lack comprehensive analysis concerning their interrelationships and applicability, thus impeding their practical utility. This paper introduces a redefined approach to MCC models for GDM issues utilizing Fuzzy Set Theory. The proposed fuzzy-based MCC (FZZ-MCC) framework enhances comprehension of MCC models and their extensions while offering a robust and adaptable methodology for tackling diverse GDM challenges.**

***Index Terms*—Minimum cost consensus, Fuzzy sets, Group decision-making, Optimization**

I. Preliminaries

Formally, a group decision-making (GDM) problem comprises a set of $n \in \mathbb{N}$ alternatives $X = \{x_1, x_2, ..., x_n\}$ and a set of $m \in \mathbb{N}$ decision-makers (DMs) $E = \{e_1, e_2, ..., e_m\}$ who assess these alternatives using a specific preference structure. Henceforth, GDM problems will be denoted by their associated pairs (E, X).

Minimum cost consensus (MCC) models are automatic consensus-reaching processes that convert a GDM problem into a mathematical programming problem. Initially proposed by Ben-Arieh and Easton [2], these models are designed to minimize the cost associated with adjusting DMs' preferences, ensuring that the resulting individualized opinions align closely (within a predefined threshold, $\varepsilon \in]0, 1]$) with the group consensus. Given initial preference values $(o_1, o_2, ..., o_m) \in \mathbb{R}^n$ and a cost vector $(c_1, c_2, ..., c_m) \in \mathbb{R}^n_+$, the MCC model is defined as:

$$\begin{aligned} &\min_{o'} \sum_{k=1}^{m} c_k \mid o'_k - o_k \mid \\ &s.t. \ \mid o'_k - \overline{o'} \mid \leq \varepsilon, k = 1, 2, \ldots, m \end{aligned} \tag{MCC}$$

where $(o'_1, \ldots, o'_m)$ are the adjusted opinions of the DMs, $\overline{o'}$ represents the collective opinion, and $\varepsilon \in]0, 1]$ is the maximum absolute deviation of each DM and the collective opinion.

This work is partially supported by the Junta de Andalucía Andalusian Plan for Research, Development, and Innovation (POSTDOC 21-00461) and by the Grants for the Requalification of the Spanish University System for 2021-2023 in the María Zambrano modality (UJA13MZ).

It is possible to find many extensions of MCC models [3], [4], all of them with common elements (see Fig. 1). However, there is a lack of a unified framework that allows systematically generating new ones.

II. Generalized Minimum Cost Consensus

Here it is presented a reformulation of MCC models considering a new definition of their elements [5] according to Fuzzy Set Theory.

A. Main elements of MCC models

The following key elements are generalized (see Fig. 1):

1) Preference Structure: Refers to the format used by the DMs to give their preferences.

Definition 1 (Discrete Numeric Preference Structure): Given a GDM problem (E, X), a numeric preference structure for the alternative set X is a subset $\mathbb{P} \subset \mathcal{F}(X^d)$, where $d = 1, 2$.

Example 1: According to the previous definitions, the Discrete Numeric Preference Structure corresponding to Fuzzy Preference Relations (FPRs) [6] is:

$$\mathbb{P}_A := \{P \in \mathcal{F}(X \times X) \ : \ P(x_i, x_j) + P(x_j, x_i) = 1, \\ i, j = \{1, 2, ..., n\}\}$$

We additionally characterize the DMs' preferences with a fuzzy set on $E \times X^d$.

Definition 2 (Numeric Ratings associated to $\mathbb{P}$): Given a GDM problem (E, X) and a discrete numeric preference structure $\mathbb{P}$, a numeric rating is a fuzzy set $P : (E, X^d) \to [0, 1]$ such that for every $k = 1, 2, ..., m$, $P(e_k, \cdot) \in \mathbb{P}$. The set containing all the possible numeric ratings for the GDM (E, X) and the preference structure $\mathbb{P}$, which is a subset of $\mathcal{F}(E, X^d)$, will be denoted by $\mathcal{P}$ and called the numeric ratings set associated to $\mathbb{P}$.

Example 2: The numeric ratings set for FPRs is given by:

$$\mathcal{P}_A := \{P \in \mathcal{F}(E, X \times X) : P(e_k, x_i, x_j) + P(e_k, x_j, x_i) = 1, \\ \forall \ i, j = \{1, 2, ..., n\} \ \forall \ k = \{1, 2, ..., m\}\}.$$

2) Aggregation of information: The primary goal of a GDM problem is to formulate a cohesive collective opinion by combining the preferences of all DMs.

Definition 3 (Collective Opinion): Let us consider a GDM problem (E, X), a preference structure $\mathbb{P} \subset \mathcal{F}(X^d)$ and an

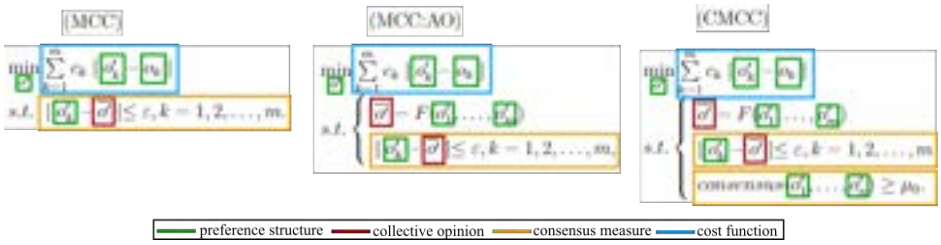


Fig. 1. Structure of classical MCC [2], MCC with generic aggregations [3], and Comprehensive MCC [4].

aggregation operator $M : [0,1]^m \to [0,1]$. For $P \in \mathcal{P}$, the fuzzy set $\overline{P} : X^d \to [0,1]$ defined as $\overline{P}(x_i) := M^m_{k=1} P(e_k, x_i)\ \forall\ i = 1,2,...,n^d$ is called the collective opinion of P under M. When an aggregation operator $M : [0,1]^m \to [0,1]$ preserves the preference structure, i.e. $\overline{P} \in \mathbb{P}$, the respective induced collective opinion operator $M : \mathbb{P}^m \to \mathbb{P}$ is well-defined.

Example 3: For FPRs, the collective opinion may be given by $A : \mathbb{P}^m_A \to \mathbb{P}_A$:

$$A(P_1,...,P_m)(x_i,x_j) = \frac{1}{m}\sum_{k=1}^{m} P_k(x_i,x_j)$$

$\forall\ i,j \in \{1,2,...,n\}\ ,\ \forall\ P_1,...,P_m \in \mathbb{P}_A$.

3) Consensus measurement: To reconceptualize the notion of a consensus measure using fuzzy sets, the main aim is to calculate a sort of distance between the opinions held by DMs.

Definition 4 (Consensus measure): Given a GDM problem (E,X) and a preference structure $\mathbb{P}$, the consensus measure for $\mathcal{P}$ associated to an aggregation operator $\hat{M} : [0,1]^{mn^d} \to [0,1]$ and the restricted dissimilarity function δ is a function $\kappa : \mathcal{P} \to [0,1]$ defined as

$$\kappa(P) = \hat{M}_{k=1,...,m;i=1,...n^d}\delta(P(e_k,x_i),\overline{P}(x_i))\ \forall\ P \in \mathcal{P},$$

where $\overline{P} \in \mathbb{P}$ is the collective opinion of P under an aggregation operator $M : [0,1]^m \to [0,1]$.

Example 4: As an example of consensus measure for FPRs, we can consider the function $\kappa : \mathcal{P}_A \to [0,1]$

$$\kappa(P) = \sum_{k=1}^{m}\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{n(n-1)} \mid P(e_k,x_i,x_j) - \overline{P}(x_i,x_j) \mid,$$

for $P \in \mathcal{P}_A$.

4) Cost function: A cost function quantifies the weighted disparity between the original opinions of DMs and their adjusted ones.

Definition 5 (Cost function): Given a GDM problem (E,X), a preference structure $\mathbb{P}$, an initial numeric rating $P_0 \in \mathcal{P}$, a relative cost vector $c \in [0,1]^m$ satisfying $\sum_{k=1}^{m} c_k = 1$ and a distance measure $D : \mathbb{P} \times \mathbb{P} \to [0,1]$. A cost function is a mapping $\xi_c : \mathcal{P} \to [0,1]$ defined as $\xi_c(P,P_0) = \sum_{k=1}^{m} c_k D(P(e_k,\cdot),P_0(e_k,\cdot))$, For simplicity, we will use the notation $\xi : \mathcal{P} \to [0,1]$ to denote a cost function in which all the relative costs are equal, i.e., $c_k = \frac{1}{m}\ \forall\ k = 1,2,...,m$.

Example 5: A cost function for FPRs is the function $\xi : \mathcal{P}_A \to [0,1]$

$\xi(P) = \sum_{k=1}^{m}\frac{1}{m}\sum_{i,j=1}^{n}\frac{1}{n(n-1)} \mid P(e_k,x_i,x_j) - P_0(e_k,x_i,x_j) \mid$

for $P \in \mathcal{P}_A$, where P_0 is an initial numeric rating.

B. Fuzzy-Set-based formulation for MCC models

In this section, we present a unified global model that incorporates the Fuzzy-Set-based reinterpretations of the classical elements found in MCC models mentioned earlier. This method extends previous suggestions and simplifies the process of adapting new MCC models to tackle particular real-world challenges.

Definition 6 (Fuzzy-Set-based MCC model): Let (E,X) be a GDM problem and a preference structure $\mathbb{P}$. Given an initial numeric rating $P_0 \in \mathcal{P}$, a cost function $\xi_c : \mathcal{P} \to [0,1]$ for P_0 and a family of $q \in \mathbb{N}$ consensus metrics $\kappa = (\kappa_1,...,\kappa_q) : \mathcal{P} \to [0,1]^q$, the corresponding Fuzzy-Set-based MCC model is given by:

$$\begin{aligned} &\min_{P\in\mathcal{P}}\ \xi_c(P,P_0) \\ &s.t.\ \ \kappa(P) \leq (\varepsilon_1,...,\varepsilon_q), \end{aligned} \qquad \text{(FZZ-MCC)}$$

where $(\varepsilon_1,...,\varepsilon_q) \in [0,1]^q$ is the vector of desired parameters to control the q consensus metrics.

III. Conclusions

This keywork has generalized classical MCC models by introducing a new formulation based on fuzzy set theory. By using this general version, it is possible to derive new MCC models that may be tailored to specific decision scenarios.

References

[1] D. García-Zamora, B. Dutta, Á. Labella, and L. Martínez, "A fuzzy-set based formulation for minimum cost consensus models," *Computers & Industrial Engineering*, vol. 181, p. 109295, 2023.

[2] D. Ben-Arieh and T. Easton, "Multi-criteria group consensus under linear cost opinion elasticity," *Decision Support Systems*, vol. 43, no. 3, pp. 713–721, 2007, integrated Decision Support.

[3] G. Zhang, Y. Dong, Y. Xu, and H. Li, "Minimum-cost consensus models under aggregation operators," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, no. 6, pp. 1253–1261, 2011.

[4] A. Labella, H. Liu, R. M. Rodríguez, and L. Martínez, "A cost consensus metric for consensus reaching processes based on a comprehensive minimum cost model," *European Journal of Operational Research*, vol. 281, p. 316–331, 2020.

[5] L. A. Zadeh, "Fuzzy sets," in *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh*, G. Klir and B. Yuan, Eds. World Scientific, 1996, pp. 394–432.

[6] S. Orlovsky, "Decision-making with a fuzzy preference relation," *Fuzzy sets and systems*, vol. 1, no. 3, pp. 155–167, 1978.

Estructuras de clausura difusas como puntos fijos de conexiones de Galois

Manuel Ojeda Hernández
Matemática aplicada
Universidad de Málaga
Málaga, España
manuojeda@uma.es

Inma P. Cabrera
Matemática aplicada
Universidad de Málaga
Málaga, España
ipcabrera@uma.es

Pablo Cordero
Matemática aplicada
Universidad de Málaga
Málaga, España
pcordero@uma.es

Emilio Muñoz Velasco
Matemática aplicada
Universidad de Málaga
Málaga, España
ejmunoz@uma.es

***Resumen*—Las conexiones de Galois parecen estar omnipresentes en las matemáticas. Se han utilizado para modelizar soluciones de problemas tanto puros como orientados a aplicaciones. A lo largo del artículo, el marco general es un retículo completo difuso sobre un retículo residuado completo. En este trabajo, se estudia la existencia de conexiones difusas de Galois (antítonas e isótonas) entre cuatro conjuntos ordenados específicos. Lo más interesante es que los sistemas de cierre difusos, los operadores de cierre difusos y las relaciones de cierre difusas fuertes son conceptos formales (puntos fijos) de estas conexiones de Galois difusas.**

***Index Terms*—Sistema de clausura, Retículo difuso, Conexión de Galois**

I. Estructuras de cierre difusas como conceptos formales

Esta contribución resume los artículos [3], [5], en los que se detalla la construcción de conexiones de Galois difusas de las cuales los puntos fijos son estructuras de clausura difusas.

A partir de las cuatro estructuras de cierre difusas principales, es decir, los operadores de clausura, sistemas de clausura, sistemas de clausura difusos y las relaciones de clausura difusas fuertes, que fueron introducidas en [2], [4], consideramos ahora las propiedades algebraicas de las aplicaciones que las relacionan. Por ejemplo, en [4] se estableció una correspondencia biyectiva entre los operadores de cierre y los sistemas de cierre difusos. Esta correspondencia puede verse como un par de aplicaciones, una definida desde el conjunto de los sistemas de cierre difusos al conjunto de los operadores de cierre, denotada por $\widehat{\mathsf{c}}$, y la otra del conjunto de los operadores de cierre al conjunto de los sistemas de cierre difusos, denotada por $\widetilde{\Psi}$. Como se ha demostrado en [4], estas aplicaciones son una la inversa de la otra. Un enfoque interesante es considerar estas aplicaciones y definirlas en dominios más generales. Por ejemplo, en lugar de considerar el conjunto de los operadores de clausura, consideramos el conjunto de todas las aplicaciones isótonas sobre A y en lugar de considerar el conjunto de los sistemas de clausura, consideramos el conjunto de todos los subconjuntos difusos de A.

El par de aplicaciones $(\widehat{\mathsf{c}}, \widetilde{\Psi})$ definido sobre estos conjuntos no sólo está bien definido sino que forma una conexión de Galois difusa [3]. Además, los pares formados por un sistema de cierre difuso y su correspondiente operador de cierre son puntos fijos de esta conexión de Galois. Existen, además, puntos fijos que no están formados por un operador de cierre ni un sistema de cierre difuso, como muestran los ejemplos en el artículo. Un razonamiento similar puede aplicarse para cada una de las correspondencias establecidas en [2], [4]. Así, obtenemos las conexiones de Galois difusas que aparecen en la figura 1. La relación entre los puntos fijos de estas conexiones de Galois y las estructuras de cierre se estudia en [3].

$$(L^A, S) \xleftarrow{(\widehat{\mathsf{c}}, \widetilde{\Psi})} (\mathrm{Isot}(A^A), \tilde{\rho}) \quad .$$

$$(L^A, S) \xleftarrow{(\kappa, \widehat{\Psi})} (\mathrm{IsotTot}(L^{A\times A}), \hat{\rho})$$

$$(\mathrm{IsotTot}(L^{A\times A}), \hat{\rho}) \underset{\prec}{\overset{(^{-1}, -^{\approx})}{\rightleftharpoons}} (\mathrm{Isot}(A^A), \tilde{\rho})$$

Figura 1. Conexiones de Galois antítonas/isótonas

Concluimos el estudio analizando la composición entre estas tres conexiones de Galois y examinamos la conmutatividad de los diagramas inducidos. El diagrama completo no es conmutativo, pero encontramos algunos resultados parciales sobre la conmutatividad de diagramas más pequeños y, como resultado final, obtenemos la caracterización de los dominios que proporcionan la conmutatividad deseada.

Nótese que todo el estudio anterior se ha hecho para álgebras difusas, el conjunto de subconjuntos difusos con la inclusión difusa, el conjunto de funciones isótonas y el conjunto de relaciones isótonas y totales; pero para ser exhaustivos, incluimos también en este problema el conjunto de los subconjuntos clásicos. Esto se realizó en un estudio independiente [5]. Dado

Este trabajo ha sido financiado parcialmente por la Agencia estatal de investigación (AEI), el Ministerio de Ciencia, Innovación y Universidades (MCIU), el Fondo Social Europeo (FEDER), la Junta de Andalucía (JA), y la Universidad Málaga (UMA) a través del contrato predoctoral FPU10/01467 (MCIU), el proyecto de investigación VALID (PID2022-140630NB-I00 financiado por MCIN/ AEI/ 10.13039/ 501100011033) y el proyecto de investigación con referencia PID2021-127870OB-I00 (MCIU/AEI/FEDER, UE).

que $(2^A, \subseteq)$ es un retículo clásico, hay dos formas principales de afrontar el problema. Por un lado, podemos considerar el 1-corte de las relaciones difusas anteriores y estudiar el problema buscando conexiones de Galois clásicas. Por otro lado, podemos considerar la *fuzzificación* de la relación de inclusión, es decir, utilizar la relación S en lugar de $\subseteq$ y estudiar el problema difuso.

En el problema clásico, tenemos que tratar con las aplicaciones representadas en la figura 2.

$$(2^A, \subseteq) \xleftarrow{(\widetilde{c}, \mathcal{F})} (\mathrm{Isot}(A^A), \tilde{\rho}^1)$$

$$(\mathrm{Ext}(L^A), \subseteq) \xleftarrow{(\widehat{c}, \widetilde{\Psi})} (\mathrm{Isot}(A^A), \tilde{\rho}^1)$$

$$(2^A, \subseteq) \underset{\leftarrow}{\overset{(-^1, -^{\approx})}{\rightleftarrows}} (\mathrm{Ext}(L^A), \subseteq)$$

$$(\mathrm{Isot}(A^A), \tilde{\rho}^1) \underset{\leftarrow}{\overset{(-^1, -^{\approx})}{\rightleftarrows}} (\mathrm{IsotTot}(L^{A\times A}), \hat{\rho}^1)$$

$$(\mathrm{Ext}(L^A), \subseteq) \xleftarrow{(\kappa, \widehat{\Psi})} (\mathrm{IsotTot}(L^{A\times A}), \hat{\rho}^1),$$

Figura 2. Conexiones de Galois clásicas

Parte de este estudio es un caso particular de las conexiones de Galois difusas mencionadas anteriormente. Los resultados de [3] se mantienen ya que el 1-corte de una conexión de Galois difusa es una conexión de Galois en el sentido clásico.

El par estudiado es, en efecto, una conexión de Galois clásica. Como era de esperar, los operadores y sistemas de cierre son de nuevo componentes de puntos fijos de la conexión de Galois. En este caso, a diferencia de los anteriores, todos los puntos fijos de esta conexión están formados por un operador de cierre y su sistema de cierre asociado.

Para que el par $(-^1, -^{\approx})$ sea una conexión de Galois clásica, necesitamos restringir, como en casos anteriores, los subconjuntos difusos a los subconjuntos extensionales. Esta restricción mantiene las conexiones de Galois descritas anteriormente, ya que las imágenes de las funciones implicadas son siempre conjuntos extensionales. Además, los subconjuntos extensionales con la inclusión forman un subretículo difuso de todos los subconjuntos difusos con la inclusión, por lo que la mayor parte del análisis puede adaptarse a este nuevo conjunto.

De nuevo, los sistemas de cierre y los sistemas de cierre difusos son componentes de puntos fijos de la adjunción, o conexión de Galois isótona. Por otro lado, no todos los puntos fijos de la adjunción están formados por un sistema de cierre o un sistema de cierre difuso. Esto se puede ver en el ejemplo 1.

Ejemplo 1. Sea $\mathbb{L}$ el intervalo unidad con la t-norma y el residuo de Łukasiewicz, $U = \{u\}$, consideremos el conjunto potencia (L^U, S) y el conjunto $X = \{\{u/_{0,5}\}\} \subseteq L^U$. Entonces,

$$\begin{aligned}(X^{\approx})(\{u/_\alpha\}) &= (\{u/_\alpha\} \approx \{u/_{0,5}\}) \\ &= \min\{1, 1-\alpha+0{,}5\} \otimes \min\{1, 1-0{,}5+\alpha\} \\ &= \min\{1, 1{,}5-\alpha\} \otimes \min\{1, 0{,}5+\alpha\} \\ &= \begin{cases} 1{,}5-\alpha, & \text{if } \alpha \geq 0{,}5 \\ 0{,}5+\alpha, & \text{if } \alpha \leq 0{,}5 \end{cases}\end{aligned}$$

Así, $(X^{\approx})(\{u/_\alpha\}) = 1$ si y sólo si $\alpha = 0{,}5$, eso es, $(X^{\approx})^1 = X$, el par $(X, X^{\approx})$ es un punto fijo de la adjunción pero X no es un sistema de clausura ya que $U \notin X$. De manera similar, por $X^{\approx}(U) = 0{,}5 \neq 1$ tenemos que $X^{\approx}$ no es un sistema de clausura difuso.

Se estudia ahora el problema difuso, es decir, consideramos la extensión difusa de la relación de inclusión e intentamos establecer una conexión de Galois difusa entre el conjunto de los subconjuntos clásicos con la inclusión difusa y las aplicaciones isótonas con el orden puntual. Sorprendentemente, esta conexión de Galois difusa se tiene si y sólo si (A, ρ) es un retículo clásico, es decir, ρ sólo toma los valores 0 y 1, lo que concluye esta vía de investigación.

De forma similar al artículo anterior, estudiamos si existe conmutatividad en el diagrama formado por las conexiones de Galois clásicas. Como estaba previsto, no la hay, pero obtenemos algunos resultados parciales y proporcionamos algunos subdiagramas conmutativos. Además, algunos de los resultados pueden ser mejorados si el retículo residuado subyacente es un álgebra de Heyting.

II. Conclusiones y trabajo futuro

En esta contribución se recopilan los resultados más relevantes de [3], [5]. En los artículos originales, se estudian los operadores, sistemas y relaciones de clausura desde el punto de vista de ser puntos fijos de ciertas conexiones de Galois. Los resultados muestran que las estructuras de cierre son puntos fijos de las conexiones de Galois propuestas, aunque el recíproco no siempre se cumple, es decir, hay puntos fijos que no son estructuras de cierre.

Como trabajo futuro, nos planteamos estudiar la naturaleza de estos puntos fijos que no son estructuras de cierre, ya que su caracterización podría dar lugar a unas estructuras de preclausura al estilo de Čech [1].

