



UNIVERSIDAD DE CÁDIZ

DEPARTAMENTO DE MATEMÁTICAS

**Optimización y caracterización
de sistemas de variables vía
teoría de conjuntos rugosos**

FERNANDO CHACÓN GÓMEZ

TESIS DOCTORAL

Enero 2024

Dra. María Eugenia Cornejo Piñero, Profesora Titular de Universidad del Departamento de Matemáticas de la Universidad de Cádiz

Dr. Jesús Medina Moreno, Profesor Catedrático de Universidad del Departamento de Matemáticas de la Universidad de Cádiz,

HACEN CONSTAR:

Que D. Fernando Chacón Gómez, Graduado en Matemáticas, ha realizado en el Departamento de Matemáticas de la Universidad de Cádiz, bajo su dirección, el trabajo de investigación correspondiente a su Tesis Doctoral titulado:

Optimización y caracterización de sistemas de variables vía teoría de conjuntos rugosos

Revisado el presente trabajo, estiman que puede ser presentado al Tribunal que ha de juzgarlo. Y para que conste a efectos de lo establecido en el Real Decreto 99/2011, autorizan la presentación de este trabajo en la Universidad de Cádiz.

Cádiz, a 22 de enero de 2024

Dra. María Eugenia Cornejo Piñero

Dr. Jesús Medina Moreno

A mi familia

Agradecimientos

Estos tres últimos años han supuesto un gran reto para mí, llenos de duro trabajo que tiene como punto y aparte esta tesis doctoral, pero también repletos de nuevas experiencias que no esperaba vivir. Por todo esto solo puedo decir: gracias.

En primer lugar, gracias a mis directores de tesis Jesús y Maru por su apoyo, ayuda, dedicación, disponibilidad, esfuerzo, paciencia y un largo etcétera. Gracias por ser fuente de inspiración y por hacerme descubrir el maravilloso (aunque a veces frustrante) mundo de la investigación.

Gracias a todos los miembros del grupo de investigación M·CIS por haberme acogido como uno más desde el principio, como un nuevo hijo que llega a la familia. También gracias a los compañeros de departamento y despacho por compartir el día a día.

Por último, pero no por ello menos importante, gracias a mis amigos y familiares por celebrar juntos los hitos y brindarme su apoyo en los momentos difíciles. En especial, gracias a mis padres por ser un espejo en el que mirarme, por animarme a estudiar lo que quisiese, por apoyarme en mis