Referencias

[1] E. Čech. *Topological spaces*. Academia, 1966.
[2] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco. Fuzzy closure relations. *Fuzzy Sets and Systems*, 450:118–132, 2022.
[3] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco. Fuzzy closure structures as formal concepts. *Fuzzy Sets and Systems*, 463:108458, 2022.
[4] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco. Fuzzy closure systems: Motivation, definition and properties. *International Journal of Approximate Reasoning*, 148:151–161, 2022.
[5] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco. Fuzzy closure structures as formal concepts II. *Fuzzy Sets and Systems*, 473:108734, 2023.

Aggregation of T-subgroups on products

Francisco Javier Talavera
Dep. Fis. y Matemat. Apl. y DATAI
Universidad de Navarra
Pamplona, Spain 31008
Email: ftalaveraan@alumni.unav.es

Sergio Ardanza-Trevijano
Dep. Fis. y Matemat. Apl. y DATAI
Universidad de Navarra
Pamplona, Spain 31008
Email: sardanza@unav.es

Jean Bragard
Dep. Fis. y Matemat. Apl. y DATAI
Universidad de Navarra
Pamplona, Spain 31008
Email: jbragard@unav.es

Jorge Elorza
Dep. Fis. y Matemat. Apl. y DATAI
Universidad de Navarra
Pamplona, Spain 31008
Email: jelorza@unav.es

***Abstract*—We analyze the conditions under which an aggregation of a set of T-subgroups defined over the same group is again a T-subgroup. The necessary and sufficient conditions for this depend on the nature of the group. We also introduce some relaxed forms of domination to address this problem.**

I. Introduction

Aggregation operators are applied in many research areas when the objective is to combine information. For this reason, the aggregation of fuzzy structures has been an intense area of research in recent years. In particular, there exist different works regarding the aggregation of T-subgroups (see for instance [2], [3], [4]). The existing relationship between T-subgroups and T-indistinguishability operators (see [5]) could provide several fields of application of the aggregation of T-subgroups such as image processing or approximate reasoning. There are two ways to define such aggregation, namely on sets and on products (see [6]). The aim of this work is to obtain a characterization of the aggregation operators that preserve T-subgroups on products of a given group depending on the structure of such a group. The following results and discussion summarize the article [8].

II. Preliminaries

In this section we present some definitions that will be important in the following discussion. The definition of fuzzy T-subgroup by means of a t-norm T was first given by Anthony and Sherwood in [1].

Definition 1. *Let G be a group, $\mu : G \to [0,1]$ a fuzzy subset of G and T a t-norm. μ is called fuzzy T-subgroup of G if:*

G1. $\mu(e) = 1$ *where $e \in G$ denotes the neutral element.*
G2. $\mu(x) = \mu(x^{-1})\ \forall x \in G$.
G3. $\mu(xy) \geq T(\mu(x), \mu(y))\ \forall x, y \in G$.

Aggregation operators are well known functions that allow to fuse information. Fuzzy set theory is one of its application areas.

Definition 2. *Let $n \in \mathbb{N}$ and $A : [0,1]^n \to [0,1]$ be a function. A is called aggregation operator or aggregation function if:*

A1. *For all $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n) \in [0,1]^n$ such that $x_i \leq y_i\ \forall i \in \{1, \ldots, n\}$, we have $A(x_1, \ldots, x_n) \leq A(y_1, \ldots, y_n)$.*
A2. *$A(0, \ldots, 0) = 0$ and $A(1, \ldots, 1) = 1$.*

We can follow [2] to define the aggregation of fuzzy T-subgroups on products.

Definition 3. *Let $A : [0,1]^n \to [0,1]$, be an aggregation operator, T a t-norm and $n \in \mathbb{N}$. Given n fuzzy T-subgroups $\mu_1, \ldots, \mu_n$ of a group G, we denote by $\tilde{\boldsymbol{\mu}}$ the map $\tilde{\boldsymbol{\mu}} : \prod_{i=1}^n G \to [0,1]^n$ with:*

$$\tilde{\boldsymbol{\mu}}(\boldsymbol{x}) = (\mu_1(x_1), \ldots, \mu_n(x_n))$$

*for $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\prod_{i=1}^n G$. We define the **aggregation of fuzzy subgroups on products** as $A \circ \tilde{\boldsymbol{\mu}}$, where:*

$$A \circ \tilde{\boldsymbol{\mu}}(\boldsymbol{x}) = A(\mu_1(x_1), \ldots, \mu_n(x_n)).$$

*We will say that **A preserves the structure of T-subgroup on products** if and only if $A \circ \tilde{\boldsymbol{\mu}}$ is a T-subgroup for any $\tilde{\boldsymbol{\mu}}$ as above.*

III. Different forms of domination

The dominance relation introduced in [7] is a a key concept regarding the preservation of different fuzzy properties as T-indistinguishability operators or T-subgroups. We are interested in the domination over a t-norm.

Definition 4. *An aggregation operator $A : [0,1]^n \to [0,1]$ dominates a t-norm T if:*

$$A(T(x_1, y_1), \ldots, T(x_n, y_n)) \geq \\ \geq T(A(x_1, \ldots, x_n), A(y_1, \ldots, y_n))$$

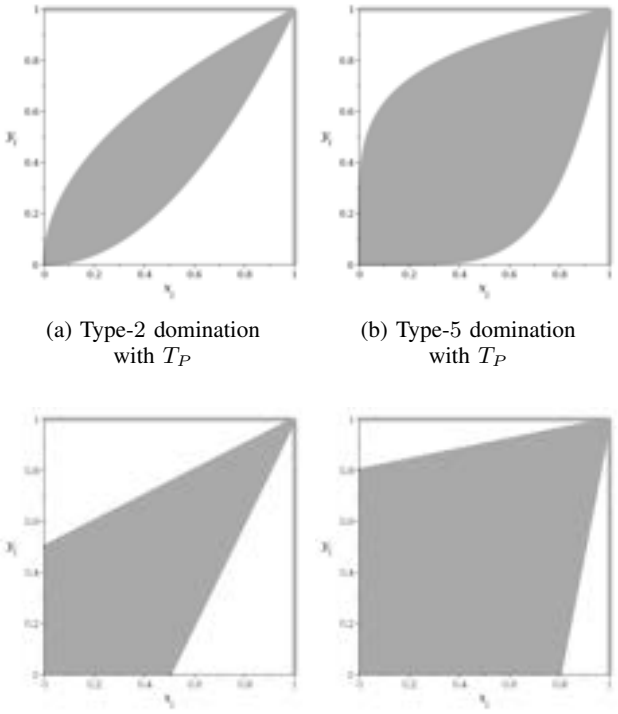
for all $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in [0,1]^n$. We denote this fact by $A \gg T$.

We now define a new concept related to domination which is key in this work. We need first to establish some notation about t-norms.

Definition 5. *Let T be a t-norm, $x \in [0,1]$ and $k \in \mathbb{N}$. We define $x_T^{(k)}$ as:*

$$x_T^{(k)} = \begin{cases} x & \text{if } k = 1, \\ T(x_T^{(k-1)}, x) & \text{if } k \neq 1. \end{cases}$$

Remark 6. *Note that, in the sense of Definition 5, the sequence $\{x_T^{(k)}\}_{k \in \mathbb{N}}$ is non-increasing due to the monotonicity of the t-norm T.*



(a) Type-2 domination with T_P

(b) Type-5 domination with T_P

(c) Type-2 domination with T_L

(d) Type-5 domination with T_L

Fig. 1: To ensure that $\boldsymbol{A} \gg_k T$, (1) must hold as long as the point (x_i, y_i) lies in the gray area of the corresponding plot for all $i \in \{1, \ldots, n\}$.

The next definition provides an interesting property that appears repeatedly in the results about preservation of T-subgroups.

Definition 7 (Type-k domination)**.** *Given $k \in \mathbb{N}$, we say that an aggregation operator A, type-k dominates a t-norm T if:*

$$A(T(x_1, y_1), \ldots, T(x_n, y_n)) \geq \geq T(A(x_1, \ldots, x_n), A(y_1, \ldots, y_n)) \quad (1)$$

for all $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in [0,1]^n$ such that for each $i \in \{1, \ldots, n\}$ one of the following conditions applies:

- $\max\{x_i, y_i\} = 1$.
- $\min\{x_i, y_i\} \geq (\max\{x_i, y_i\})_T^{(k)}$.

We denote this fact by $A \gg_k T$.

For $k = 0$ we say that A type-0 dominates T if (1) holds for all $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in [0,1]^n$ such that $\max\{x_i, y_i\} = 1$ for each $i \in \{1, \ldots, n\}$ and we denote it $A \gg_0 T$.

Remark 8. *Note that type-k domination relaxes the conditions of domination since it reduces the domain where inequality (1) is required (see Figure 1). Moreover, given $k_1, k_2 \in \mathbb{N}$ where $k_1 \geq k_2$, we have the following chain of implications:*

$$A \gg T \Rightarrow A \gg_{k_1} T \Rightarrow A \gg_{k_2} T \Rightarrow A \gg_0 T.$$

IV. Aggregation of T-subgroups on products

In this section, we study the necessary and sufficient conditions for the structure of T-subgroup to be preserved under a specific aggregation operator. The next result shows that the dominance property and type-k domination play a fundamental role depending on the ambient group. We have compiled the results included in [8] into a single theorem.

Theorem 9. *Given a group G, an aggregation operator A and a t-norm T. The following statements hold:*

1) *If $G \cong \mathbb{Z}_2$, then A preserves T-subgroups of G on products if and only if $A \gg_0 T$.*
2) *If $G \cong \mathbb{Z}_4$, then A preserves T-subgroups of G on products if and only if $A \gg_1 T$.*
3) *If $G \cong \mathbb{Z}_p$ being $p > 2$ a prime number, then A preserves T-subgroups of G on products if and only if $A \gg_{\frac{p-3}{2}} T$.*
4) *For the rest of the groups, A preserves T-subgroups of G on products if and only if $A \gg T$.*

V. Conclusions and future work

We study the cases where the aggregation of T-subgroups on products is again a T-subgroup. We have shown that any aggregation operator preserving T-subgroups of a group with non prime order and non-isomorphic to $\mathbb{Z}_4$ on products must dominate the t-norm T. For the rest of the cases, it is necessary and sufficient for that operator to type-k dominate T with the adequate election of k. Additionally, there is another way to define the aggregation of fuzzy structures, known as *on sets*. However, this concept will be studied in a future work that is about to be finished.

Acknowledgment

The authors acknowledge the finacial support of the Spanish Ministerio de Ciencia e Innovación/AEI/EU FEDER Funds (Grant PID2021-122905NB-C22). The first author also thanks the support of the "Asociación de Amigos" of the University of Navarra.

References

[1] Anthony, J. M., and Sherwood, H. (1979). Fuzzy groups redefined. *Journal of mathematical analysis and applications*, 69(1), 124-130.
[2] Ardanza-Trevijano, S., Chasco, M. J., Elorza, J., de Natividade, M. and Talavera, F. J. (2023). Aggregation of T-subgroups. *Fuzzy Sets and Systems*, 463, 108390.
[3] Bejines, C., Chasco, M. J., and Elorza, J. (2021). Aggregation of fuzzy subgroups. *Fuzzy Sets and Systems*, 418, 170-184.
[4] Bejines, C., Chasco, M. J., Elorza, J., and Recasens, J. (2019). Preserving fuzzy subgroups and indistinguishability operators. *Fuzzy Sets and Systems*, 373, 164-179.
[5] Boixader, D., and Recasens, J. (2019). On the relationship between fuzzy subgroups and indistinguishability operators. *Fuzzy Sets and Systems*, 373, 149-163.
[6] Pedraza, T., Rodríguez-López, J., and Valero, Ó. (2021). Aggregation of fuzzy quasi-metrics. *Information Sciences*, 581, 362-389.
[7] Saminger, S., Mesiar, R., and Bodenhofer, U. (2002). Domination of aggregation operators and preservation of transitivity. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(supp01), 11-35.
[8] Talavera, F. J., Ardanza-Trevijano, S., Bragard, J. and Elorza, J. (2023). Aggregation of T-subgroups of groups whose subgroup lattice is a chain. *Fuzzy Sets and Systems*, 473, 108717.

A Complementary Study on General Interval Type-2 Fuzzy Sets

Pablo Hernández
Dpto. Ciencias Exactas, Universidad San Sebastián,
Santiago, Chile,
pablo.hernandez@uss.cl

Susana Cubillo
Dpto. Matemática Aplicada a las TIC, Universidad Politécnica de Madrid,
Madrid, España
scubillo@fi.upm.es

Carmen Torres-Blanc
Dpto. Matemática Aplicada a las TIC, Universidad Politécnica de Madrid,
Madrid, España
ctorres@fi.upm.es

***Abstract*—The current research work is a comprehensive summary of our previous article published in 2022 [2]. To respect, this work studies the structure of the set of functions from [0,1] to $\{0,1\}$ (expanding the set considered by Bustince et al. in 2015), from which we have removed the constant function 0, obtaining a different study to the one carried out by Walker and Walker. More specifically, we consider join and meet operations, partial order derived from each one, and the negation operators in that set. Among other results, we provide new characterisations of join and meet operations, and of partial orders on the set of functions from [0,1] to $\{0,1\}$; we also present the first negation operators on this set.**

***Index Terms*—Fuzzy truth values, Interval type-2 fuzzy sets, join and meet operators, negations, partial orders, subalgebras.**

I. Introduction

Type-2 fuzzy sets (T2FSs) were introduced by L.A. Zadeh in 1975 [7] as an extension of type-1 fuzzy sets (FSs). Whereas for FSs the degree of membership of an element is determined by a value in the interval $[0,1]$, the degree of membership of an element for T2FSs is a fuzzy set in [0,1]. Then, a T2FS, A, is determinated by a *membership function* $\mu_A\colon X \to \mathbf{M}$, where $\mathbf{M}$ = Map([0,1],[0,1]), and $\mu_A(x)$ the degree in which x belongs to the set, is merely a label of the linguistic variable "TRUTH". Moreover, the membership degrees in interval type-2 fuzzy sets (IT2FSs) only take their values in $\{0,1\}$. There are many open problems regarding a theoretical structure of IT2FSs. To respect, we continue with the study started in papers [1], [4], [5], delving further into the properties and operations of IT2FSs. Thus, in this work we study the partial orders in this set, with the maximum and minimum elements of each one; we also analyse the possible lattice structure and negations. All this will enable us to tackle contradiction, similarity, entropy, etc., in future research. To do so, we must first consider that the support of each membership degree could be a union of intervals, some of them open or half-open, a case that has not yet been considered in previous papers. We see how this fact affects the characterisations and results presented in [4].

This work has been partially supported by Universidad San Sebastián (Chile), the Government of Spain (grant PID2021-122905NB-C22) and Universidad Politécnica de Madrid (Spain).

II. Preliminaries

For more details, in [2], one can consult the definitions, operations and properties summarized in this Section II.

Definition 1: A T2FS, A, is said to be an IT2FS if $(\mu_A(x))(u) = f_x(u) \in \{0,1\}$ for all x and u, and $f_x \neq \mathbf{0}$ ($\mathbf{0}$ is the constant function $\mathbf{0}(y) = 0$, $\forall y \in [0,1]$), that is, $f_x \in \mathbf{K} = Map\,([0,1],\{0,1\}) \setminus \{\mathbf{0}\}$.

Why do we demand $f_x \neq \mathbf{0}$? As any other knowledge-based system, fuzzy systems may be affected by the lack of information. In particular, in the case of type-2 fuzzy systems, any membership degree constant $\mathbf{a}$ (with $a \in [0,1]$) represents a lack of information. Moreover, when we concentrate on $Map\,([0,1],\{0,1\})$, the constant function $\mathbf{a}$ is limited to two cases, $\mathbf{1}$ and $\mathbf{0}$. Hence, so as not to consider two different labels to represent the lack of information, we must choose only one of them, and as a label of the "TRUTH" variable must have some value in the truth degrees, that is, with a non empty support, then we eliminate the case $\mathbf{0}$.

A function $f \in \mathbf{M}$ is normal if $sup\{f(x) : x \in [0,1]\} = 1$, and it is convex if for any $x \leq y \leq z$, the inequality $f(y) \geq f(x) \wedge f(z)$ holds. $\mathbf{L}$ denotes the set of all normal and convex functions of $\mathbf{M}$. We consider the subset of $\mathbf{K}$, denoted by $\mathbf{K}_c^F$, constituted by the functions whose support is the finite union of closed intervals.

Definition 2: The operations $\sqcup$, $\sqcap$, $\neg$, the elements $\bar{0}$, $\bar{1}$, and the auxiliaries functions f^L, f^R, are defined on $\mathbf{M}$ as follows:

$$(f \sqcup g)(x) = sup\{f(y) \wedge g(z) : y \vee z = x\},$$

$$(f \sqcap g)(x) = sup\{f(y) \wedge g(z) : y \wedge z = x\},$$

$$\neg f(x) = sup\{f(y) : 1 - y = x\} = f(1-x),$$

$$\bar{0}(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}, \quad \bar{1}(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{if } x \neq 1. \end{cases},$$

$$f^L(x) = sup\{f(y) : y \leq x\}, \; f^R(x) = sup\{f(y) : y \geq x\},$$

where $\vee$ and $\wedge$ are, respectively, the maximum and minimum operations on the lattice $([0,1],\leq)$.

$\mathbb{M} = (\mathbf{M}, \sqcup, \sqcap, \neg, \bar{0}, \bar{1})$ does not have a lattice structure [6], however, the operations $\sqcup$ and $\sqcap$ have the properties required to each define a partial order on $\mathbf{M}$.

Definition 3: The partial orders defined on **M** are as follows:

$$f \sqsubseteq g \quad \text{if } f \sqcap g = f; \quad f \preceq g \quad \text{if } f \sqcup g = g.$$

These two partial orders do not generally coincide. But, both partial orders are equivalent ($\sqsubseteq \equiv \preceq$) on **L**, and $\mathbb{L} = (\mathbf{L}, \sqcup, \sqcap, \neg, \bar{0}, \bar{1})$ has a bounded lattice structure.

Definition 4: ([3]) Let $(A, \leq_A, 0_{\leq_A}, 1_{\leq_A})$ be a bounded partial order set (poset). A function $N : A \to A$ is a negation if it is decreasing and satisfies $N(0_{\leq_A}) = 1_{\leq_A}$ and $N(1_{\leq_A}) = 0_{\leq_A}$. If, additionally, $N(N(x)) = x \ \forall x \in A$, it is said to be a strong negation.

In order to find negations, the authors introduced and studied the following operation in [3].

Definition 5: ([3]) Let n be a negation in $([0,1], \leq)$, such that n is a surjective function. The operation $N_n \colon \mathbf{M} \to \mathbf{M}$ is given by, $(N_n(f))(x) = sup\{f(y) : n(y) = x\}, \ \forall x \in [0,1]$.

III. Some Original Results

Below we present the most remarkable results obtained in our work [2].

A. Results About the Algebra of the IT2FSs

- $\mathbb{K} = (\mathbf{K}, \sqcup, \sqcap, \neg, \bar{0}, \bar{1})$ and $\mathbb{K}_c^F = (\mathbf{K}_c^F, \sqcup, \sqcap, \neg, \bar{0}, \bar{1})$, are subalgebras of $\mathbb{M}$, but neither of them has a lattice structure. Moreover, $\sqsubseteq \neq \preceq$ in these subalgebras.
- The posets $(\mathbf{K}, \preceq)$, $(\mathbf{K}_c^F, \preceq)$, $(\mathbf{K}, \sqsubseteq)$ and $(\mathbf{K}_c^F, \sqsubseteq)$ are bounded, and $\bar{0}$ and $\bar{1}$ are, respectively, the smallest and the largest elements.
- Let $f, g \in \mathbf{K}$. Then

$$(f \sqcup g)(x) = \begin{cases} f(x) \vee g(x) & \text{if } x \in Supp(f^L) \cap Supp(g^L) \\ 0 & \text{otherwise,} \end{cases}$$

$$(f \sqcap g)(x) = \begin{cases} f(x) \vee g(x) & \text{if } x \in Supp(f^R) \cap Supp(g^R) \\ 0 & \text{otherwise.} \end{cases}$$

- Let $f, g \in \mathbf{K}$, then
  $f \sqsubseteq g \iff I_f \leq_{I^*} I_g$ and $f(x) \geq g(x), \forall x \in I_f$.
  $f \preceq g \iff I_f \leq_{I^*} I_g$ and $g(x) \geq f(x), \forall x \in I_g$.
  where $I_f = Supp(f^L) \cap Supp(f^R) = /a, b/ \leq_{I^*} I_g = Supp(g^L) \cap Supp(g^R) = /c, d/$ if and only if $a \leq c$, $b \leq d$, $\overline{/c, d/}(a) \leq \overline{/a, b/}(a)$, and $\overline{/a, b/}(d) \leq \overline{/c, d/}(d)$.

  The notation for intervals between two slashes, for example $/a, b/$, refers to any non-empty interval in $[0,1]$,and $\overline{/a, b/}$ is its characteristic function.

B. Results About Negations on the IT2FSs

- The operation $\neg$, and generally, N_n are not negations on $(\mathbf{K}, \preceq)$, $(\mathbf{K}, \sqsubseteq)$, $(\mathbf{K}_c^F, \preceq)$, and $(\mathbf{K}_c^F, \sqsubseteq)$.
- Let $f \in \mathbf{K}$, $v_i = \inf\{Supp(f)\}, v_s = \sup\{Supp(f)\}$, n_1, n_2 negations in [0,1], such that $n_1(x) \leq n_2(x)$ for all $x \in [0,1]$. Then, the operations

$$N_{n_1,n_2}(f) = \overline{[n_1(v_s), n_2(v_i)]},$$

and if n_1, n_2 are strictly decreasing

$$N^*_{n_1,n_2}(f) = \overline{/n_1(v_s), n_2(v_i)/},$$

are negations, but not strong, on $(\mathbf{K}, \sqsubseteq)$, $(\mathbf{K}, \preceq)$, $(\mathbf{K}_c^F, \sqsubseteq)$ and $(\mathbf{K}_c^F, \preceq)$.

- The operations

$$N_\wedge(f) = \begin{cases} \bar{1}, & \text{if } f = \bar{0} \\ \bar{0}, & \text{if } f \neq \bar{0} \end{cases}, \qquad N_\vee(f) = \begin{cases} \bar{0}, & \text{if } f = \bar{1} \\ \bar{1}, & \text{if } f \neq \bar{1} \end{cases}$$

are, respectively, the minimum and the maximum negations in $(\mathbf{K}, \sqsubseteq)$, $(\mathbf{K}, \preceq)$, $(\mathbf{K}_c^F, \sqsubseteq)$, and $(\mathbf{K}_c^F, \preceq)$.

- Let n be a negation in $[0,1]$. Then the operations

$$N_r^n(f) = \begin{cases} \bar{0}, & \text{if } f(1) = 1 \\ \bar{1}, & \text{if } f = \bar{0} \\ \mathbf{1} - f^R \circ n, & \text{otherwise} \end{cases},$$

$$N_l^n(f) = \begin{cases} \bar{1}, & \text{if } f(0) = 1 \\ \bar{0}, & \text{if } f = \bar{1} \\ \mathbf{1} - f^L \circ n, & \text{otherwise} \end{cases},$$

are negations, not strong, on $(\mathbf{K}, \sqsubseteq)$ and $(\mathbf{K}, \preceq)$, But, they are not negations on $(\mathbf{K}_c^F, \sqsubseteq)$ and $(\mathbf{K}_c^F, \preceq)$.

IV. Conclusions

In our article published in the journal "IEEE Transaction on Fuzzy Systems" [2] we carried out a detailed study of the algebraic structure of $\mathbb{K}$ and $\mathbb{K}_c^F$, showing that they are subalgebras of $\mathbb{M}$ and bounded posets, obtaining the maximum and minimum elements. We also achieved characterizations of the operations $\sqcup, \sqcap$, and of the partial orders $\sqsubseteq, \preceq$ on **K** and $\mathbf{K}_c^F$. These characterisations often make some processes easier.

Furthermore, for the first time, negations in the sets **K** and $\mathbf{K}_c^F$ in relation to each partial order, have been presented.

This study opens the door to future research on other topics, such as strong negations, continuous negations, t-norms, t-conorms, aggregation, contradiction, similarity, etc., on the sets **K** and $\mathbf{K}_c^F$.

References

[1] H. Bustince, J. Fernandez, H. Hagras, F. Herrera, M. Pagola and E. Barrenechea, "Interval Type-2 Fuzzy Sets are Generalization of Interval-Valued Fuzzy Sets: Toward a Wider View on Their Relationship", *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 5, pp. 1876–1882, 2015.

[2] P. Hernández, S. Cubillo and C. Torres-Blanc, "A Complementary Study on General Interval Type-2 Fuzzy Sets, " *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 11, pp. 5034–5043, november 2022.

[3] P. Hernández, S. Cubillo and C. Torres-Blanc, "Negations on type-2 fuzzy sets", *Fuzzy Sets Syst.*, vol. 252, pp. 111–124, 2014.

[4] G. Ruiz, H. Hagras, H. Pomares, I. Rojas and H. Bustince, "Join and meet operations for type-2 fuzzy sets with nonconvex secondary memberships", *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 4, pp. 1000–1008, 2016.

[5] G. Ruiz-García, H. Hagras, I. Rojas and H. Pomares, "Towards a Framework for Singleton General Forms of Interval Type-2 Fuzzy Systems", *Lecture Notes in Computer Science*, pp. 3–26, 2017.

[6] C. Walker and E. Walker, "The algebra of fuzzy truth values", *Fuzzy Sets Syst.*, vol. 149, pp. 309–347, 2005.

[7] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I", *Inf. Sci.*, vol. 8, pp. 199–249, 1975.

El f-índice de inclusión como par adjunto óptimo para modus ponens difuso

Nicolás Madrid
Matemática Aplicada
Universidad de Málaga
Málaga, España
nicolas.madrid@uma.es

Manuel Ojeda-Aciego
Matemática Aplicada
Universidad de Málaga
Málaga, España
aciego@uma.es

***Resumen*—Continuamos estudiando las propiedades del f-índice de inclusión y mostramos que, dado un par fijo de conjuntos difusos, su f-índice de inclusión puede vincularse a una conjunción difusa que forma parte de un par adjunto. También mostramos que, cuando este par se utiliza como estructura subyacente para proporcionar una interpretación difusa de la regla de inferencia *modus ponens*, proporciona el máximo valor de verdad posible en la conclusión entre todos los valores obtenidos por *modus ponens* difuso utilizando cualquier otro par adjunto.**

Keywork: N. Madrid and M. Ojeda-Aciego. The f-index of inclusion as optimal adjoint pair for fuzzy modus ponens. *Fuzzy Sets and Systems*, 466:Article 108474, 2023

I. Introducción

El origen del f-índice de inclusión se remonta a la incorporación de la negación en programas lógicos multiadjuntos [8] que, como es sabido, puede hacer surgir inconsistencias. Al estudiar los programas lógicos normales residuados bajo la semántica de conjunto-respuesta [2], quedó claro que la (in)consistencia no debe considerarse como una noción nítida cuando se aplica en teorías de lógica difusa (general). Como resultado, introdujimos la noción de *contradicción débil* en [1] en el marco general de la teoría de conjuntos difusos.

Animados a buscar algunos aspectos motivacionales a la noción de *contradicción débil*, nos dimos cuenta de que estaba estrechamente relacionada con un tipo de inclusión. Por lo tanto, poco después de introducir medidas para la contradicción débil, empezamos a imaginar algún tipo de enfoque *basado en funciones* para medir la inclusión entre conjuntos difusos, y presentamos las primeras ideas sobre el f-índice de inclusión en [6], donde introdujimos la noción de inclusión mediante la asignación de una *función* a cada par de conjuntos difusos. Posteriormente, en [5], analizamos el f-índice de inclusión en términos de las definiciones axiomáticas más comunes de medida de inclusión, y probamos que es compatible con los conocidos axiomas propuestos en [9]. En este trabajo analizamos la relación del f-índice de inclusión con las medidas difusas estándar de inclusión basadas en la implicación residuada.

Parcialmente financiado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación y los fondos FEDER a través del proyecto VALID, PID2022-140630NB-I00/AEI/10.13039/501100011033/ FEDER, UE.

II. f-índice de inclusión

El f-índice de inclusión se define en tres pasos. En el primero, definimos cuales son los "grados" con los que representaremos o mediremos la inclusión entre conjuntos difusos.

Definición 1: El *conjunto de f-índices*, denotado por Ω, es el conjunto formado por las funciones crecientes $f\colon [0,1] \to [0,1]$ que cumplen $f(x) \leq x$ para todo $x \in [0,1]$.

En el segundo de los pasos, definimos cuándo un par de conjuntos difusos satisface un "grado" de f-inclusión.

Definición 2: Sean A y B dos conjuntos difusos y sea $f \in \Omega$. Decimos que A *está f-incluido en* B (denotado por $A \subseteq_f B$) si y solo si la desigualdad $f(A(u)) \leq B(u)$ se cumple para todo $u \in \mathcal{U}$.

Finalmente, hay que tener en cuenta que no representaremos la inclusión con una única f-inclusión, sino con todas ellas. En concreto la idea es representar la inclusión siguiendo la idea de que "cuantas más f-inclusiones se satisfagan entre dos conjuntos difusos, mayor será la inclusión". Puede demostrarse [5] que el conjunto de las f-inclusiones que satisfacen dos conjuntos difusos tiene estructura de sub-retículo y, de hecho, está caracterizado por el máximo de todas ellas. Por ese motivo, el tercero de los pasos consiste en definir el f-índice de inclusión de la siguiente forma:

Definición 3 (f-índice de inclusión): Sean A y B dos conjuntos difusos, el *f-índice de inclusión* de A en B, denotado por $\mathrm{Inc}(A,B)$, se define como

$$\mathrm{Inc}(A,B) = \max\{f \in \Omega \mid A \subseteq_f B\}$$

Es fácil comprobar que la inclusión difusa estándar

$$S(A,B) = \bigwedge_{u \in \mathcal{U}} A(u) \to B(u) \tag{1}$$

donde $\to$ es una implicación residuada,[1] es equivalente a un grado particular de f-inclusión. En concreto, dado un par adjunto $(*, \to)$, tenemos para todo $\alpha \in [0,1]$ la equivalencia

$$\alpha \leq A(u) \to B(u) \iff A(u) * \alpha \leq B(u).$$

[1]En lo que sigue asumimos que el lector conoce la terminología básica de retículos residuados y, en particular, la propiedad de adjunción.

Por tanto, si consideramos $f_\alpha(x) = x * \alpha$, que obviamente pertenece a Ω, tenemos:

$$\alpha \leq \bigwedge_{u\in\mathcal{U}} A(u) \to B(u) \iff \alpha \leq A(u) \to B(u)a \quad \forall u \in \mathcal{U}$$
$$\iff A \text{ está } f_\alpha\text{-incluido en } B$$

Como resultado, la desigualdad $\alpha \leq S(A, B)$ equivale a decir que A está f_α-incluido en B. Por lo tanto, el procedimiento estándar para asignar grados de verdad a una fórmula del tipo $\forall u\big(A(u) \to B(u)\big)$ mediante pares adjuntos, coincide con un caso específico de f-inclusión.

Podemos definir una variación del f-índice de inclusión restringida a un determinado conjunto de índices de inclusión de la siguiente manera:

Definición 4: Sean A y B dos conjuntos difusos y Θ un sub-retículo de Ω; es decir, Θ es cerrado bajo supremos arbitrarios y contiene a $\bot$ y id. Entonces, el *f-índice de inclusión restringido a* Θ, denotado por $\text{Inc}_\Theta(A, B)$, se define como

$$\text{Inc}_\Theta(A, B) = \sup\{f \in \Theta \mid A \subseteq_f B\}.$$

El siguiente resultado muestra que $\text{Inc}_\Theta(A, B)$ generaliza efectivamente la medida estándar de inclusión dada por la Ecuación (1) cuando se basa en pares adjuntos.

Teorema 1: Sea $(*, \to)$ un par adjunto, A y B conjuntos difusos y $\Theta = \{x * \alpha \mid \alpha \in [0, 1]\}$, entonces

$$\text{Inc}_\Theta(A, B)(x) = x * \left(\bigwedge_{u\in\mathcal{U}} A(u) \to B(u)\right).$$

A continuación, mostramos una especie de afirmación inversa, a saber, que cuando el universo es finito, el f-índice original de inclusión siempre puede estar vinculado a un par adjunto. Para demostrar tal relación, necesitamos introducir en primer lugar el siguiente resultado.

Proposición 1: Sean A y B dos conjuntos difusos sobre un universo finito $\mathcal{U}$. Entonces, existe una función $\overline{\text{Inc}}(A, B)\colon [0, 1] \to [0, 1]$ tal que $(\text{Inc}(A, B), \overline{\text{Inc}}(A, B))$ forma un par adjunto.

Ahora ya enunciamos el resultado recíproco, que tiene importantes repercusiones, ya que muestra que en el caso de un universo finito, el f-índice de inclusión puede asociarse al menos a un par adjunto formado por una conjunción difusa y una implicación difusa.

Teorema 2: Sean A y B dos conjuntos difusos sobre un universo finito $\mathcal{U}$. Entonces, existe un par adjunto $(*, \to)$ con $*$ una conjunción difusa conmutativa y $\alpha \in [0, 1]$ tal que

$$\text{Inc}(A, B)(x) = x * \alpha.$$

III. El f-índice de inclusión como operador óptimo para modus ponens difuso

Si identificamos los valores de $\bigwedge_{u\in\mathcal{U}} A(u) \to B(u)$ y $A(u_0)$ con α y β (en $[0, 1]$), una opción para representar el *modus ponens* difuso es la siguiente:

$$\begin{array}{ll} A \Rightarrow B & \equiv \alpha \\ A(u) & \equiv \beta \\ \hline \therefore\ B(u) & \geq \beta * \alpha \end{array} \tag{2}$$

y la corrección de tal inferencia viene dada por la propiedad de adjunción del par $(*, \to)$.

El siguiente teorema nos indica la maximalidad de la respuesta obtenida al aplicar modus ponens difuso usando el f-índice de inclusión, con respecto del valor obtenido mediante cualquier otro par adjunto.

Teorema 3: Sean A y B dos conjuntos difusos definidos sobre un universo finito $\mathcal{U}$ y sea $u_0 \in \mathcal{U}$, entonces:

$$B(u_0) \geq \text{Inc}(A, B)(A(u_0)) \geq A(u_0) * \left(\bigwedge_{u\in\mathcal{U}} A(u) \to B(u)\right)$$

para todo par de pares residuados $(*, \to)$ y $x \in [0, 1]$.

IV. Conclusiones y trabajo futuro

Se ha presentado una versión muy preliminar de un posible sistema de inferencia difusa basado en el f-índice de inclusión y su uso para recuperar datos que faltan. Para completar dicho sistema, es necesario comprobar diferentes posibilidades en la técnica de construcción y probar los resultados de los datos que no se han utilizado en el conjunto de datos de entrenamiento. Desde un punto de vista teórico, estudiaremos el problema de construir una t-norma $*$ compatible con el Teorema 2.

Que el f-índice de inclusión se comporte de forma similar a una implicación abre la posibilidad de incluirlo en una estructura residuada. En tal caso, podríamos utilizar el f-índice de inclusión directamente como una implicación y aplicarlo en todas aquellas áreas en las que la estructura subyacente sea residuada, por ejemplo, el Análisis de Conceptos Formales Difuso [7] o la Morfología Matemática Difusa [3], [4].

Referencias

[1] H. Bustince, N. Madrid, and M. Ojeda-Aciego. *The Notion of Weak-Contradiction: Definition and Measures*. IEEE Trans. Fuzzy Systems, **23**(4) (2015), 1057–1069.

[2] C.V. Damásio, N. Madrid and M. Ojeda-Aciego *On the Notions of Residuated-Based Coherence and Bilattice-Based Consistence.* Lecture Notes in Computer Science **6857** (2011) 115–122.

[3] B. De Baets, E. Kerre and M. Gupta. *The fundamentals of fuzzy mathematical morphology, part 1 : basic concepts.* International Journal of General Systems. 23(2) (1995). 155–171

[4] N. Madrid, J. Medina, M. Ojeda-Aciego, and I. Perfilieva. *L-fuzzy relational mathematical morphology based on adjoint triples*. Information Sciences, **474** (2019) 75–89.

[5] N. Madrid and M. Ojeda-Aciego. *A view of f-indexes of inclusion under different axiomatic definitions of fuzzy inclusion*. Lecture Notes in Artificial Intelligence, **10564** (2017), 307–318.

[6] N. Madrid, M. Ojeda-Aciego, and I. Perfilieva. *f-inclusion indexes between fuzzy sets*. Proc. of IFSA-EUSFLAT 2015.

[7] 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.

[8] J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Similarity-based unification: a multi-adjoint approach. Fuzzy Sets and Systems, 146(1):43-62, 2004.

[9] D. Sinha and E. R. Dougherty. *Fuzzification of set inclusion: Theory and applications*. Fuzzy Sets and Systems, **55**(1) (1993), 15–42.

Distribución No Uniforme de la Granularidad de la Información para Mejorar la Consistencia y el Consenso en Toma de Decisiones en Grupo

Juan Carlos González-Quesada
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
juancarlosgq@ugr.es

Ignacio Javier Pérez
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
ijperez@decsai.ugr.es

Sergio Alonso
dept. Lenguajes y Sistemas Informáticos
University of Granada
Granada, España
zerjioi@ugr.es

Juan Miguel Tapia
dept. Métodos Cuantit. Econo. y Empresa
Universidad de Granada
Granada, España
jmtaga@ugr.es

Enrique Herrera-Viedma
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
viedma@decsai.ugr.es

Francisco Javier Cabrerizo
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
cabrerizo@decsai.ugr.es

***Abstract*—Este trabajo es un resumen del artículo *Eng. Appl. Artif. Intell.*, vol. 130, art. nro. 107737, 2024. En ese artículo, se presenta un nuevo enfoque basado en computación granular que está compuesto por dos procesos para resolver problemas de toma de decisiones en grupo con múltiples criterios. Así, en primer lugar, se desarrolla un proceso automático basado en una asignación no uniforme de la granularidad de la información para maximizar tanto el consenso como la consistencia de las relaciones de preferencia difusas que modelan las opiniones de los participantes en el problema de decisión. Sobre esta base, en segundo lugar, se introduce un proceso interactivo que requiere la implicación de los integrantes del grupo y que también pretende maximizar tanto el consenso como la consistencia. Para mostrar la esencia de este nuevo enfoque, así como su rendimiento y flexibilidad, se realizan diversos estudios experimentales. Finalmente, con el objetivo de validar su eficacia y viabilidad en la resolución de problemas prácticos del mundo real, se aplica para resolver un problema de decisión sobre rehabilitación de edificios.**

***Index Terms*—Consenso, consistencia, relación de preferencia difusa, granularidad de la información, toma de decisiones en grupo**

I. Propuesta

Un problema de toma de decisiones en grupo con múltiples criterios se define como una situación en la que una serie de personas, $P = \{p_1, \ldots, p_n\}$, tienen que seleccionar de forma colectiva la mejor alternativa de entre un conjunto de estas, $A = \{a_1, \ldots, a_m\}$, teniendo en cuenta para ello una serie de criterios, $C = \{c_1, \ldots, c_o\}$. Para resolver este tipo de problemas, se debe hacer frente, entre otras, a dos cuestiones: la consistencia individual de las opiniones proporcionadas por las personas participantes y el consenso alcanzado entre ellas.

Esta publicación es parte del proyecto de I+D+i PID2022-139297OB-I00, financiado por MICIU/AEI/10.13039/501100011033 y por FEDER/UE. También se ha realizado gracias a la ayuda PREP2022-000352 financiada por MICIU/AEI /10.13039/501100011033.

En vista de que es inusual que las opiniones individuales sean totalmente consistentes y, sobre todo, llegar a un consenso al inicio del proceso de decisión, se han propuesto diferentes métodos que pretenden mejorar la consistencia, el consenso, e incluso ambos, particularmente en entornos con incertidumbre. Es decir, entornos en los que los conjuntos difusos y sus extensiones se han mostrado útiles para modelar las opiniones de los participantes en el problema de toma de decisiones [1]. Estos métodos se esfuerzan en mejorar la consistencia y el consenso. Sin embargo, la mayoría consiguen este objetivo de manera que producen desviaciones grandes entre la opinión inicial dada por el participante y las modificadas. Esto conduce a una importante pérdida de información. Afortunadamente, se pueden introducir ciertas restricciones en este proceso de modificación de las opiniones para tratar de minimizar esta pérdida de información.

La granularidad de la información [2], concepto clave dentro del paradigma de la computación granular, se ha empleado en los últimos años para controlar la diferencia entre las opiniones iniciales y las modificadas. La idea consiste en modelar las opiniones en términos de gránulos de información (representados generalmente como intervalos), los cuales proporcionan la flexibilidad necesaria para incrementar el acuerdo y la consistencia mientras se controla la pérdida de información [3]. Siguiendo este razonamiento, se han propuesto varios modelos de toma de decisiones en grupo cuya característica común es que emplean una distribución uniforme de la granularidad de la información. Aunque estos modelos permiten aumentar la consistencia y el consenso, a la vez que restringen el rango en el que pueden cambiar las opiniones, gracias a los gránulos de información, se ha demostrado que su rendimiento y flexibilidad podría mejorarse con una distribución no uniforme de la granularidad de la información [4].

Existen, sin embargo, cuestiones sin resolver en los modelos basados en una distribución no uniforme de la granularidad de la información. Por ejemplo, algunos modelos solamente mejoran la consistencia [5], o el consenso [6]. Además, se basan en un proceso automático de modificación de las opiniones, el cual no implica la participación de los miembros del grupo. Sin embargo, en problemas reales, las opiniones deberían modificarse y volver a evaluarse con la participación esencial de estos. Por su parte, otros modelos mejoran tanto el consenso como la consistencia, pero no en problemas con múltiples criterios [4], [7]. Por último, existen modelos que mejoran tanto la consistencia como el consenso en problemas de decisión en grupo con múltiples criterios [8], [9]. Sin embargo, estos modelos no implican la participación de los miembros del grupo en el proceso de modificación.

Con el fin de mejorar los modelos existentes basados en una asignación no uniforme de la granularidad de la información, la propuesta presentada en [10] presenta el primer enfoque basado en una distribución de la granularidad de la información que mejora tanto la consistencia como el consenso en problemas de toma de decisiones en grupo con múltiples criterios y que permite la participación de los miembros del grupo en este proceso de modificación. Asumiendo relaciones de preferencia difusas para modelar las opiniones de los miembros del grupo e intervalos para modelar los gránulos de información, este objetivo general condujo a las siguientes contribuciones principales:

- Desarrollo de un proceso automático de mejora de la consistencia y el consenso que asume un nivel medio de granularidad de la información para asignarla de forma no uniforme entre las entradas de las relaciones de preferencia difusas. El nivel medio de granularidad de la información admitido por la persona es sinónimo de su flexibilidad, es decir, cuanto mayor sea el valor admitido, mayor será la cooperación de esta persona para mejorar el consenso y la consistencia. Existe un desajuste entre la pérdida de información y la mejora de la consistencia y el consenso, es decir, un incremento de la consistencia y el consenso significa también un incremento de la pérdida de información. No obstante, el nivel medio de granularidad de la información, el cual está acotado, permite maximizar el consenso y la consistencia con una pérdida de información limitada entre las relaciones de preferencia difusas iniciales y las modificadas.
- Desarrollo de un proceso interactivo que implica la participación de los miembros del grupo. Este proceso se basa en el anterior y proporciona un marco de aplicación completo para la resolución de problemas de toma de decisiones en grupo con múltiples criterios en entornos del mundo real.

Para ilustrar la esencia del enfoque propuesto, así como su flexibilidad y rendimiento, se llevan a cabo diversos estudios experimentales y se realiza un análisis comparativo. En comparación con los enfoques basados en una distribución uniforme de la granularidad de la información, el enfoque propuesto en este trabajo alcanza mayores valores para el consenso y la consistencia. Esto se debe a que una distribución no uniforme de la granularidad de la información permite un uso más eficiente de los recursos al asignar mayores niveles de granularidad a las entradas de las relaciones de preferencia difusas que más lo requieren. Esto conduce a mejores resultados en la toma de decisiones, ya que se alcanzan mayores valores de consistencia y consenso.

Además, se presenta un caso de estudio sobre la reconstrucción de un edificio no residencial en un edificio residencial, el cual ayuda a entender la eficacia y la aplicabilidad de este nuevo enfoque granular en la resolución de problemas del mundo real. En concreto, el proyecto preveía la decoración exterior del edificio, la instalación de ventanas de metal y plástico, ventilación (aire acondicionado), un sistema de suministro de agua, calefacción y tejas. Sin embargo, no preveía la intervención o el traslado de ninguna estructura situada en el lugar del diseño ni la eliminación de la vegetación existente. Por tanto, había que renovar el sistema de suministro de agua, la calefacción, la ventilación, las ventanas y el tejado del edificio. Entre ellos, el artículo se centra en la renovación de la ventilación, en donde había que elegir entre cinco empresas diferentes (Danfoss, KERMI, BT-Service, VIKMA LTD, Behtc) considerando para ello cuatro indicadores (nivel de ruido, durabilidad, precio, presión máxima).

REFERENCES

[1] H. Bustince, E. Berrenechea, M. Pagola, J. Fernández, Z.S. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, B. De Baets, "A historical account of types of fuzzy sets and their relationships," IEEE Trans. Fuzzy Syst., vol. 24, pp. 179–194, Feb. 2016.

[2] W. Pedrycz, "Allocation of information granularity in optimization and decision making models: Towards building the foundations of granular computing," European J. Oper. Res., vol. 232, pp. 137-145, Jan. 2014.

[3] J. Qin, L. Martínez, W. Pedrycz, X. Ma, Y. Liang, "An overview of granular computing in decision-making: Extensions, applications, and challenges," Inf. Fusion, vol. 98, art. no. 101833, Oct. 2023.

[4] B. Zhang, W. Pedrycz, A.R. Fayek, Y.C. Dong, "A differential evolution-based consistency improvement method in AHP with an optimal allocation of information granularity," IEEE T. Cybern., vol. 52, pp. 6733–6744, Jul. 2022.

[5] F.J. Cabrerizo, J.C. González-Quesada, J.A. Morente-Molinera, I.J. Pérez, E. Herrera-Viedma, E., W. Pedrycz, "An improvement of consensus in group decision-making through an optimal distribution of information granularity," in Proc. IEEE Symp. Ser. Comput. Intell. SSCI, Singapur, pp. 119–124, 2022.

[6] F.J. Cabrerizo, A. Kaklauskas, I.J. Pérez, E. Herrera-Viedma, "A granular-based approach to address multiplicative consistency of reciprocal preference relations in decision-making," in Proc. 56th Hawaii Int. Conf. Syst. Sci., Maui, Hawaii, USA, pp. 1541–1550, 2023.

[7] B. Zhang, Y.C. Dong, W. Pedrycz, "Consensus model driven by interpretable rules in large-scale group decision making with optimal allocation of information granularity," IEEE Trans. Syst. Man Cybern. -Syst., vol. 53, pp. 1233–1245, Feb. 2023.

[8] J. Qin, X. Ma, Y. Liang, "Building a consensus for the best-worst method in group decision-making with an optimal allocation of information granularity," Inform. Sci., vol. 619, pp. 630–653, Jan. 2023.

[9] J. Qin, X. Ma, W. Pedrycz, "A granular computing-driven best-worst method for supporting group decision making," IEEE Trans. Syst. Man Cybern. Syst., vol. 53, pp. 5591–5603, Sep. 2023.

[10] J.C. González-Quesada, A. Velykorusova, A. Banaitis, A. Kaklauskas, F.J. Cabrerizo, "Non-uniform allocation of information granularity to improve consistency and consensus in multi-criteria group decision-making: Application to building refurbishment," Eng. Appl. Artif. Intell., vol. 130, art. no. 107737, Apr. 2024.

Un método de ayuda a la toma de decisiones en grupo basado en el comportamiento de los expertos durante el debate

1st José Ramón Trillo
Dpt. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
jrtrillo@ugr.es

2nd Enrique Herrera-Viedma
Dpt. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
viedma@decsai.ugr.es

3rd Juan Antonio Morente-Molinera
Dpt. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
jamoren@decsai.ugr.es

4th Francisco Javier Cabrerizo
Dpt. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
cabrerizo@decsai.ugr.es

5th Ignacio Javier Pérez
Dpt. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
ijperez@decsai.ugr.es

***Abstract*—Esta contribución trata de difundir el artículo "A group decision-making method based on the experts' behaviour during the debate" que ha sido recientemente publicado en la revista "IEEE Transactions on Systems, Man, and Cybernetics" [1]. El debate es un proceso que consiste en llegar a una decisión razonada sobre una serie de alternativas en el que los individuos deben ser capaces de defender sus propias opiniones. Se ha utilizado en problemas de toma de decisiones en grupo (TDG) para ayudar a los expertos a tomar mejores decisiones. Sin embargo, si los expertos se enzarzan en un debate enérgico, puede dar lugar al uso de un lenguaje agresivo que disminuya el consenso, que es el principal objetivo de la TDG. Para evitarlo, presentamos un método novedoso para problemas de TDG que puede identificar comentarios agresivos durante el debate incorporando un clasificador basado en técnicas de análisis de sentimientos. A diferencia de los métodos de TDG existentes, este nuevo enfoque puede aprovechar la información extraída durante el debate (es decir, el comportamiento de los expertos) a lo largo de todo el proceso de decisión, lo que lo hace más similar a los procesos de TDG del mundo real.**

***Index Terms*—Toma de decisiones en grupo, Consenso, Análisis de sentimientos**

I. Introduccion

En este trabajo hemos presentado un nuevo método de TDG basado en el comportamiento de los expertos durante la etapa de debate llevada a cabo antes de proporcionar las evaluaciones. Los métodos de TDG existentes [2] no tienen en cuenta el comportamiento de los expertos durante la etapa de debate en el resto del proceso de toma de decisiones y, al no considerarlo, se pierde información que podría ser de interés. Sin embargo, este nuevo método de TDG incorpora un clasificador basado en técnicas de análisis de sentimientos [3] que puede extraer datos adicionales generados durante la etapa de debate y aprovecharlo. Concretamente, basándonos en el comportamiento de los expertos, hemos desarrollado dos procedimientos para asignar pesos a los expertos y se han propuesto dos nuevas medidas de consenso.

Esta publicación es parte del proyecto de I+D+i PID2022-139297OB-I00, financiado por MICIU/AEI/10.13039/501100011033 y por FEDER/UE y del proyecto C-ING-165-UGR23, cofinanciado por la Consejería de Universidad, Investigación e Innovación y por la Unión Europea con cargo al Programa FEDER Andalucía 2021-2027

II. Metodología y Resultados

Aunque los modelos para TDG traten diferentes aspectos del proceso de toma de decisiones, la comparación de los resultados devueltos por un modelo con otros no es una tarea sencilla. Las características consideradas por los modelos son diferentes y, como consecuencia, una comparación cuantitativa no sería significativa. De todas formas, a continuación, analizamos algunas ventajas y limitaciones del enfoque propuesto:

- Al incorporar un clasificador, que se basa en técnicas de análisis de sentimientos y hace uso de diferentes características lingüísticas, el modelo propuesto puede categorizar el comportamiento de los expertos durante la etapa de debate. Dicho clasificador está compuesto por las siguientes etapas (ver Fig. 1): (i) detección de lenguaje agresivo durante el debate, (ii) provisión de evaluaciones, (iii) cálculo de los pesos de los expertos, (iv) análisis de consenso, (v) cálculo de una evaluación colectiva, y (vi) clasificación de las alternativas.
  En comparación con los métodos existentes podemos observar una ventaja notable porque se ha demostrado que, primero, un lenguaje agresivo afecta negativamente a la confiabilidad del hablante y a su credibilidad y,

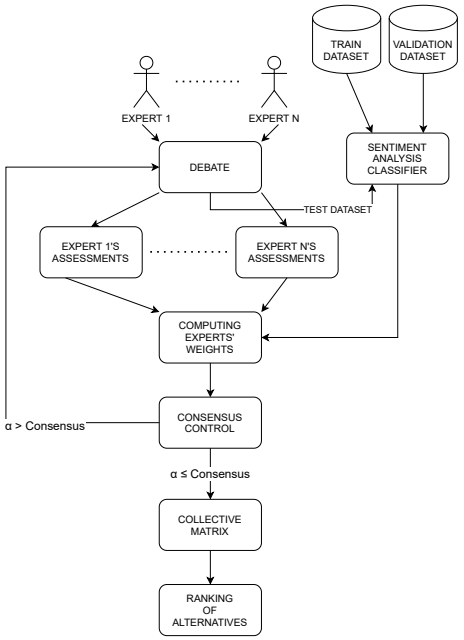


Fig. 1. Esquema general del método propuesto.

segundo, también conduce al deterioro del debate. Ambos escenarios deben evitarse porque dificultan el logro del consenso, que es el punto más importante en los procesos de TDG. Por lo tanto, el modelo propuesto en este estudio, al explotar la información devuelta por el clasificador basado en análisis de sentimientos, puede modelar mejor el proceso de alcanzar consenso llevado a cabo en un proceso de TDG.

- Además, este nuevo modelo, introduce dos nuevos procedimientos para asignar pesos a los expertos según su comportamiento durante la etapa de debate. Esto permite considerar el comportamiento de los expertos en el resto del proceso de TDG, específicamente, en el análisis del consenso (al calcular el consenso alcanzado entre las evaluaciones de los expertos) y en el cálculo de la evaluación colectiva (cuando se agregan las evaluaciones individuales de los expertos). Por una parte, el procedimiento para calcular el peso global obtiene un promedio ponderado de los comentarios agresivos proporcionados por un experto. Su ventaja es que incluso si todos los comentarios proporcionados en una ronda son agresivos, el peso del experto puede compensarse en otras rondas donde no se proporcione ningún comentario agresivo. Sin embargo, tiene la desventaja de que es más general, lo que implica que el comportamiento del experto solo se puede ver durante el debate en su conjunto y no en cada ronda. Por otro lado, el procedimiento para calcular el peso iterativo utiliza el peso de la ronda anterior para calcular el nuevo. Esto tiene la ventaja de que se puede analizar en cada ronda de una manera más detallada porque el peso relacionado con el experto se altera según la ronda anterior. Sin embargo, es más estricto porque en el caso de que todos los comentarios proporcionados por un experto sean agresivos, el peso de este experto será igual a cero, aunque, en las otras rondas, el experto no proporcione ningún comentario agresivo. Por lo tanto, mediante el uso de estos dos nuevos procedimientos, se proponen dos nuevas medidas de consenso y dos nuevos métodos para calcular la evaluación colectiva que consideran el comportamiento de los expertos durante la etapa de debate.
- El clasificador basado en técnicas de análisis de sentimientos utiliza diferentes características lingüísticas que proporcionan información relevante al clasificar los comentarios. Presenta las siguientes ventajas: Primero, detecta comentarios agresivos utilizando diferentes características lingüísticas que completan el preprocesamiento. Algunas de ellas ya se han utilizado, como el uso de POS-tagging, pero otras son nuevas, como el número de palabras completas escritas en letras mayúsculas o el número consecutivo de exclamaciones repetidas, por citar algunos ejemplos. Segundo, utiliza una matriz de preprocesamiento optimizada construida mediante diferentes métodos para eliminar palabras que no proporcionan información útil y para reducir palabras a sus raíces (reduce el número de elementos en el conjunto de palabras). Sin embargo, su inconveniente es que podría no descubrir todas las características lingüísticas que están presentes. Como resultado, la matriz numérica sería menos completa y esto podría generar algunos falsos negativos, es decir, un comentario agresivo podría ser clasificado como un comentario no agresivo.
- Medidas de consenso. Por lo general, el consenso se determina observando la diferencia entre las evaluaciones de los expertos. Sin embargo, esto es incompleto ya que el comportamiento de los expertos puede ser tan relevante como las evaluaciones. Por ejemplo, dos expertos pueden estar de acuerdo en que una determinada alternativa es mejor que otra, pero las razones dadas por cada experto pueden ser diferentes. En consecuencia, ambos expertos tienen la misma opinión, pero los argumentos considerados para dar esa opinión son diferentes. Por esta razón, las medidas de consenso propuestas en este estudio no solo consideran las evaluaciones, sino también el comportamiento mostrado para proporcionarlas. Esto hace que las medidas de consenso sean más completas que si solo se consideraran las evaluaciones de los expertos.

References

[1] J. R. Trillo, E. Herrera-Viedma, J. A. Morente-Molinera, and F. J. Cabrerizo, "A group decision-making method based on the experts' behavior during the debate," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 9, pp. 5796–5808, 2023.

[2] J. A. Morente-Molinera, G. Kou, K. Samuylov, F. J. Cabrerizo, and E. Herrera-Viedma, "Using argumentation in expert's debate to analyze multi-criteria group decision making method results," *Information Sciences*, vol. 573, pp. 433–452, 2021.

[3] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14–46, 2015.

Retículos de conceptos multiadjuntos con intensificadores

M. Eugenia Cornejo
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
mariaeugenia.cornejo@uca.es

Jesús Medina
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
jesus.medina@uca.es

Francisco José Ocaña
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
franciscojose.ocana@uca.es

Resumen*—Este trabajo recopila resultados recientes sobre el uso de intensificadores en el marco de los retículos de conceptos multiadjuntos que se han presentado en el artículo científico *Attribute implications in multi-adjoint concept lattices with hedges, Fuzzy Sets and Systems 479:108854, 2024.

***Palabras clave*—Retículos de conceptos multiadjuntos, operadores de formación de conceptos, intensificadores.**

I. Introducción

Desde que se introduce el Análisis de Conceptos Formales (FCA, de sus siglas en inglés) en los ochenta [1], se han propuesto numerosas generalizaciones difusas, a fin de poder trabajar con la información imprecisa y compleja presente en las bases de datos. El uso de intensificadores en FCA ha propiciado mejorar el modelado de bases de datos [4]. El marco difuso de los retículos de conceptos multiadjuntos [2] surge como un nuevo ambiente más general y flexible que engloba otros enfoques difusos presentes en la literatura. La incorporación de intensificadores en los operadores de formación de conceptos de los retículos multiadjuntos ha sido recientemente estudiada en [3]. A continuación, se presenta un resumen de las propiedades obtenidas para los operadores de formación de conceptos con intensificadores.

II. Retículos de conceptos multiadjuntos con intensificadores

Los intensificadores permiten abordar con mayor flexibilidad la modelización de la información deseada, lo que hace interesante su incorporación en los retículos de conceptos multiadjuntos [2]. En este trabajo, consideramos una generalización de la noción de intensificador introducida por Hájek [5], en términos de triples adjuntos [6].

Definición 1: Dado un conjunto parcialmente ordenado acotado superiormente $(P, \leq, \top)$ y un triple adjunto $(\&, \swarrow, \nwarrow)$ con respecto a $(P, \leq, \top)$, un $\swarrow$*-intensificador* es una función unaria $*\colon P \to P$ que satisface las siguientes propiedades, para todo $x, y, z \in P$:

$$*(\top) = \top \qquad \text{(condición de frontera)}$$
$$*(x) \leq x \qquad \text{(condicion subdiagonal)}$$
$$*(z \swarrow y) \leq *(z) \swarrow *(y) \qquad \text{(condición de } \swarrow\text{-regularidad)}$$
$$*(*(x)) = *(x) \qquad \text{(idempotencia)}$$

Análogamente, se define un $\nwarrow$*-intensificador* para $(P, \leq, \top)$.

Para presentar formalmente el marco de trabajo de los retículos de conceptos multiadjuntos con intensificadores, es necesario incluir previamente las nociones de marco multiadjunto y contexto.

Definición 2:

- $(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$ es un *marco multiadjunto* donde $(L_1, \preceq_1)$ y $(L_2, \preceq_2)$ son retículos completos, $(P, \leq)$ es un conjunto parcialmente ordenado y $(\&_i, \swarrow^i, \nwarrow_i)$ es un triple adjunto con respecto a $(L_1, \preceq_1)$, $(L_2, \preceq_2)$, $(P, \leq)$, para todo $i \in \{1, \ldots, n\}$.
- (A, B, R, σ) es un *contexto* donde A y B son conjuntos no vacíos (atributos y objetos, respectivamente), la aplicación $R\colon A \times B \to P$ es una relación P-difusa y $\sigma\colon A \times B \to \{1, \ldots, n\}$ asocia cualquier elemento en $A \times B$ con un triple adjunto particular del marco.

A continuación, presentamos las nociones de operadores de formación de conceptos y retículo de conceptos multiadjunto, considerando intensificadores.

Definición 3: Consideremos un marco multiadjunto $(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$, un contexto (A, B, R, σ), $*_A$ una familia arbitraria de $\swarrow$-intensificadores en L_1 y $*_B$ una familia arbitraria de $\nwarrow$-intensificadores en L_2, definidas como:

$$*_A = \{*_a\colon L_1 \to L_1 \mid *_a \text{ es un } \swarrow\text{-intensificador, para todo } a \in A\}$$
$$*_B = \{*_b\colon L_2 \to L_2 \mid *_b \text{ es un } \nwarrow\text{-intensificador, para todo } b \in B\}$$

- Los operadores de formación de conceptos con intensificadores $^{\uparrow_{*\sigma}}\colon L_2^B \to L_1^A$ y $^{\downarrow^{*\sigma}}\colon L_1^A \to L_2^B$ se definen, para todo $g \in L_2^B$, $f \in L_1^A$, $a \in A$, $b \in B$, como:

$$g^{\uparrow_{*\sigma}}(a) = \text{ínf}\{R(a, b') \swarrow^{\sigma(a,b')} *_{b'}(g(b')) \mid b' \in B\}$$
$$f^{\downarrow^{*\sigma}}(b) = \text{ínf}\{R(a', b) \nwarrow_{\sigma(a',b)} *_{a'}(f(a')) \mid a' \in A\}$$

Parcialmente subvencionado por el proyecto PID2019-108991GB-I00 financiado por MICIU/AEI/10.13039/501100011033, por el proyecto PID2022-137620NB-I00 financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE, por el proyecto TED2021-129748B-I00 financiado por MICIU/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR, y por el contrato predoctoral industrial PU/EPIF-FPI-GRUPOENERGETICOPUERTOREAL/CP/2022-051, correspondiente al Programa de Fomento e Impulso de la Investigación y la Transferencia de la Universidad de Cádiz 2018/2019.

- El contexto asociado con los operadores de formación de conceptos previos se denota como $(A_{*_A}, B_{*_B}, R, \sigma)$. Un *concepto* es un par $\langle g, f\rangle$ satisfaciendo que $g \in L_2^B$, $f \in L_1^A$ y $g^{\uparrow_{*\sigma}} = f$, $f^{\downarrow^{*\sigma}} = g$.
- El *retículo de conceptos multiadjunto con intensificadores*, denotado como $(\mathcal{M}_*, \preceq)$, asociado al marco multiadjunto y al contexto $(A_{*_A}, B_{*_B}, R, \sigma)$, es el conjunto:

$$\mathcal{M}_* = \{\langle g, f\rangle \mid g \in L_2^B, f \in L_1^A, g^{\uparrow_{*\sigma}} = f, f^{\downarrow^{*\sigma}} = g\}$$

  donde el orden viene definido por $\langle g_1, f_1\rangle \preceq \langle g_2, f_2\rangle$ si y solo si $g_1 \preceq_2 g_2$ (o equivalentemente, $f_2 \preceq_1 f_1$).

A fin de simplificar la notación, escribiremos $(^{\uparrow_*}, ^{\downarrow^*})$ en lugar de $(^{\uparrow_{*\sigma}}, ^{\downarrow^{*\sigma}})$. Conviene mencionar que estos operadores no forman una conexión de Galois antítona, en general, como se muestra en el siguiente ejemplo.

Ejemplo 1: Consideremos el marco multiadjunto $([0,1]_2, \leq, (\&_{\mathrm{DP}}, \swarrow^{\mathrm{DP}}, \nwarrow_{\mathrm{DP}}))$, los contextos (A, B, R, σ) y $(A_{*_A}, B_{*_B}, R, \sigma)$, donde $[0,1]_2 = \{0, 0.5, 1\}$ es una partición del intervalo unidad en dos trozos, $A = \{a_1, a_2, a_3\}$, $B = \{b_1, b_2, b_3\}$, la relación $R\colon A \times B \to [0,1]_2$ viene dada por la Tabla I, $(\&_{\mathrm{DP}}, \swarrow^{\mathrm{DP}}, \nwarrow_{\mathrm{DP}})$ es el triple adjunto definido a partir de la discretización de la t-norma producto [3] en $[0,1]_2$, σ es constantemente $(\&_{\mathrm{DP}}, \swarrow^{\mathrm{DP}}, \nwarrow_{\mathrm{DP}})$ y las familias

$$\begin{aligned} *_A &= \{*_a\colon [0,1]_2 \to [0,1]_2 \mid *_a \text{ es el } \swarrow\text{- intensificador globalización, para todo } a \in A\} \\ *_B &= \{*_b\colon [0,1]_2 \to [0,1]_2 \mid *_b \text{ es el } \nwarrow\text{-intensificador identidad, para todo } b \in B\} \end{aligned}$$

Tabla I

R	a_1	a_2	a_3
b_1	0.5	1	0.5
b_2	1	0.5	0
b_3	0	0.5	1

Valores para calcular $f^{\downarrow^*}$	a_1	a_2	a_3
$*_a(f)$	1	0	0
$R(a,b_1) \nwarrow_{\mathrm{DP}} *_a(f(a))$	0.5	1	1
$R(a,b_2) \nwarrow_{\mathrm{DP}} *_a(f(a))$	1	1	1
$R(a,b_3) \nwarrow_{\mathrm{DP}} *_a(f(a))$	0	1	1

Valores para calcular $f^{\downarrow^*\uparrow_*}$	b_1	b_2	b_3
$*_b(f^{\downarrow^*})$	0.5	1	0
$R(a_1,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b))$	1	1	1
$R(a_2,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b))$	1	0.5	1
$R(a_3,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b))$	1	0	1

Dado $f = \{1/a_1, 0.5/a_3\}$, calculamos su clausura $f^{\downarrow^*\uparrow_*}$. Aplicando la definición de los operadores $^{\downarrow^*}$ y $^{\uparrow_*}$, así como teniendo en cuenta los datos de la Tabla I, obtenemos:

$$\begin{aligned} f^{\downarrow^*}(b_1) &= \text{ínf}\{R(a,b_1) \nwarrow_{\mathrm{DP}} *_a(f(a)) \mid a \in A\} = 0.5 \\ f^{\downarrow^*}(b_2) &= \text{ínf}\{R(a,b_2) \nwarrow_{\mathrm{DP}} *_a(f(a)) \mid a \in A\} = 1 \\ f^{\downarrow^*}(b_3) &= \text{ínf}\{R(a,b_3) \nwarrow_{\mathrm{DP}} *_a(f(a)) \mid a \in A\} = 0 \\ f^{\downarrow^*\uparrow_*}(a_1) &= \text{ínf}\{R(a_1,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b)) \mid b \in B\} = 1 \\ f^{\downarrow^*\uparrow_*}(a_2) &= \text{ínf}\{R(a_2,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b)) \mid b \in B\} = 0.5 \\ f^{\downarrow^*\uparrow_*}(a_3) &= \text{ínf}\{R(a_3,b) \swarrow^{\mathrm{DP}} *_b(f^{\downarrow^*}(b)) \mid b \in B\} = 0 \end{aligned}$$

Por tanto, $f^{\downarrow^*\uparrow_*} = \{1/a_1, 0.5/a_2\}$. En consecuencia, existe $f \in [0,1]_2^A$ tal que $f \not\leq f^{\downarrow^*\uparrow_*}$, contradiciendo las propiedades exigidas en la definición de conexión de Galois. Este hecho nos lleva a concluir que $(^{\uparrow_*}, ^{\downarrow^*})$ no es una conexión de Galois antítona, en general.

III. Propiedades de los operadores de formación de conceptos con intensificadores

En esta sección se presentan propiedades básicas de los operadores $^{\uparrow_*}$ y $^{\downarrow^*}$ que tienen cierta similitud con las propiedades de las conexiones de Galois antítonas. Antes de introducir las propiedades mencionadas, se hace necesario considerar el siguiente subconjunto difuso de atributos:

$$\bar{*}_A(f)\colon A \to L_1, \text{ definido como } \bar{*}_A(f)(a) = *_a(f(a)), \text{ para todo } *_a \in *_A,\ a \in A \text{ y } f \in L_1^A.$$

y el siguiente subconjunto difuso de objetos:

$$\bar{*}_B(g)\colon B \to L_2, \text{ definido como } \bar{*}_B(g)(b) = *_b(g(b)), \text{ para todo } *_b \in *_B,\ b \in B \text{ y } g \in L_2^B.$$

Proposición 1: Consideremos un marco multiadjunto $(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$ y los contextos (A, B, R, σ) y $(A_{*_A}, B_{*_B}, R, \sigma)$. Los operadores de formación de conceptos $^{\uparrow_*}$ y $^{\downarrow^*}$ satisfacen las siguientes propiedades, para todo $g \in L_2^B$, $f \in L_1^A$, $a \in A$, $b \in B$:

$g^{\uparrow} \preceq_1 g^{\uparrow_*}$	$f^{\downarrow} \preceq_2 f^{\downarrow^*}$
$\bar{*}_B(g) \preceq_2 g^{\uparrow_*\downarrow^*}$	$\bar{*}_A(f) \preceq_1 f^{\downarrow^*\uparrow_*}$
$g^{\uparrow_*} = \bar{*}_B(g)^{\uparrow_*}$	$f^{\downarrow^*} = \bar{*}_A(f)^{\downarrow^*}$

Además, si asumimos que $L_1 = L_2 = P$ y que los conjuntores de los triples adjuntos del marco multiadjunto satisfacen las condiciones de frontera, para todo $g, g_1, g_2 \in L_1^B$ y $f, f_1, f_2 \in L_1^A$, se cumple que:

If $g_1 \preceq_1 g_2$ then $g_2^{\uparrow_*} \preceq_1 g_1^{\uparrow_*}$	If $f_1 \preceq_1 f_2$ then $f_2^{\downarrow^*} \preceq_1 f_1^{\downarrow^*}$
$\bar{*}_A(g^{\uparrow_*}) \preceq_1 g^{\uparrow_*\downarrow^*\uparrow_*} \preceq_1 g^{\uparrow_*}$	$\bar{*}_B(f^{\downarrow^*}) \preceq_1 f^{\downarrow^*\uparrow_*\downarrow^*} \preceq_1 f^{\downarrow^*}$
$\bar{*}_A(g^{\uparrow_*}) = \bar{*}_A(g^{\uparrow_*\downarrow^*\uparrow_*})$	$\bar{*}_B(f^{\downarrow^*}) = \bar{*}_B(f^{\downarrow^*\uparrow_*\downarrow^*})$
$g^{\uparrow_*\downarrow^*} = g^{\uparrow_*\downarrow^*\uparrow_*\downarrow^*}$	$f^{\downarrow^*\uparrow_*} = f^{\downarrow^*\uparrow_*\downarrow^*\uparrow_*}$

IV. Conclusiones y trabajo futuro

En este artículo se han presentado propiedades de los operadores de formación de conceptos con intensificadores en el marco multiadjunto. Como trabajo futuro, se profundizará en el estudio de estos operadores, a fin de poder definir la estructura algebraica de los retículos de conceptos multiadjuntos con intensificadores.

Referencias

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

[2] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño, "Formal concept analysis via multi-adjoint concept lattices," *Fuzzy Sets and Systems*, vol. 160, no. 2, pp. 130–144, 2009.

[3] M. E. Cornejo, J. Medina, and F. J. Ocaña, "Attribute implications in multi-adjoint concept lattices with hedges," *Fuzzy Sets and Systems*, vol. 479, p. 108854, 2024.

[4] R. Bělohlávek and V. Vychodil, "Formal concept analysis and linguistic hedges," *Int. J. General Systems*, vol. 41, no. 5, pp. 503–532, 2012.

[5] P. Hájek, "On very true," *Fuzzy Sets and Systems*, vol. 124, no. 3, pp. 329 – 333, 200.

[6] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa, "Algebraic structure and characterization of adjoint triples," *Fuzzy Sets and Systems*, vol. 425, pp. 117–139, 2021.

Clasificación de nuevos objetos en teoría de conjuntos rugosos difusos

Fernando Chacón-Gómez
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
fernando.chacon@uca.es

M. Eugenia Cornejo
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
mariaeugenia.cornejo@uca.es

Jesús Medina
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
jesus.medina@uca.es

***Resumen*—Las reglas de decisión son una herramienta útil tanto para la extracción de información de conjuntos de datos, como para inferir el comportamiento de nuevos datos. Este trabajo propone dos métodos para clasificar nuevos objetos en el marco de la teoría de conjuntos rugosos difusos. Ambos métodos tienen como punto de partida la comparativa entre el nuevo dato a clasificar y las reglas de decisión consideradas en el estudio. La principal diferencia radica en la noción difusa en la que se sustentan, grado de idoneidad y grado de representatividad.**

***Palabras clave*—Teoría de conjuntos rugosos difusos, reglas de decisión, métodos de clasificación.**

I. Introducción

La teoría de conjuntos rugosos [7]–[9] y la teoría de conjuntos rugosos difusos (FRST, de sus siglas en inglés, *Fuzzy Rough Set Theory*) [4], [6], [11], son herramientas matemáticas eficaces para el tratamiento de información incompleta o imprecisa en conjuntos de datos relacionales, interpretados como tablas de decisión. Ambas teorías se apoyan en reglas de decisión [5] para el manejo de dichas tablas en términos lógicos, facilitando así considerablemente el análisis de la información y la extracción del conocimiento. Este trabajo presenta dos métodos de clasificación de nuevos objetos, propuestos en [1], que se sustentan en la filosofía de FRST y en las reglas de decisión. La clasificación de nuevos objetos es una tarea íntimamente relacionada con la toma de decisiones, siendo esta última un problema de gran interés en la actualidad, puesto que aborda multitud de procesos relacionados con el descubrimiento de patrones y su aplicación a nuevos escenarios, entre otros.

II. Teoría de conjuntos rugosos difusos

Esta sección recuerda las nociones más importantes de FRST [2] para el desarrollo de este trabajo. Para comenzar, presentamos las tablas de decisión, que permiten el tratamiento de conjuntos de datos en términos matemáticos.

Definición 1. *Sean U y $\mathcal{A}$ conjuntos no vacíos de objetos y atributos, respectivamente. Una* tabla de decisión *es una tupla $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ tal que $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ con $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, donde V_a es el conjunto de valores asociados al atributo a sobre U, y $\overline{\mathcal{A}_d} = \{\bar{a}\colon U \to V_a \mid a \in \mathcal{A}_d\}$.*

A continuación, introducimos diversas nociones fundamentales para expresar la información contenida en tablas de decisión en términos lógicos. Es conveniente destacar que, con el objetivo de realizar estudios más precisos, en lo que sigue consideraremos relaciones de tolerancia separables [10].

Definición 2. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $T = \{T_a\colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ una familia de $[0,1]$-relaciones de tolerancia difusas separables, $B \subseteq \mathcal{A}_d$ y $C \subseteq \mathcal{A}$.*

- *El* conjunto de fórmulas asociadas a B, *denotado como $For(B)$, se construye a partir de pares atributo valor (a, v), donde $a \in B$ y $v \in V_a$, por medio del conectivo lógico de conjunción $\wedge$.*
- *La aplicación $\|\cdot\|_S^T\colon For(B) \to [0,1]^U$ se define inductivamente como:*

$$\|\Phi\|_S^T(x) = T_a(\bar{a}(x), v)$$

para todo $x \in U$ y $\Phi = (a, v)$, donde $a \in B$ y $v \in V_a$. Por tanto, $\|\Phi\|_S^T(x)$ es el grado de satisfacibilidad del objeto x con respecto a la fórmula Φ.
- *Para cada $\Phi, \Psi \in For(B)$, la* conjunción de fórmulas *se define, para todo $x \in U$, como sigue:*

$$\|\Phi \wedge \Psi\|_S^T(x) = \min\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$

- *Una* regla de decisión *en S es una expresión $\Phi \to \Psi$, donde $\Phi \in For(C), \Psi \in For(\{d\})$ son el antecedente y consecuente de la regla de decisión, respectivamente.*

Finalmente, presentamos la noción de T-soporte en FRST. Este indicador de relevancia proporciona la representatividad de la regla de decisión dada en la tabla de decisión bajo consideración, y viene dado en función del cardinal difuso [3].

Definición 3. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $T = \{T_a\colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}_d\}$ una familia de $[0,1]$-relaciones de tolerancia difusas separables y $\Phi \to \Psi$ una regla de decisión en S. Se denomina* T-soporte *de la regla de decisión $\Phi \to \Psi$ al valor:*

$$supp_S^T(\Phi, \Psi) = card_F(\|\Phi \wedge \Psi\|_S^T)$$

Parcialmente subvencionado por el proyecto PID2019-108991GB-I00 financiado por MICIU/AEI/10.13039/501100011033, por el proyecto PID2022-137620NB-I00 financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE, por el proyecto TED2021-129748B-I00 financiado por MICIU/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR, por el proyecto PR2023-009 financiado por la Universidad de Cádiz y por el Programa de Fomento e Impulso de la Investigación y la Transferencia de la Universidad de Cádiz.

Tabla I
RESUMEN DE CADA MÉTODO DE CLASIFICACIÓN PRESENTADO EN LA SECCIÓN III.

	Método de Idoneidad	**Método de ε-representatividad**
Noción difusa	$\sigma_{Dec(S)}$	$\rho^{\varepsilon}_{Dec(S)}$
Paso 1	Calcular la noción difusa anterior para cada decisión posible	
Paso 2	Comparar los valores obtenidos en el paso 1 y tomar decisión si alguna es concluyente	
Paso 3	Agregar más valores, comparar los resultados y tomar decisión si alguna es concluyente	Cambiar el umbral ε y volver al paso 1
Paso 4	Considerar otra agregación y volver al paso 2	

donde $card_F(\cdot)$ denota el cardinal de un conjunto difuso.

III. MÉTODOS DE CLASIFICACIÓN

En este trabajo, la clasificación de nuevos datos se basa en la información extraída de una tabla de decisión por medio de un conjunto de reglas de decisión. De esta forma, dado un objeto no perteneciente a dicha tabla, y del que solo se desconoce el valor del atributo de decisión d, se pretende determinar el mejor valor en dicho atributo. Para ello, la información conocida sobre este objeto se compara con la dada por cada una de las reglas anteriores. De ahora en adelante, el universo de objetos se define como unión de dos conjuntos, $\mathcal{U} = U \cup U^c$, donde U son los objetos de la tabla de decisión dada y U^c los nuevos objetos a clasificar. El primer método de clasificación introducido en [1] se basa en la siguiente definición.

Definición 4. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $Dec(S)$ un conjunto de reglas de decisión. Para cada nuevo objeto $x \in U^c$ y $v \in V_d$, el* grado de idoneidad de la decisión $\bar{d}(x) = v$ de acuerdo a $Dec(S)$ *viene dado por la aplicación $\sigma_{Dec(S)} \colon U^c \times V_d \to [0,1]$, definida como:*

$$\sigma_{Dec(S)}(x, v) = \text{máx}\{\|\Phi\|^T_S(x) \mid \Phi \to \Psi \in Dec(S), \Psi = (d, v)\}$$

El primer método de clasificación propuesto comienza calculando el grado de idoneidad de cada decisión $v \in V_d$ para el objeto nuevo $x \in U^c$. Si existe una decisión cuyo grado de idoneidad sea elevado y suficientemente diferente del resto entonces se toma dicha decisión. En caso caso contrario, en la definición de $\sigma_{Dec(S)}$ se puede cambiar el máximo por otro operador de agregación que no tenga en cuenta solo el máximo, y se vuelven a comparar los resultados obtenidos.

El segundo método de clasificación se introduce en [1] con el fin de tener en cuenta un indicador tan importante como es el T-soporte de las reglas en los casos en los que el primer método no arroje una clasificación clara del nuevo objeto.

Definición 5. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $Dec(S)$ un conjunto de reglas de decisión. Dado $\varepsilon \in [0,1]$, para cada nuevo objeto $x \in U^c$ y $v \in V_d$, el* grado de ε-representatividad de la decisión $\bar{d}(x) = v$ de acuerdo a $Dec(S)$ *viene dado por la aplicación $\rho^{\varepsilon}_{Dec(S)} \colon U^c \times V_d \to [0,1]$, definida como:*

$$\rho^{\varepsilon}_{Dec(S)}(x, v) = \frac{\sum_{\{\Phi \to \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|^T_S(x), \Psi = (d,v)\}} \|\Phi\|^T_S(x) \cdot supp^T_S(\Phi, \Psi)}{\sum_{\{\Phi \to \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|^T_S(x)\}} \|\Phi\|^T_S(x) \cdot supp^T_S(\Phi, \Psi)}$$

El umbral ε es fundamental para seleccionar aquellas reglas con antecedentes más similares al objeto considerado, y permite al usuario adaptar los estudios a sus necesidades. La filosofía del segundo método es similar a la del primero, calculando en primer lugar el grado de ε-representatividad de cada decisión posible y tomando una decisión si su grado es elevado y suficientemente diferente del resto. En otro caso, el usuario debe cambiar el umbral ε para recalcular el grado de ε-representatividad y poder tomar una decisión robusta.

Los métodos de clasificación presentados se muestran resumidos en la Tabla I y se desarrollan con mayor detalle en [1].

IV. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se ha abordado el estudio de la toma de decisiones en FRST por medio de la clasificación de nuevos objetos. Con este propósito, se han presentado dos métodos que afrontan este proceso desde perspectivas diferentes ofreciendo conclusiones complementarias. Como trabajo futuro, se investigarán nuevos métodos de clasificación empleados en otros ámbitos, como en la minería de datos, y se extenderán al marco de trabajo de FRST. También se analizará cómo afecta la reducción de atributos en la toma de decisiones, prestando especial atención al uso de reductos y algoritmos de decisión que conserven la eficiencia tras realizar la reducción correspondiente.

REFERENCIAS

[1] F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Decision making in Fuzzy Rough Set Theory. *Mathematics*, 11(19), 2023.

[2] F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Rough set decision algorithms for modeling with uncertainty. *Journal of Computational and Applied Mathematics*, 437:115413, 2024.

[3] A. De Luca and S. Termini. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, 20(4):301–312, 1972.

[4] D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17(2-3):191–209, 1990.

[5] J. Li, C. Mei, and Y. Lv. Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction. *International Journal of Approximate Reasoning*, 54(1):149–165, 2013.

[6] A. Nakamura. Fuzzy rough sets. *Note on Multiple-Valued Logic in Japan*, 9:1–8, 1988.

[7] Z. Pawlak. Information systems theoretical foundations. *Information Systems*, 6(3):205 – 218, 1981.

[8] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.

[9] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992.

[10] L. Zedam, H. Bouremel, and B. De Baets. Left- and right-compatibility of order relations and fuzzy tolerance relations. *Fuzzy Sets and Systems*, 360:65–81, 2019. Theme: Fuzzy Relations.

[11] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

Enriquecimiento de explicaciones interactivas en asistentes conversacionales usando redes de restricciones temporales difusas

Jose M. Alonso-Moral, Alejandro Catala, Alberto Bugarín-Diz
Centro Singular de Investigación en Tecnologías Intelixentes (CiTIUS), Universidade de Santiago de Compostela
Santiago de Compostela, Spain
{josemaria.alonso.moral, alejandro.catala, alberto.bugarin.diz}@usc.es

***Resumen*—En este *keywork* presentamos un resumen del trabajo titulado *Enriching interactive explanations with fuzzy temporal constraint networks*, publicado originalmente en la revista *International Journal of Approximate Reasoning* en enero de 2024 [1]. La versión inicial del trabajo fue enviada el 13 de enero de 2023, fue aceptado finalmente el 14 de enero de 2024 y está disponible en línea desde el 17 de enero de 2024 en acceso abierto.**

***Index Terms*—Redes de Restricciones Temporales Difusas, Razonamiento Temporal Difuso, Grafos de Conocimiento, Grandes Modelos de Lenguaje, Agentes Conversacionales**

I. Introducción

Los modelos de lengua preentrenados a gran escala (Large Language Models, LLM) dan excelentes resultados en muchas tareas lingüísticas, pero presentan algunos inconvenientes como su falta de transparencia y de capacidad de razonamiento temporal exhaustivo. Por lo tanto, el uso de tales modelos puede provocar diálogos incoherentes o incorrectos en el contexto de los agentes conversacionales cuyo objetivo sea proporcionar a los usuarios de sistemas inteligentes explicaciones interactivas. En el trabajo "Enriching interactive explanations with fuzzy temporal constraint networks"publicado originalmente en la revista *International Journal of Approximate Reasoning* [1], propusimos un modelo de razonamiento temporal difuso para superar algunas inconsistencias detectadas en la aplicación de LLM al diseño de un agente conversacional. Más concretamente, partiendo de un grafo que proporciona una representación intuitiva de las entidades y relaciones en el dominio de aplicación, describimos cómo mapear la información temporal en una red difusa de restricciones temporales. Este formalismo nos permitió representar información temporal imprecisa y proporcionar mecanismos para comprobar coherencia en las conversaciones. Además, como prueba de concepto, desarrollamos el agente conversacional TimeVersa, que integra el modelo propuesto en un dominio de aplicación (información horaria de transportes) que requiere el manejo de restricciones temporales imprecisas. Ilustramos en un caso de uso cómo el agente puede identificar inconsistencias temporales y responder a consultas relacionadas con información temporal correctamente. Los resultados de un estudio de usuarios indican que la percepción de coherencia por parte de los usuarios es significativamente mayor en una conversación con TimeVersa que en una conversación similar utilizando el LLM GPT-3, cuando se trata de información temporal imprecisa. Nuestro enfoque es un paso adelante para el desarrollo de agentes conversacionales que operen en ámbitos que requieren razonamiento temporal en condiciones de incertidumbre.

A pesar de sus impresionantes capacidades, los LLM plantean graves problemas. En primer lugar, su entrenamiento requiere enormes capacidades de computación, lo que supone un enorme consumo de energía y un reto desde el punto de vista de la sostenibilidad. En segundo lugar, debido a su naturaleza opaca, no se entiende bien cómo funcionan, cómo se han entrenado y validado, ni qué propiedades emergentes presentan. En casos concretos como GPT-3 (único modelo de OpenAI disponible de forma gratuita), se han identificado otras limitaciones como que los resultados pueden carecer de coherencia semántica, lo que da lugar a textos incomprensibles, especialmente cuando se alarga la conversación, los resultados incorporan todos los sesgos que pueden encontrarse en los datos de entrenamiento, o pueden corresponder a afirmaciones que no están en consonancia con la verdad. Además, el rendimiento mostrado en tareas de Procesamiento del Lenguaje Natural con lenguas de escasos recursos (por ejemplo, euskera, gallego o gaélico) es mucho peor que en inglés, porque dichas lenguas estaban infrarrepresentadas en los datos de entrenamiento. Cabe señalar que, aunque versiones posteriores como GPT-4 anuncian que resuelven algunas de las limitaciones anteriores, su naturaleza opaca permanece, ya que GPT-4 sigue siendo un modelo de caja negra distribuido en forma de software como servicio y es imposible validar su rendimiento desde un punto de vista científico riguroso. Estas deficiencias son comunes a otros LLM que también podrían devolver resultados incoherentes en algunos casos. La posible incoherencia en los resultados de los LLM podría explicarse por su falta de razonamiento lógico o porque sus

Esta publicación es parte del proyecto de I+D+i TED2021-130295B-C33, financiado por MCIN/AEI/10.13039/501100011033/ y por la "Unión Europea NextGenerationEU/PRTR". Además, esta investigación contribuye a los Proyectos PID2021-123152OB-C21 y PID2020-112623GB-I00 financiados por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa. Se reconoce el apoyo de la Consellería de Educación, Universidades y Formación Profesional de la Xunta de Galicia y del Fondo Europeo de Desarrollo Regional (ayudas ED431G2019/04 y ED431C2022/19). La investigación cuenta con financiación del "European Union's Horizon 2020 Research and Innovation Programme"(Marie Skłodowska-Curie Grant Agreement No. 860621).

capacidades deductivas e inductivas son muy simples. Por ejemplo, esos modelos no comprenden bien el razonamiento temporal y su memoria a veces se queda corta, por lo que no pueden tener en cuenta incluso una restricción temporal simple que pudiera haberse introducido unos pocos mensajes antes. También sufren el llamado olvido catastrófico cuando se afinan para tareas específicas. Además, una dificultad añadida es la continua presencia de expresiones imprecisas en el uso del lenguaje natural por parte de los humanos, que a menudo utilizamos expresiones que incluyen términos vagos o imprecisos como “más o menos”, “un poco más”, “no tanto”, “aproximadamente”, etc. Estas expresiones captan la imprecisión del lenguaje y resumen la información de forma adecuada para la comunicación humana. Hay que tener en cuenta que el razonamiento automático y, más concretamente, el razonamiento temporal se vuelven más difíciles de manejar cuando se trata con este tipo de expresiones. Teniendo en cuenta todo lo anterior, la principal contribución del trabajo [1] fue enriquecer las explicaciones interactivas con un modelo que permita razonar con información temporal imprecisa. Con este objetivo, propusimos un modelo que combina una Red Difusa de Restricciones Temporales con un grafo de conocimiento, y describimos cómo opera en un agente conversacional cuidadosamente diseñado para proporcionar a los usuarios de sistemas inteligentes explicaciones interactivas dentro de un dominio de aplicación específico. En concreto, desarrollamos un caso de uso ilustrativo para el sector turístico para mostrar cómo el agente maneja adecuadamente algunos problemas potenciales de inconsistencia temporal. Además, para llevar a cabo un proceso de validación exhaustivo del modelo descrito, también llevamos a cabo un estudio empírico de usuarios en el que los participantes pudieron evaluar algunos aspectos de las conversaciones relacionados con el razonamiento temporal.

II. Propuesta

Nuestro modelo para gestionar expresiones temporales difusas y detectar incoherencias entre ellas parte de un grafo de conocimiento (KG) representativo del dominio específico, y proyecta la información temporal sobre un grafo difuso de restricciones temporales, que proporciona mecanismos de razonamiento para comprobar la coherencia y responder a consultas explicativas. Más concretamente, suponemos que el KG es conocido, está dado de antemano, representa el conocimiento sobre un dominio de aplicación específico y proporciona una abstracción concisa e intuitiva de los datos implicados. A continuación, utilizamos el conocido formalismo de las Redes de restriccions temporales difusas (FTCN), para disponer de un mecanismo que nos permite representar información temporal difusa y razonar con ella. Seguidamente, describimos un algoritmo para proyectar la información de dominio contenida en el KG sobre la FTCN, lo que nos va a permitir formular y responder consultas acerca de la información contenida en ella. El procedimiento de proyección comienza por los nodos del KG conectados por una relación temporal y continúa con la proyección de las relaciones no temporales que puedan existir.

III. Caso de uso y Validación

En el trabajo presentamos un caso de uso ilustrativo para mostrar cómo nuestro modelo aborda varias incoherencias relacionadas con el razonamiento temporal. Para facilitar la reproducibilidad, y de acuerdo con las directrices para el desarrollo de una IA responsable, hemos puesto a disposición el código para ejecutar este caso de uso ilustrativo. El caso de uso es el agente conversacional TimeVersa para informar, recomendar, reservar y dar información sobre horarios de viajes, al objeto de centrar el caso en un contexto de información temporal. El agente puede responder adecuadamente a preguntas sobre duración y horario de los barcos que realizan un trayecto, disponibilidad de restaurantes, y duración de rutas de senderismo. También ayuda a los usuarios a reservar el restaurante o comprar los billetes para el viaje. Como validación comparamos el comportamiento del modelo propuesto con el de uno de de los modelos lingüísticos preentrenados disponibles gratuitamente y más populares (GPT-3) en un caso de uso similar y planteamos un estudio empírico con evaluadores humanos para examinar las diferencias en la percepción que tienen los usuarios en cuanto a la consistencia de una conversación con un agente, dependiendo de si el agente incluía nuestro modelo de razonamiento temporal difuso (agente TimeVersa) o no lo incluía (agente OpenAI, con el LLM GPT-3). El objetivo fue validar las dos hipótesis de investigación siguientes:

- H1: Cuando se incluye información temporal imprecisa, la interacción con el agente TimeVersa se percibe más coherente que con el agente OpenAI.
- H2: Cuando se incluye información temporal imprecisaa, la interacción con el agente TimeVersa que incorpora el modelo de razonamiento temporal se percibe como más coherente que cuando incorpora la versión nítida del modelo.

Los resultados del estudio, en el que participaron 130 usuarios humanos a través de la plataforma Prolific, indicaron que las diferencias observadas en cuanto a la consistencia de la conversación con el agente TimeVersa frente al agente OpenAI fueron estadísticamente significativas a favor de TimeVersa. Asimismo, encontramos una mejor percepción al utilizar un modelo difuso frente a un modelo nítido en cuanto a la coherencia de la conversación, con diferencias estadísticamente significativas, si bien resultaron pequeñas en términos absolutos.

Referencias

[1] M. Canabal-Juanatey, J. M. Alonso-Moral, A. Catala, A. Bugarín-Diz, ”Enriching interactive explanations with fuzzy temporal constraint networks”, International Journal of Approximate Reasoning, 2024, 109128, ISSN 0888-613X. https://doi.org/10.1016/j.ijar.2024.109128.

Aggregation functions for random elements over bounded lattices

Juan Baz
Department of Statistics and O.R and D.M.
University of Oviedo
Oviedo, Spain
bazjuan@uniovi.es

Irene Díaz
Deparment of Computer Science
University of Oviedo
Gijón, Spain
sirene@uniovi.es

Susana Montes
Department of Statistics and O.R and D.M.
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—Aggregation theory is devoted to the fusion of information from different sources in a unique final value. In some applied problems, the aggregated information is associated to an empirical measurement that can be interpreted as a random sampling of a population. With this point of view, it is quite natural to consider the inputs of the aggregation function to be realizations of random elements. In this paper, we define the concept of aggregation of random elements over bounded lattices from the very beginning. The boundary and monotonicity conditions are generalized by using stochastic orders. In addition, several properties and particular examples are provided.**

***Index Terms*—Aggregation, Probability Theory, Partial orders**

I. Introduction

Aggregation theory focuses on the study of functions that take n elements of certain type and return another element of the same type, fulfilling some boundary and monotonicity conditions. Besides its use in fuzzy set theory or dependence modeling, they appear frequently in data analysis [1, 2]. Considering the usual probabilistic approach made in Statistics, we can consider the inputs of the aggregations to have a random behavior. In this direction, the notion of aggregation of random variables was introduced in [3], redefining the boundary and monotonicity conditions by means of stochastic orders.

The notions and results summarized in this paper are deeply studied in [4], in which aggregations of random elements over bounded lattices were introduced. This paper is a summary of that work and it is organized as follows. The construction of the definition of aggregations of random elements over bounded lattices is provided in Section II. Section III is devoted to the study of some desirable properties. Finally, some examples and the conclusions are provided in Sections IV and V.

II. Aggregation of random elements over bounded lattices

Let us consider $(S, \leq)$ as a partially ordered set. We recall that an upper set of $(S, \leq)$ is a subset $U \subseteq S$ such that for any $u \in U$, if $s \geq u$ then $s \in U$. Similarly a lower set is defined as $L \subseteq S$ such that for any $l \in L$, if $s \leq l$ then $s \in L$.

The publication was supported by the Spanish Ministry of Economy and Competitiveness (PGC2018-099402-B-I00) and the Principality of Asturias (BP21048).

In order to define random elements over S, it must be endowed with a sigma-algebra F. For constructing this sigma-algebra, we can consider first the semi sigma-algebra (Proposition 3.1 in [4]):

$$F_s = \{A \subseteq S : A = A_L \cap A_U, \\ A_L \text{ lower set of } (S, \leq), A_U \text{ upper set of } (S, \leq)\}$$

and then the sigma-algebra

$$F = \{B \subseteq S : B = \cup_{i=1}^{\infty} B_i, B_i \in F_s\}.$$

The resulting pair (S, F) is a measurable space. For any pair of elements $r, t \in S$ such that $r \leq t$, the subset consisting of the elements greater than r and smaller than t will be denoted as $[r, t] = \{s \in S : r \leq s \leq t\}$. We can extend the partially ordered set and the measurable space structure to the Cartesian product of S, $S^n = S \times \cdots \times S$. The partially ordered set $(S^n, \leq_{po})$ is defined by considering the product order on S^n using as reference the initial order $\leq$. The sigma-algebra F^n is defined using a semi sigma-algebra consisting on Cartesian products of elements of F_s (Proposition 3.2 in [4]).

Starting from a probabilistic space (Ω, Σ, P), we define as $L^n_{[r,t]}$ the set of measurable functions from Ω to S^n with support $[r, t]^n$. In symbols,

$$L^n_{[r,t]} = \{f : \Omega \to S^n : f \text{ is measurable, } f^{-1}([r, t]^n) = \Omega\}$$

To provide a way to order these functions, which can be considered as random vectors of elements of the bounded lattice, we consider the following stochastic order.

Definition 2.1: Let $f, g \in L^n_{[s,t]}$. If $P(f^{-1}(M)) \leq P(g^{-1}(M))$ for every M measurable upper set of S^n, then it is said that f is smaller than or equal to g in the usual stochastic order and it is denoted by $f \leq_{st} g$.

This stochastic order is transitive, i.e. $f \leq_{st} g, g \leq_{st} h \implies f \leq_{st} h$, fulfills that $f \leq_{st} g, g \leq_{st} h$ if and only if f and g have the same distribution and in coherent with respect to the initial order $\leq$ when considering degenerate distributions (Proposition 3.3 in [4]).

With all this settlement, we can define the concept of aggregation of random elements over bounded lattices.

Definition 2.2: The function $A : L^n_{[r,t]} \to L_{[r,t]}$ is an aggregation of random elements over a bounded lattice with respect to $\leq_{st}$ if the following conditions are fulfilled:

- For any $f, g \in L^n_{[r,t]}$ such that $f \leq_{st} g$, $A(f) \leq_{st} A(g)$.
- If $f(\Omega) = \{(r, \dots, r)\}$, $A(f)(\Omega) = \{r\}$.
- If $f(\Omega) = \{(t, \dots, t)\}$, $A(f)(\Omega) = \{t\}$.

III. Additional properties and families

The concept of aggregation of random elements over bounded lattices goes beyond the simply composing usual aggregation functions over lattices and random vectors, since its definition is not very restrictive. In the following, we define some families of aggregations of random elements over bounded lattices and some coherent properties that can be considered when additional restrictions are needed, for instance, in applied problems.

Definition 3.1: Let $A : L^n_{[r,t]} \to L_{[r,t]}$ an aggregation of random elements over a bounded lattice. Then:

- If there exists an aggregation function $\hat{A} : [r,t]^n \to [r,t]$ such that $A(\vec{X}) = \hat{A} \circ f$ for any $f \in L^n_{[r,t]}$, A is said to be induced (by $\hat{A}$).
- If there exists $f \in L^n_{[r,t]}$ and $\vec{v} \in f(\Omega)$ with $P(f = \vec{v}) > 0$ such that the conditional distribution of $A(f)$ given $f = \vec{v}$ is not degenerate, A is said to be random.
- If $A(f)$ has degenerate distribution for any $f \in L^n_{[r,t]}$, A is said to be degenerate.
- If for any $f, g \in L^n_{[r,t]}$ such that f and g have the same distribution, $A(f)$ and $A(g)$ have the same distribution, A is said to be coherent in distribution.
- If for any $\vec{v} \in [r,t]^n$ and any $f \in L^n_{[r,t]}$ such that $\vec{v} \in f(\Omega)$, $P(f = \vec{v}) > 0$ the conditional distribution of $A(f)$ given $f = \vec{v}$ is always the same, A is said to be conditional coherent.
- If for any $f \in L^n_{[r,t]}$, $P(f = (r, \dots, r)) \leq P(A(f) = r)$ and $P(f = (t, \dots, t)) \leq P(A(f) = t)$, A is said to be boundary coherent.

Let us discuss very briefly some results and relations between the introduced concepts. It can be proved that families of induced, random and degenerate aggregations of random elements over bounded lattices are disjoint (Proposition 3.4 in [4]). Distribution coherence is always fulfilled (Proposition 3.5 in [4]). If $r \neq t$, degenerate aggregations cannot be boundary coherent (Proposition 3.7 in [4]) . Finally, induced aggregations fulfill all three coherence properties (Theorem 3.2 in [4]).

IV. Some examples

The construction of aggregations of random elements over bounded lattices may seem very abstract, but there are many examples of applications for different structures.

The most prominent example is the aggregation of random variables, in which $(S, \leq)$ is considered as the real line $\mathbb{R}$ with the usual order of the real numbers. In the following, we provide an example of non-trivial aggregation of random variables.

Example 4.1: Let $r, t \in \mathbb{R}$ with $r < t$ and $\left(U_{\vec{X}}, \vec{X} \in L^n_{[r,t]}\right)$ a family of standard uniform random variables such that $U_{\vec{X}}$ is independent of $\vec{X}$ for any $\vec{X} \in L^n_{[s,t]}$. Consider the function $A : L^n_{[r,t]} \to L_{[r,t]}$ that takes a random vector $\vec{X}$ and returns the random variable $A(\vec{X}) = (\max(\vec{X}) - \min(\vec{X}))U_{\vec{X}} + \min(\vec{X})$ is an aggregation of random variables.

The latter aggregation is random, distribution coherent, conditional coherent and boundary coherent and it can be seen as a generalization of the estimation procedure for an uniform random variable.

Another interesting example is the aggregation of random graphs, that can be considered, for instance, in social choice theory or in reliability theory [5]. In this case, S is the set of graphs with a fixed set of nodes V and the order is considered to be the inclusion order of the edge sets, i.e. $(V, E_1) \leq (V, E_2) \iff E_1 \subseteq E_2$. The most important aggregations are the union and intersection of the edges.

As the last example, we can consider the aggregation of semi-positive definite matrices. Random semi-positive matrices appear in the estimation of the covariance matrix of a population, and its aggregation is relevant in some statistical models such as the Multivariate Analysis of Variance (MANOVA) [6], in which the covariance estimations in the different groups are fused in an unique estimation (under homoscedasticity). In this case, the natural order is the Loewner order, defined as $M_1 \leq M_2$ if and only if $M_2 - M_1$ is semi-positive definite.

V. Conclusions

In this paper, the concept of aggregation of random elements over bounded lattices is constructed. In addition, some properties, families and examples are provided. We aim this work to be one of the first steps in the inclusion of probability and statistical procedures in the area of aggregations functions over bounded lattices. For a more detailed explanation of the concepts, additional results and more examples, we refer the reader to the original paper [4].

References

[1] Martha K Nungesser, Linda A Joyce, and A David McGuire. Effects of spatial aggregation on predictions of forest climate change response. *Climate research*, 11(2):109–124, 1999.

[2] Daniel Paternain, Javier Fernández, H Bustince, Radko Mesiar, and Gleb Beliakov. Construction of image reduction operators using averaging aggregation functions. *Fuzzy Sets and Systems*, 261:87–111, 2015.

[3] Juan Baz, Irene Díaz, and Susana Montes. The choice of an appropriate stochastic order to aggregate random variables. In *International Conference on Soft Methods in Probability and Statistics*, pages 40–47. Springer, 2023.

[4] Juan Baz, Irene Díaz, and Susana Montes. Aggregation of random elements over bounded lattices. *International Journal of Approximate Reasoning*, 166:109–122, 2024.

[5] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. John Wiley & Sons, 2011.

[6] James H Bray and Scott E Maxwell. *Multivariate analysis of variance*. Number 54. Sage Publications INC, Thousand Oaks, USA, 1985.

De Funciones de Equivalencia Restringida en L^n a Medidas de Similitud entre multiconjuntos difusos

M. Ferrero-Jaurrieta, I. Rodriguez-Martinez, A. Bernardini, J. Fernández, C. Lopez-Molina, H. Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España

Z. Takáč
Instit. of Applied Informatics, Automation and Mechatronics
Slovak University of Technology
Trnava, Eslovaquia

C. Marco-Detchart
Valencian Research Institute for Artificial Intelligence (VRAIN)
Universitat Politècnica de València
Valencia, España

***Resumen*—Este artículo es un resumen del trabajo publicado en la revista *IEEE Transactions on Fuzzy Systems* [1]. En este trabajo, presentamos una contribución a la teoría de las Funciones de Equivalencia Restringida (REF), que permite comparar elementos multivaluados. Extendemos el concepto de REF de L a L^n y presentamos una nueva construcción de similitud en L^n. A partir de esta filosofía se construyen medidas de similitud entre multiconjuntos difusos y se presenta un ejemplo aplicado en en el contexto de la difusión anisotrópica de imágenes en color.**

***Palabras clave*—multiconjuntos difusos, función de equivalencia restringida, comparación multivaluada, medida de similitud**

I. Introducción

Las Funciones de Equivalencia Restringida (REF) son unas conocidas funciones para la comparación de números en el intervalo unidad. Fueron presentadas en [2] para la comparación de grados de pertenencia de conjuntos difusos, mediante la adaptación de los axiomas originales de las Funciones de Equivalencia, dados por Fodor y Roubens [3]. Tomamos $L = [0, 1]$.

Definición 1.1: Una función $R : L \times L \to L$ es una REF, si para todo $x, y, z \in L$ satisface:

(R1) $R(x, y) = 1$ si y sólo si $x = y$.
(R2) $R(x, y) = 0$ si y sólo si $\{x, y\} = \{0, 1\}$.
(R3) $R(x, y) = R(y, x)$.
(R4) Si $x \leq y \leq z$, entonces $R(x, z) \leq R(x, y)$ y $R(x, z) \leq R(y, z)$.

Las medidas de similitud han sido extensamente estudiadas en la literatura en diversas aplicaciones [4], [5]. Dada la diversidad de elementos a comparar, se han requerido distintos axiomas y propiedades para la construcción de las mismas. Teniendo en cuenta el concepto inicial de medida de similitud entre conjuntos difusos [6], se adaptaron los axiomas de la siguiente manera [7].

Definición 1.2: Sea U un conjunto no vacío y $F(U)$ el conjunto de todos los conjuntos difusos. Una medida de similitud entre dos conjuntos difusos es una función $S : F(U) \times F(U) \to L$ tal que para todos los conjuntos difusos $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \in F(U)$ satisface las siguientes propiedades:

(S1) $S(\mathfrak{A}, \mathfrak{B}) = S(\mathfrak{B}, \mathfrak{A})$.
(S2) $S(\mathfrak{A}, \mathfrak{B}) = 0$ si y sólo si $\{\mathfrak{A}(u), \mathfrak{B}(u)\} = \{0, 1\}$ para todo $u \in U$.
(S3) $S(\mathfrak{A}, \mathfrak{B}) = 1$ si y sólo si $\mathfrak{A}(u) = \mathfrak{B}(u)$, para todo $u \in U$.
(S4) Si $\mathfrak{A} \leq \mathfrak{B} \leq \mathfrak{C}$ entonces $S(\mathfrak{A}, \mathfrak{C}) \leq S(\mathfrak{A}, \mathfrak{B})$ y $S(\mathfrak{C}, \mathfrak{A}) \leq S(\mathfrak{C}, \mathfrak{B})$.

Las medidas de similitud entre conjuntos difusos pueden construirse a partir de funciones de agregación y REFs [7].

II. Funciones de Equivalencia Restringida en L^n

En este trabajo se presentan dos extensiones de las REF en vectores de L^n diferenciadas por sus axiomatizaciones. Las diferencias entre dichas propuestas radican en la consideración de qué elementos son considerados complementarios entre sí, así como en el orden de los elementos multivalorados. Es importante señalar que para las REFs en L^n, no es necesario un orden total y con un orden parcial entre vectores es suficiente para su definición. Consideramos el orden parcial entre dos vectores $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n) \in L^n$ como $X \leq_P Y$ si y sólo si $x_i \leq y_i$, para todo $i \in \{1, \ldots, n\}$.

En la primera propuesta, se consideran como elementos complementarios el $\mathbf{0} = (0, \ldots, 0)$ con el $\mathbf{1} = (1, \ldots, 1)$.

Definición 2.1: Sea n un entero positivo. Una función $R_{L_1} : L^n \times L^n \to L^n$ es denominada REF en L^n de tipo 1, si satisface, para todo $X, Y, Z \in L^n$ las siguientes propiedades:

(RL1$_1$) $R_{L_1}(X, Y) = \mathbf{1}$ si y sólo si $X = Y$.
(RL2$_1$) $R_{L_1}(X, Y) = \mathbf{0}$ si y sólo si $\{X, Y\} = \{\mathbf{0}, \mathbf{1}\}$.
(RL3$_1$) $R_{L_1}(X, Y) = R_{L_1}(Y, X)$.
(RL4$_1$) Si $X \leq_P Y \leq_P Z$, entonces $R_{L_1}(X, Z) \leq_P R_{L_1}(X, Y)$ y $R_{L_1}(X, Z) \leq_P R_{L_1}(Y, Z)$.

Proyecto PID2022-136627NB-I00 financiado por MCIN/AEI/10.13039/501100011033/FEDER, UE; Proyectos TED2021-131295B-C32, PID2021-123673OB-C31 del MCIN/AEI/10.13039/501100011033 Proyecto PROMETEO CIPROM/2021/077 Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital - Generalitat Valenciana Ayuda PAID-06-23 del Vicerrectorado de Investigación de la U. Politècnica de València (UPV)

Por otro lado, para la segunda propuesta, en lugar de considerar los elementos **0** y **1** como complementarios, consideramos cualquier par complementario de elementos crisp, que conducen al valor mínimo de similitud.

Definición 2.2: Sea n un entero positivo. Una función $R_{L_2} : L^n \times L^n \to L^n$ es denominada REF en L^n de tipo 2, si satisface, para todo $X, Y, Z \in L^n$ las siguientes propiedades:

(RL1_2) $R_{L_2}(X,Y) = \mathbf{1}$ si y sólo si $X = Y$.
(RL2_2) $R_{L_2}(X,Y) = \mathbf{0}$ si y sólo si $\{x_i, y_i\} = \{0,1\}$.
(RL3_2) $R_{L_2}(X,Y) = R_{L_2}(Y,X)$
(RL4_2) Si $\min(x_i, z_i) \leq y_i \leq \max(x_i, z_i)$, entonces $R_{L_2}(X,Z) \leq_P R_{L_2}(X,Y)$ y $R_{L_2}(X,Z) \leq_P R_{L_2}(Y,Z)$, para todo $i \in \{1, \ldots, n\}$.

III. Medidas de similitud entre multiconjuntos difusos

Los tipos de axiomas introducidos difieren en la forma de entender la complementariedad en los entornos de n-tuplas, siendo los primeros más restictivos que los segundos. Estas consideraciones permiten introducir dos tipos diferentes de medidas de similitud entre multiconjuntos difusos. Los multiconjuntos difusos considerados en este trabajo tienen una única pertenencia y una dimensionalidad fija n. Es decir, consideramos un multiconjunto sobre U como una función $\mathcal{A} : U \to L^n$.

Definición 3.1: Sea U un conjunto no vacío y $LF(U)$ el conjunto de todos los multiconjuntos difusos. Una función $S_{L_1} : LF(U) \times LF(U) \to L^n$ es denominada una medida de similitud de tipo 1 entre dos multiconjuntos difusos si satisface los siguientes axiomas para todo $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \in LF(U)$:

(SL1_1) $S_{L_1}(\mathcal{A}, \mathcal{B}) = \mathbf{1}$ si y sólo si $\mathcal{A} = \mathcal{B}$.
(SL2_1) $S_{L_1}(\mathcal{A}, \mathcal{B}) = \mathbf{0}$ si y sólo si $\mathcal{A}$ y $\mathcal{B}$ son $\{\mathcal{A}(u), \mathcal{B}(u)\} = \{\mathbf{0}, \mathbf{1}\}$, para todo $u \in U$.
(SL3_1) $S_{L_1}(\mathcal{A}, \mathcal{B}) = S_{L_1}(\mathcal{B}, \mathcal{A})$.
(SL4_1) Si $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{C} \subseteq \mathcal{D}$ entonces $S_{L_1}(\mathcal{A}, \mathcal{D}) \leq_P S_{L_1}(\mathcal{B}, \mathcal{C})$.

Definición 3.2: Sea U un conjunto no vacío y $LF(U)$ el conjunto de todos los multiconjuntos difusos. Una función $S_{L_1} : LF(U) \times LF(U) \to L^n$ es denominada una medida de similitud de tipo 2 entre dos multiconjuntos difusos si satisface los siguientes axiomas para todo $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \in LF(U)$:

(SL1_2) $S_{L_2}(\mathcal{A}, \mathcal{B}) = \mathbf{1}$ si y sólo si $\mathcal{A} = \mathcal{B}$.
(SL2_2) $S_{L_2}(\mathcal{A}, \mathcal{B}) = \mathbf{0}$ si y sólo si $\mathcal{A}(u) \in \{0,1\}^n$ y $\mathcal{B}(u) = \mathbf{1} - \mathcal{A}(u)$, para todo $u \in U$
(SL3_2) $S_{L_2}(\mathcal{A}, \mathcal{B}) = S_{L_2}(\mathcal{B}, \mathcal{A})$.
(SL4_2) Si para todo $u \in U$, $\min(\mathcal{A}(u)_i, \mathcal{D}(u)_i) \leq \mathcal{B}(u)_i$ y $\max(\mathcal{A}(u)_i, \mathcal{D}(u)_i) \geq \mathcal{C}(u)_i$, entonces $S_{L_2}(\mathcal{A}, \mathcal{D}) \leq_P S_{L_2}(\mathcal{B}, \mathcal{C})$ (donde $\mathcal{A}(u)_i$ es el i-ésimo elemento de una n-tupla $\mathcal{A}(u)$).

Se ha presentado un método de construcción para cada medida de similitud, utilizando un tipo de REF en L^n para cada una. Para ello, consideramos un entero positivo m, $U = \{u_1, \ldots, u_m\}$ y una función $R_L : L^n \times L^n \to L^n$ REFs en L^n de tipo 1 ó 2. Sea $G_L : (L^n)^m \to L^n$ una funcion de agregación n-dimensional m-aria, tal que, para todo $X_1, \ldots, X_m \in L^n$, $G_L(X_1, \ldots, X_m) = \mathbf{0}$ implica $X_1 = \ldots = X_m = \mathbf{0}$ y $G_L(X_1, \ldots, X_m) = \mathbf{1}$ implica $X_1 = \ldots = X_m = \mathbf{1}$. Entonces la función $S_L : LF(U) \times LF(U) \to L^n$ dada por $S_{L_1}(\mathcal{A}, \mathcal{B}) = G_L\left(R_L(\mathcal{A}(u_1), \mathcal{B}(u_1)), \ldots, R_L(\mathcal{A}(u_m), \mathcal{B}(u_m))\right)$ es una medida de similitud entre multiconjuntos difusos, de tipo 1 si $R_L = R_{L_1}$, o de tipo 2 si $R_L = R_{L_2}$.

IV. Aplicación

Como caso de estudio hemos seleccionado el modelo de difusión anisotrópica Perona-Malik (PMAD), un algoritmo de procesamiento de imagen para el suavizado sensible al contenido. Éste método se basa en la comparación de datos semi-local para estimar la necesidad de suavizado alrededor de cada píxel de la imagen. Para medir su eficacia, utilizamos una medida de homogeneidad para obtener lo similar que son los colores en la vecindad de un píxel, obteniendo una imagen de homogeneidad que refleja dónde las regiones son homogéneas y dónde hay cambios relevantes.

Comparando nuestro enfoque con otros métodos con Mean-Shift y el Filtrado Bilateral vemos que los resultados son similares a los que no utilizan un automorfismo específico. Por un lado, en términos de suavizado de objetos estos métodos son equivalentes a nuestra propuesta, que tiene un valor de homogeneidad ligeramente mejor. Por otro lado, el suavizado del fondo es peor que nuestro método, especialmente cuando se utiliza el Filtrado Bilateral.

V. Conclusión

Por último, podemos ver que la extensión de las REF al dominio L^n es una medida de comparación adecuada entre elementos multivaluados, dando pie a la generación de otras medidas de comparación entre estructuras multivaluadas, como los multiconjuntos difusos. Podemos afirmar que diferentes parametrizaciones de estas funciones conducen a resultados de difusión anisotrópica distintos en la aplicación específica propuesta, verificando la sensibilidad de estas funciones a la configuración de los parámetros y su adecuación a diferentes escenarios.

References

[1] M. Ferrero-Jaurrieta, Z. Takáč, I. Rodriguez-Martinez, C. Marco-Detchart, A. Bernardini, J. Fernández, C. Lopez-Molina, and H. Bustince, "From restricted equivalence functions on l^n to similarity measures between fuzzy multisets," *IEEE Transactions on Fuzzy Systems*, vol. 31, no. 8, pp. 2709–2721, 2023.
[2] H. Bustince, E. Barrenechea, and M. Pagola, "Restricted equivalence functions," *Fuzzy Sets and Systems*, vol. 157, no. 17, pp. 2333–2346, 2006.
[3] J. Fodor and M. Roubens, *Fuzzy preference modelling and multicriteria decision support*. Kluwer Academic Publishers, 1994.
[4] D. Van Der Weken, M. Nachtegael, and E. Kerre, "Using similarity measures and homogeneity for the comparison of images," *Image and Vision Computing*, vol. 22, pp. 695–702, 08 2004.
[5] T. Chaira and A. K. Ray, "Fuzzy measures for color image retrieval," *Fuzzy Sets Syst.*, vol. 150, no. 3, pp. 545–560, 2005.
[6] L. Xuecheng, "Entropy, distance measure and similarity measure of fuzzy sets and their relations," *Fuzzy Sets and Systems*, vol. 52, no. 3, pp. 305–318, 1992.
[7] H. Bustince, E. Barrenechea, and M. Pagola, "Image thresholding using restricted equivalence functions and maximizing the measures of similarity," *Fuzzy Sets and Systems*, vol. 158, no. 5, pp. 496–516, 2007.

Key work paper "A Methodology to Quickly Perform Opinion Mining and Build Supervised Datasets Using Social Networks Mechanics"

Published in IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, number 35 issue 9, pages 9797-9808, SEPT 1 2023

Manuel Francisco
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain

Juan Luis Castro
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain
castro@decsai.ugr.es

Abstract—**Social Networking Sites (SNS) offer a full set of possibilities to perform opinion studies such as polling or market analysis. Normally, artificial intelligence techniques are applied, and they often require supervised datasets. The process of building them is complex, time-consuming and expensive. In this paper, it is proposed to assist the labelling task by taking advantage of social network mechanics. In order to do that, it is introduced the co-retweet relation to build a graph that allows to propagate user labels to their similarity neighbourhood. Therefore, it is possible to build supervised datasets with significant less human effort and with higher accuracy than other weak-supervision techniques. The proposal was tested with 3 datasets labelled by an expert committee, and results showed that it outperforms other weak-supervision techniques. This methodology may be adapted to other social networks and topics, it is relevant for applications like informed decision-making (e.g. content moderation), specially when interpretability is required.**

Keywords— Terms—human-in-the-loop labelling, opinion mining, social network analysis, user profiling, supervised learning.

I. Introduction

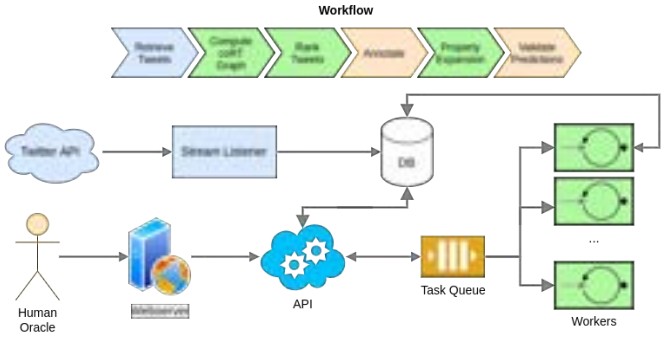
In this article it is proposed a methodology to reduce the effort required to produce a supervised dataset. It is made by using semantic network representations and label propagation mechanisms. This methodology can be divided in different steps that start with data retrieval. Meta-data regarding user interaction is used to compute a similarity graph, and most relevant documents are labelled using oracles. These labels are aggregated and propagated through the graph to obtain information regarding unknown users. Since not all labels are accurate, oracles are asked to validate new information and check for potential conflicts. This makes our proposal a human-in-the-loop approach to produce weak-supervised dataset with improved quality.

The main contributions of this article are (1) a methodology to reduce the effort when producing labelled dataset from SNS data, (2) a novel similarity-based network representation (the co-retweet graph), and (3) a label propagation mechanism that takes into account the relevance and coherence of the network. Both the network representation and the propagation mechanism are two possible implementations of our methodology, that we tested using (4) a proof-of-concept platform that follows aforementioned steps to build weak-supervised datasets from Twitter. Results show that this methodology is able to reduce labour costs by 75 percent and it outperforms other weak-supervision techniques.

II. Proposed Methodology

We propose the use of a system that would allow us to infer properties of unknown users from other previously annotated documents. The basic workflow would be:

1) Rank tweets by utility.
2) Ask an oracle to annotate the top n tweets.
3) Expand properties to other user profiles using a deep relation.
4) Rank automatically-generated user annotations by utility.
5) Ask an oracle to validate the top m autoannotations.
6) Repeat from step 2 until necessary.



Grant PID2021-125788OB-Ioo funded by MICCIU/AEI/10.1309/501100011033 and by ERDF/EU.

Fig. 1. Workflow of the system

And the keys of the system are:

A. Similarity Graph

It calculates the retweets that any two users have in common. As we stated above, the retweet mechanic implies that the user is interested in the topic and also that they agree with the opinion of the original author. Therefore, if users A and B have retweeted something from C, the retweet graph would have three directed links, from A to C and from B to C; with our proposal, both links would be undirected and a new edge would arise between A and B. Note that original tweets are considered copies (retweets) of themselves. Consequently, each node in the co-retweet graph would stand for a Twitter user and edges connecting

B. Property expansion

Given a similarity graph (in our case, co-retweet graph), it is possible to infer property values for new users in the neighbourhood of known users through a process of weighted

III. Development and Experiments

In order to check the performance of our proposal against current weak-supervision techniques. We applied several classic machine learning algorithms (Support Vector Machine (SVM), Random Forest (RF), AdaBoost (AB) and Multinomial Naive Bayes (MNB). We applied our methodology to datasets of tweets regarding Spanish National Elections (spanish), Madrid Regional Elections (madrid) and USA Presidential Elections (usa).

IV. Results and Discussion

Results shows that the number of automatic annotations grows quickly with the first annotations due to the ranking strategy, and it stabilises after the annotation process reach non-influential users. For example, for the Spanish National Elections dataset, the top 25% has more than 191 connected users, hence the grow. The ratio between number of accepted annotations versus total annotations has a mean value of 0.89, with an almost negligible standard deviation of 0.0045. It keeps steady regardless of the number of manual annotations. This points out that the chosen user representation (co-retweet graph) behaves coherently. Our proposal significantly improves the results of other weak-supervision techniques when the number of annotations is low. When annotating tweets, it is common to assume that the cost of evaluating a tweet is uniform [42], since documents have 240 characters as much. The number of manual annotations is called effort. Table 6 shows the effort required to reach at least .75 in f1-score. Note that we stopped several experiments before they reached the threshold. Our proposal is the method that requires less human annotations therefore it is ideal to reduce labelling costs.

V. Conclusion

SNS are used frequently to study the opinion of customers, citizens, voters and almost any other role a human can assume while using Social Media. However, these analysis present limitations that should be considered, such as users age range and digital literacy. Normally, artificial intelligence techniques can be used to perform these analyses, which often require supervised datasets. Labelling datasets is an expensive task that needs a lot of resources, regardless that it is conducted in-house or outsourced to companies or freelancers.

Most SNS have similar mechanics, such as liking and befriending, that may offer more information than the contentitself. We introduced the co-retweet, that is built upon user interactions and represents the network as a similarity graph. Therefore, it constitutes a way to infer knowledge of unknown users from others in their neighbourhood. In order to do this, we presented a methodology that iteratively propagates labels through a similarity graph, generating predictions that can be reviewed automatically, in most cases, or manually. Our experiments show that our proposal outperforms other weak-supervision techniques when the effort is low, and it behaves at least as well as the number of manual annotations grows.

These results are relevant in the field of opinion mining for applications like market analysis, recommendation system and trend predictions, and it is particularly useful since it significantly improves the required time to build a supervised dataset without sacrificing interpretability or quality.

Acknowledgment

Grant PID2021-125788OB-Ioo unded by MICIU/AEI/ 10.13039/501100011033 and by “ERDF/EU”.

Reference

[1] M. Francisco, and J..L. Castro, “A Methodology to Quickly Perform Opinion Mining and Build Supervised Datasets Using Social Networks Mechanics”, IEEE Trans. on Knowledge and Data Engineering 35 9, 9797-9808, 2023.

Grant PID2021-125788OB-Ioo funded by MICCIU/AEI/10.1309/501100011033 and by ERDF/EU.

Sobre el problema de ordenación de Z-números basados en números borrosos discretos

M. González-Hidalgo *† §S. Massanet*†§, A. Mir*†§, A. Mir-Fuentes † ‡, J. V. Riera*†§, L. De Miguel¶
* *Grupo de investigación en Soft Computing, Procesamiento de Imágenes y Agregación (SCOPIA)*
‡ *Grupo de investigación Modelos para el Tratamiento de la Información Borrosa (MOTIBO)*
Departamento de Ciencias Matemáticas e Informática
Universitat de les Illes Balears, 07122 Palma, España
†*Instituto de Investigación Sanitaria de las Islas Baleares (IdISBa), 07010 Palma, España*
§*Instituto de Investigación en Inteligencia Artificial (IAIB), 07122 Palma, España*
¶ *Departamento de Estadística, Informática y Matemáticas,*
Universidad Pública de Navarra, Campus Arrosadia, 31006 Pamplona, Spain
E-mails: {manuel.gonzalez, s.massanet, arnau.mir, a.mir, jvicente.riera}@uib.es, laura.demiguel@unavarra.es

***Resumen*—Este trabajo es un resumen del trabajo publicado en las actas del congreso de la 13th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2023) [3] y sometido a la revista [4] para su presentación en la Multiconferencia CAEPIA'24 (ESTYLF 2024) KeyWorks.**

***Index Terms*—Z-números de Zadeh, Número borroso discreto, Orden Total, Orden admisible, Ranking de Z-números.**

I. Resumen

Los Z-números de Zadeh se han posicionado como un interesante modelo lingüístico computacional que se ha utilizado en múltiples aplicaciones y, en particular, en problemas de toma de decisiones. Formalmente [7], una pareja ordenada de números borrosos (A, B) se entenderá como un Z-número. Cuando un Z-número es asociado a una variable de incertidumbre real valorada X, el triplete ordenado (X, A, B) es entendido como una Z-valoración, donde la primera componente A es interpretada como una restricción sobre los valores que puede tomar X y la segunda componente B nos indica el grado de certeza (seguridad, confianza, posibilidad, probabilidad...) que se tiene sobre el valor descrito por A. Este modelo permite modelizar oraciones del tipo "Usualmente, el tiempo de viaje entre Madrid y Barcelona es aproximadamente tres horas", siendo la Z-valoración asociada a esta frase Z_1=(X= "tiempo de viaje entre Madrid y Barcelona", A="aproximadamente tres horas", B="usualmente").

Si bien este modelo lingüístico permite modelar muchas oraciones imprecisas del lenguaje natural, tiene grandes desventajas a la hora de realizar operaciones con este tipo de estructuras. Este problema ya fue mencionado por el propio L. Zadeh en [7] "Problems involving computation with Z-numbers are easy to state but far from easy to solve". Con el objetivo de reducir el coste computacional cuando se efectúan operaciones entre Z-valoraciones, algunos autores han propuesto una nueva perspectiva de los Z-números de Zadeh [1]. En esta nueva formulación, los Z-números se representan por pares de números borrosos discretos (A, B) donde la segunda componente describe el grado de certeza (certidumbre, confianza, fiabilidad, ...) de la primera componente y además no se interpreta desde un punto de vista estrictamente probabilístico, sino como una evaluación basada en un número borroso discreto [6] facilitando los cálculos entre Z-valoraciones utilizando operadores de agregación sobre el conjunto de números borrosos discretos [2]. Esta nueva perspectiva resuelve en gran medida el problema de los elevados costes computacionales asociados a la definición clásica de Z-números, manteniendo la esencia lingüística establecida en la idea original propuesta por Zadeh. Recordemos a continuación la notación básica que se empleará en este trabajo así una definición esencial para el mismo.

Denotemos por $\mathcal{A}_1^{L_n}$, donde $L_n = \{0, \cdots, n\}$ es una cadena finita, el conjunto de números borrosos discretos cuyo soporte ($\{x \in L_n$ tales que $A(x) > 0\}$) es un intervalo cerrado de la cadena finita L_n y por $\mathcal{A}_1^{L_n \times Y_m}$ el subconjunto de números borrosos discretos $A \in \mathcal{A}_1^{L_n}$ tales que su función de pertenencia $A(x) \in Y_m = \{y_1 = 0, y_2, \cdots, y_m = 1\}$. Recordemos que el núcleo de cualquier elemento $A \in \mathcal{A}_1^{L_n}$ es el conjunto Núcleo$(A) = \{x \in L_n$ tales que $A(x) = 1\}$.

Definición 1: Sean L_n y L_k dos cadenas finitas. Una pareja ordenada de números borrosos discretos (A, B) con $A \in \mathcal{A}_1^{L_n}$, $B \in \mathcal{A}_1^{L_k}$ la llamaremos (L_n, L_k)-Z-número borroso discreto. Si asociamos un (L_n, L_k)-Z-número borroso discreto a una variable de incertidumbre X, la primera componente A desempeñará la función de restricción borrosa sobre X, mientras que el número borroso discreto B describirá una estimación imprecisa de la confianza sobre A. El triplete ordenado (X, A, B) será llamado Z-valoración y se escribirá X es (A, B). El conjunto de Z-números discretos será denotado por $\mathcal{Z}(L_n, L_k)$ y un elemento arbitrario de este conjunto por $Z(A, B)$. En particular, denotaremos por $\mathcal{Z}(L_n, L_k \times Y_m)$ al conjunto de (L_n, L_k)-Z-números borrosos discretos cuya segunda componente sea finita valorada, esto es, si $Z(A, B) \in \mathcal{Z}(L_n, L_k \times Y_m)$, se tiene que $A \in \mathcal{A}_1^{L_n}$ y $B \in \mathcal{A}_1^{L_k \times Y_m}$.

El principal objetivo de los trabajos [3], [4] es el estudio de ordenes totales en el conjunto $\mathcal{Z}(L_n, L_k \times Y_m)$ que sean con-

sistentes con el significado lingüístico de las dos componentes del Z-número discreto de acuerdo a la propuesta original establecida por Zadeh [7], que es una de las principales carencias que tienen los órdenes sobre Z-números conocidos en la literatura [5].

Para ello, un primer paso fue estudiar el cardinal del conjunto $\mathcal{A}_1^{L_k \times Y_m}$, siendo su valor $\binom{k+2m-2}{2m-2}$. Este cardinal nos permite definir la función de ordenación relativa que será básica para definir un orden total en el conjunto $\mathcal{Z}(L_n, L_k \times Y_m)$.

Definición 2: Sea $\preceq$ un orden total en $\mathcal{A}_1^{L_k \times Y_m}$ y $C_A = \{X \in \mathcal{A}_1^{L_k \times Y_m} \mid X \preceq A\}$. Se define la función ω que determina el rango relativo de un número borroso discreto A tal como sigue:

$$\begin{array}{rccl} \omega: & \mathcal{A}_1^{L_k \times Y_m} & \longrightarrow & (0,1] \\ & A & \longrightarrow & \omega(A) = \frac{|C_A|}{\left|\mathcal{A}_1^{L_k \times Y_m}\right|}, \end{array}$$

siendo $|C_A|$ el cardinal de C_A.

Una propiedad esencial de esta función ω es que es compatible con el orden total $\preceq$ fijado en $\mathcal{A}_1^{L_k \times Y_m}$, esto es, $A \preceq B$ si y solo si $\omega(A) \leq \omega(B)$. Notar que para calcular el rango relativo de un número borroso discreto, es necesario calcular todos los números borrosos discretos de $\mathcal{A}_1^{L_k \times Y_m}$. Para ello, se han establecido tres algoritmos que permiten generarlos y ordenarlos. Haciendo uso de la función ω, se construye la siguiente función que nos permitirá establecer finalmente una relación de orden total en el conjunto $\mathcal{Z}(L_n, L_k \times Y_m)$.

Definición 3: Sea $Z(A,B) \in \mathcal{Z}(L_n, L_k \times Y_m)$. Se define A_B, la función transformada de A por B como:

$$A_B(x) = \begin{cases} A(x) + f(\omega(B))(1 - A(x)), & \text{si } x \in \text{soporte}(A), \\ 0 & \text{en otro caso,} \end{cases}$$

donde $f : (0,1] \longrightarrow [0,1)$ es una función estrictamente decreciente tal que $f(1) = 0$ y ω es el rango relativo en $\mathcal{Z}(L_n, L_k \times Y_m)$.

Es importante destacar que A_B no sólo mantiene el mismo núcleo y soporte del número borroso discreto inicial A. Como puede verse en su definición, el valor positivo $f(\omega(B))(1 - A(x))$ se añade a la pertenencia de cada elemento del soporte de A. Este valor es mayor cuanto menor es el rango relativo de B en $L_k \times Y_m$, es decir, cuanto menor es la fiabilidad/confianza de A. De este modo, esta transformación modifica la incertidumbre de A (aumentando los valores de pertenencia fuera del núcleo) en función de la fiabilidad/confianza de A proporcionada por B. Ahora, estamos en condiciones de definir en $\mathcal{Z}(L_n, L_k \times Y_m)$ un orden total. Para ello, consideremos la siguiente relación binaria:

Definición 4: Consideremos dos ordenes totales, $\preceq_{\mathcal{A}}$, $\preceq_{\mathcal{B}}$, en $\mathcal{A}_1^{L_n}$ y en $\mathcal{A}_1^{L_k \times Y_m}$ respectivamente. Definimos la relación binaria en $\mathcal{Z}(L_n, L_k \times Y_m)$ asociada a $\preceq_{\mathcal{A}}$ and $\preceq_{\mathcal{B}}$ de la siguiente manera:

Dados $Z(A,B), Z(A',B') \in \mathcal{Z}(L_n, L_k \times Y_m)$,

$$Z(A,B) \preceq Z(A',B') \Leftrightarrow ((A_B \prec_{\mathcal{A}} A'_{B'}) \text{ o } (A_B = A'_{B'} \text{ y } B \preceq_{\mathcal{B}} B')).$$

La idea de esta relación es la siguiente:

Si se consideran las Z-valoraciones $Z(A,B), Z(A',B') \in \mathcal{Z}(L_n, L_k \times Y_m)$, primero se comparan A_B y $A'_{B'}$ ya que estos dos números borrosos discretos representan, en cada caso, la información dada por la restricción borrosa sobre la variable que se quiere modelar transformada por medio de la credibilidad de esta restricción borrosa. En el caso en que estos números sean iguales, se considerará más relevante el que tenga mayor credibilidad. Finalmente, se demuestra que dicha relación es un orden total.

Teorema 5: La relación binaria establecida anteriormente es un orden total en el conjunto $\mathcal{A}_1^{L_k \times Y_m}$.

En conclusión, en este trabajo se ha presentado un método de construcción de órdenes totales en el conjunto $\mathcal{Z}(L_n, L_k \times Y_m)$ que mantienen la coherencia lingüística de las frases que se modelan de acuerdo al modelo lingüístico computacional inicialmente propuesto por L. Zadeh en [7] y que, además puede ser usado con facilidad en problemas de toma de decisiones. Como futuro trabajo queremos definir funciones de agregación para Z-números discretos basados en el orden total que hemos establecido. Además, pretendemos ampliar este orden para los llamados Z-números mixtos $Z(A,B)$, en los que A no es necesariamente un número borroso discreto sino que puede ser cualquier número borroso.

Agradecimientos

M. González-Hidalgo, S. Massanet, A. Mir y J.V. Riera están parcialmente subvencionados por el proyecto R+D+i PID2020-113870GB-I00-"Desarrollo de herramientas de Soft Computing para la Ayuda al Diagnóstico Clínico y a la Gestión de Emergencias (HESOCODICE)", financiado por MCIN/AEI/10.13039/501100011033/.

L. De Miguel está parcialmente subvencionada por el proyecto R+D+i PID2022-136627NB-I00-"Técnicas de reducción del número de datos numéricos y datos multiatributo heterogéneos para la mejora de su fusión en problemas de inteligencia artificial" financiado por MCIN/AEI/10.13039/501100011033/.

Referencias

[1] Massanet, S., Riera, J.V., Torrens, J.: A new vision of Zadeh's Z-numbers. In: Carvalho, J.P., Lesot, M.J., Kaymak, U., Vieira, S., Bouchon-Meunier, B., Yager, R.R. (eds.) Information Processing and Management of Uncertainty in Knowledge-Based Systems. pp. 581–592. Springer International Publishing, Cham (2016).

[2] Riera, J.V., Torrens, J.: Aggregation of subjective evaluations based on discrete fuzzy numbers. Fuzzy Sets and Systems 191, 21–40 (2012)

[3] Mir-Fuentes, A., De Miguel, L., Massanet, S., Mir, A. and Riera, J. V.: On the problem for ordering Z-numbers based on discrete fuzzy numbers . Book of Abstracts, Palma, Spain 13th Conference of the European Society for Fuzzy Logic 12th International Summer School on Aggregation Operators (2023).

[4] Mir-Fuentes, A., De Miguel, L., Massanet, S., Mir, A. and Riera, J.V.: On a total order on the set of Z-numbers based on discrete fuzzy numbers sometido a la revista Computational and Applied Mathematics (2024).

[5] Cheng, R., Zhang, J., Kang, B.: Ranking of Z-numbers based on the developed golden rule representative value. IEEE Transactions on Fuzzy Systems 30(12), 5196–5210 (2022).

[6] Massanet S., Riera J. V., Torrens J., Herrera-Viedma E., "A new linguistic computational model based on discrete fuzzy numbers for computing with words", Inf. Sci. 258, pp. 277–290, 2014.

[7] Zadeh, L.A.: A note on Z-numbers. Information Sciences **181**(14), 2923–2932 (2011).

Generalizando el pooling máximo por funciones (a, b)-grouping en Redes Neuronales Convolucionales

Iosu Rodriguez-Martinez
Dpto. Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España

Tiago da Cruz Asmus
Inst. Matemática, Estadística e Física
Universidade Federal do Rio Grande
Rio Grande, Brasil

Graçaliz Pereira Dimuro
Centro de ciencia Computacionais
Universidade Federal do Rio Grande
Rio Grande, Brasil

Francisco Herrera
DASCI
Universidad de Granada
Granada, España

Zdenko Takáč
Inst. Appl. Informatics, Automation, Mechatronics
Slovak University of Technology
Trnava, Eslovaquia

Humberto Bustince
Dpto. Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España

***Abstract*—Este artículo es un resumen del trabajo publicado en la revista Information Fusion [1]. En este artículo explorábamos el reemplazo del operador de pooling máximo comunmente empleado en redes neuronales convolucionales (CNNs) por funciones (a, b)-grouping. Estas funciones extienden el concepto de función de grouping clásica [2] a un intervalo cerrado $[a, b]$, siguiendo la filosofía de [3]. En el contexto del operador de pooling, estas nuevas funciones ayudan a la optimización de los modelos suavizando los gradientes en el proceso de retropropagación y obteniendo resultados competitivos con métodos más complejos.**

***Index Terms*—Redes neuronales convolucionales, funciones grouping, funciones pooling, clasificación de imagen**

I. Introducción

En el contexto de un problema de clasificación de imagen, las redes neuronales convolucionales CNN [4] operan tradicionalmente siguiendo un proceso alterno consistente en (1) la extracción de características visuales mediante la convolución de filtros de imagen y (2) la reducción de la información extraída mediante estos filtros en las llamadas capas de "pooling".

Este segundo proceso lleva a cabo un proceso de reducción de imagen sustituyendo ventanas disjuntas de la imagen por una agregación de sus valores como la media aritmética o, más comunmente, el máximo. No obstante, en el contexto de un problema de optimización convexa, el máximo tiene el problema de hacer cero el gradiente de todos los valores agregados a excepción de uno, lo que puede afectar al proceso de optimización.

Pese a ello, priorizar las activaciones altas extraídas por la CNN es un comportamiento deseable, especialmente al emplear la funciones rectificadas como función de activación. Esto nos lleva a explorar otras funciones pertenecientes a la familia de las funciones grouping [2], a la que el propio máximo pertenece.

Definición 1.1: Una función de grouping n-dimensional $G : [0, 1]^n \to [0, 1]$ es una función de agregación que cumple que para todo $\mathbf{x} \in [0, 1]^n$:

- (G1) G es simétrica;
- (G2) $G(\mathbf{x}) = 0$ si y solo si $x_i = 0$ para todo $i \in \{1, \ldots, n\}$;
- (G3) $G(\mathbf{x}) = 1$ si y solo si existe $i \in \{1, \ldots, n\}$ tal que $x_i = 1$;
- (G4) G es creciente;
- (G5) G es continua.

II. Funciones de (a, b)-grouping

Las entradas y representaciones internas de una red neuronal son valores reales no acotados al intervalo $[0, 1]$. En la práctica es posible asumir que los valores generados tras cada capa se encuentran acotados a un intervalo $[a, b]$ con $a, b \in \mathbb{R}$. En [3] presentamos una estrategia que permite extender el dominio de cualquier función de agregación a un intervalo $[a, b]$, poniendo especial énfasis en las propiedades fundamentales de cada familia de funciones.

Definición 2.1: Una función $G^{a,b} : [a, b]^n \to [a, b]$ es una función de grouping (a, b) si, para todo $\mathbf{x} \in [a, b]^n$ se cumple:

- (GAB1) $G^{a,b}$ es simétrica;
- (GAB2) $G^{a,b}(\vec{x}) = a$ si y solo si $x_i = a$ para todo $i \in \{1, \ldots, n\}$;
- (GAB3) $G^{a,b}(\vec{x}) = b$ si y solo si existe $i \in \{1, \ldots, n\}$ tal que $x_i = b$;
- (GAB4) $G^{a,b}$ es creciente;
- (GAB5) $G^{a,b}$ es continua.

La anterior definición hace sencilla la construcción de funciones de (a, b)-grouping a partir de funciones de grouping

PID2022-136627NB-I00 financiado por MCIN/AEI/10.13039/501100011033/FEDER, UE. Proyecto CNPq (301618/2019-4). Andalusian Excellence project P18-FR-4961. Proyecto VEGA 1/0267/21. Gobierno de Navarra - Departamento de Universidad, Innovación y Transformación Digital. Tracasa Instrumental.

mediante el uso de un isomorfismo $\phi : [a, b] \to [0, 1]$ de la siguiente forma:

Teorema 2.1: Sean una función de fusión $G : [0, 1]^n \to [0, 1]$, una función creciente y biyecctiva $\phi : [a, b] \to [0, 1]$ y una función de fusión $(a, b)G^{a,b} : [a, b]^n \to [a, b]$ dada, para todo $\mathbf{x} = (x_1, \ldots, x_n) \in [a, b]$, por

$$G^{a,b}(\mathbf{x}) = \phi^{-1}\left(G\left(\phi(x_1), \ldots, \phi(x_n)\right)\right), \qquad (1)$$

Entonces, $G^{a,b}$ es una función de (a, b)-grouping n-dimensional si y solo si G es una función de grouping n-dimensional.

Al igual que las funciones de grouping n-dimensionales, las funciones de (a, b)-grouping son autocerradas con respecto a composición. Además, la combinación convexa de distintas funciones de (a, b)-grouping, da como resultado otra función (a, b)-grouping. Estos métodos de construcción alternativos nos dan flexibilidad a la hora de construir nuevas funciones de grouping (a, b) en entornos aplicados.

III. Resultados experimentales

Con el objetivo de validar nuestra propuesta, llevamos a cabo una serie de validaciones experimentales usando modelos de CNN y dataset estándar. Debido a limitaciones de espacio, solo incluimos en este trabajo un resumen superficial de las pruebas realizadas. Se pueden consultar los detalles de la validación experimental en [1].

Para este estudio sustituimos las capas de pooling de los modelos LeNet-5 [5] y VGG16 [6] por una serie de funciones (a, b)-groupings. De modo similar, también sustituimos las convoluciones encargadas del proceso de reducción de imagen en un modelo ResNet-56 [6] por las mismas funciones. Además de comparar nuestra propuesta con el pooling máximo y promedio, también la comparamos con alternativas como "mixed" pooling y "gated" pooling [7] o el pooling basado en el mecanismo de atención [8].

Los resultados obtenidos con los mejores (a, b)-groupings para el conjunto CIFAR-10 [9] se presentan en la Tabla I. En general, los resultados obtenidos con (a, b)-groupings son comparables con los obtenidos con métodos que dependen de parámetros adicionales como los pooling "gated" o el pooling basado en atención, pese a no requerir ningún parámetro extra.

Adicionalmente, motivado por los buenos resultados obtenidos por el "mixed" pooling, completamos nuestra propuesta probando a combinar distintos (a, b)-groupings con la media aritmética mediante una combinación convexa. Esta modificación facilita la utilización de algunas funciones de (a, b)-grouping que generan gradientes no estables en la retropropagación por sí mismos. En general, el rendimiento del método mejora, sin suponer un sobrecoste en tiempo de ejecución al algoritmo.

TABLE I
Comparación entre el mejor (a, b)-grouping y el resto de métodos

	CIFAR-10		
	LeNet-5	VGG16	ResNet
Avg	0.825 ± 0.003	0.915 ± 0.001	0.919 ± 0.004
Max	0.837 ± 0.003	0.911 ± 0.003	0.919 ± 0.003
Best (a, b)-grouping	0.839 ± 0.003	$\mathbf{0.916 \pm 0.002}$	$\mathbf{0.923 \pm 0.001}$
Mixed pooling	$\mathbf{0.842 \pm 0.001}$	$\mathbf{0.916 \pm 0.002}$	0.922 ± 0.002
Gated pooling	$\mathbf{0.842 \pm 0.003}$	0.913 ± 0.003	0.922 ± 0.002
Attention pooling	0.836 ± 0.002	0.884 ± 0.008	$\mathbf{0.923 \pm 0.003}$

IV. Conclusión

Este artículo demuestra que existen alternativas sencillas que pueden mejorar el funcionamiento de redes neuronales complejas cuando se estudian las propiedades necesarias. En concreto, las funciones de (a, b)-grouping representan una buena alternativa al operador de pooling máximo que permiten incorporar información de más de uno de los elementos de entrada, aún priorizando los más elevados. Además, también comprobamos que combinar algunas de estas expresiones con la media aritmética suaviza los gradientes generados durante el proceso de retropropagación, facilitando el aprendizaje del modelo.

References

[1] I. Rodriguez-Martinez, T. da Cruz Asmus, G. P. Dimuro, F. Herrera, Z. Takáč, and H. Bustince, "Generalizing max pooling via (a, b)-grouping functions for convolutional neural networks," *Information Fusion*, vol. 99, p. 101893, 2023.

[2] H. Bustince, M. Pagola, R. Mesiar, E. Hullermeier, and F. Herrera, "Grouping, overlap, and generalized bientropic functions for fuzzy modeling of pairwise comparisons," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 3, pp. 405–415, 2011.

[3] T. d. C. Asmus, G. P. Dimuro, B. Bedregal, J. A. Sanz, J. Fernandez, I. Rodriguez-Martinez, R. Mesiar, and H. Bustince, "A construtive framework to define fusion functions with floating domains in arbitrary closed real intervals," *Information Sciences*, vol. 610, pp. 800–829, 2022.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*, 2015.

[7] C.-Y. Lee, P. Gallagher, and Z. Tu, "Generalizing pooling functions in cnns: Mixed, gated, and tree," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 863–875, 2018.

[8] Q. Bi, K. Qin, H. Zhang, J. Xie, Z. Li, and K. Xu, "Apdc-net: Attention pooling-based convolutional network for aerial scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 9, pp. 1603–1607, 2019.

[9] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," University of Toronto, Toronto, Ontario, Tech. Rep. 0, 2009.

Applying a Fuzzy-logic Based System for Pattern Mining to Diagnostics and Co-morbidity in a Distributed Medical Environment

1st Carlos Fernandez-Basso
Computer Science and Artificial Intelligence
University of Granada
Granada, España
cjferba@decsai.ugr.es

2nd Karel Gutiérrez-Batista
Computer Science and Artificial Intelligence
University of Granada
Granada, España
karel@decsai.ugr.es

3rd Roberto Morcillo-Jiménez
Computer Science and Artificial Intelligence
University of Granada
Granada, España
robermorji@ugr.es

4th Maria Dolores Ruiz
Computer Science and Artificial Intelligence
University of Granada
Granada, España
mdruiz@decsai.ugr.es

5th Maria J. Martin-Bautista
Computer Science and Artificial Intelligence
University of Granada
Granada, España
mbautis@decsai.ugr.es

***Abstract*—This paper discusses utilizing data mining techniques and fuzzy logic for extracting hidden knowledge from medical records. Existing challenges such as lack of interpretability and data understanding due to overemphasis on classification, prediction and knowledge extraction are recognized. The proposed system addresses these challenges by a) preprocessing the database, b) fuzzifying the data to account for natural uncertainty, c) discovering patterns and relationships among different data features and d) visualizing results for easier analysis and interpretation. Increased comprehension and interpretability of medical data for end-users has been achieved, with real-life case studies included to demonstrate the system's effectiveness.**

***Index Terms*—association rules fuzzy logic data mining medical records**

I. Introduction

Increasing interest is being shown by corporations and researchers in the development of robust systems for health data analysis due to the expanding need in this field. The crucial challenges lie in extracting concealed knowledge from health data, conducting diagnostic and co-morbidity analyses, optimizing healthcare services, and enhancing the interpretability of data mining systems [1].

High interpretability allows users to comprehend the rationale behind decisions or consistently predict model outcomes, and is substantially improved using techniques like association rules extraction and fuzzy logic [2]. The former helps explore large datasets to find unique patterns, while the latter enables better representation and analysis of real-world data by managing incomplete or vague information [3].

Within medical databases, there are often incomplete and imprecise data. As a solution, fuzzy association rules, a suitable tool for discretizing numerical values and revealing hidden relationships in a user-friendly manner, have been employed [4]. Noting that standard association rule mining algorithms often overlook correlations between tasks, we have focused on a distributed approach to resolving multitasking rule problems, bringing together several rules, and dividing the problem into individual tasks [5], [6].

Given that medical databases mainly store patients' historical data, implementing algorithms in a distributed environment is necessary when dealing with voluminous data [reference].

In this context, we propose a distributed, fuzzy-based medical system for pattern mining. This system, using tools like association rules, fuzzy logic, and Big Data, enables end-users to uncover and interpret hidden patterns from health-related data, aiding diagnostic analysis, subsequent treatment, and early potential disease prevention. The proposed system is entirely automatic and can support both labeled and unlabeled data. Our primary contributions are data enrichment, feature fuzzification, and efficient visualization [references].

Section-wise, this paper first reviews significant work in this field, presents the proposed fuzzy-based system for pattern mining, provides a case study from real-world medical data and discusses the resulting findings, and finally presents the derived conclusions.

II. Background

In this paper, we propose a novel fuzzy information-based system for uncovering hidden patterns and connections to improve the interpretability and diagnostic and co-morbidity analysis of medical data [7], [8]. A central challenge in this field is how results are presented to the end user, leading our research to focus on enhancing the interpretability of findings via data mining techniques such as association rule discovery and fuzzy logic.

Practical interpretation of medical data often requires employment of fuzzy techniques to convert vague data into decipherable information for end users. Prominent studies have introduced measures of accuracy for the extraction of fuzzy association rules from medical relational databases, streamlined prescriptions for diabetic patients using wearable sensors, and applied a fuzzy programming algorithm to negate uncertainty [9]–[11].

Considering co-morbidity, or the occurrence of a second disease due to the existence of a primary disease in the same patient, we are able to interpret patients as simple, medium, or complex based on their count of co-morbidity diagnoses [12], [13].

This paper presents a comprehensive system for managing and extracting information in healthcare systems employing big data architecture. This includes a transformative process of data integration, enrichment, and the novel implementation of fuzzy techniques to address data uncertainty. The system employs an algorithm for the extraction of fuzzy association rules and implements visualization tools to enhance end-user interpretation of results.

III. Pattern Mining System with Fuzzy Logic in Big Data

We propose a robust system for managing and extracting data in big health systems through the implementation of big data architecture. The process integrates data, refines it, and introduces fuzzy techniques to handle uncertainty. It also embeds an algorithm to extract fuzzy association rules and visualization tools to improve interpretability of the results by the users.

Implemented in two Spanish hospitals, the system is capable of collecting and merging varied patient data from multiple hospital sources, creating a patient-centric model. Data preprocessing and enrichment is performed using information from external taxonomies and databases, improving the interpretability of variables, especially those involving diagnoses, patient origin, reasons for admission or discharge, surgical procedures, and the like. This constellates into a unique 'fuzzification' process, guided by expert knowledge, that paves the way for more interpretable results.

IV. Conclusion

This research developed a data mining system using the Big Data framework, applied to datasets from two hospitals in southern Spain. Improvements were made in data interpretability by enriching features and applying a fuzzification algorithm. The system was deployed using the Spark platform, facilitating the analysis of large volumes of data generated in hospitals. Practical experimentation with real data demonstrated the system's efficacy, discovering meaningful rules connecting various patient conditions, which could be utilized in predicting diseases, exploring relationships across features, and analyzing co-morbidity.

Acknowledgment

This research paper is part of the COPKIT project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 786687. This paper was also supported by grant PID2021-123960OB-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe and by grant TED2021-1289402B-C21 founded by MCIN/AEI/10.13039/501100011033 and, BIGDATAMED project, which has received funding from the Andalusian Government (Junta de Andalucía) under grant agreement No P18-RT-1765 and by the European Union NextGenerationEU/PRTR.

References

[1] A. R. Feinstein, "The pre-therapeutic classification of co-morbidity in chronic disease," *Journal of Chronic Diseases*, vol. 23, no. 7, pp. 455–468, 1970.

[2] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," *Artificial Intelligence*, vol. 267, pp. 1–38, 2019.

[3] B. Kim, R. Khanna, and O. O. Koyejo, "Examples are not enough, learn to criticize! criticism for interpretability," *Advances in neural information processing systems*, vol. 29, 2016.

[4] P. Y. TAŞER, K. U. BİRANT, and D. Birant, "Multitask-based association rule mining," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 28, no. 2, pp. 933–955, 2020.

[5] D. d. Castro Rodrigues, V. Siqueira, F. Tavares, M. Lima, F. Oliveira, L. Osco, W. Junior, R. Costa, and R. Barbosa, "Discovering associative patterns in healthcare data," in *Proceedings of Sixth International Congress on Information and Communication Technology*. Springer, 2022, pp. 371–379.

[6] I. R. Mewes, H. Jenzer, and F. Einsele, "A study about discovery of critical food consumption patterns linked with lifestyle diseases for swiss population using data mining methods." in *HEALTHINF*, 2021, pp. 30–38.

[7] J. Yoon, D. Jeong, C.-H. Kang, and S. Lee, "Forensic investigation framework for the document store nosql dbms: Mongodb as a case study," *Digital Investigation*, vol. 17, pp. 53–65, 2016.

[8] M. D. Ruiz, J. Gomez-Romero, C. Fernandez-Basso, and M. J. Martin-Bautista, "Big data architecture for building energy managament systems," *IEEE Transactions on Industrial Informatics*, 2021.

[9] A. Beam and I. Kohane, "Big data and machine learning in health care," *JAMA - Journal of the American Medical Association*, vol. 319, no. 13, pp. 1317–1318, 2018.

[10] C. Lee and H.-J. Yoon, "Medical big data: Promise and challenges," *Kidney Research and Clinical Practice*, vol. 36, no. 1, pp. 3–11, 2017.

[11] M. Sabzevari and E. Imani, "Separation of movement direction concepts based on independent component analysis algorithm, linear discriminant analysis, deep belief network, artificial and fuzzy neural networks," *Biomedical Signal Processing and Control*, vol. 62, 2020.

[12] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, "Data processing and text mining technologies on electronic medical records: A review," *Journal of Healthcare Engineering*, vol. 2018, 2018.

[13] S. Lakshmanaprabu, S. Mohanty, S. S., S. Krishnamoorthy, J. Uthayakumar, and K. Shankar, "Online clinical decision support system using optimal deep neural networks," *Applied Soft Computing Journal*, vol. 81, 2019.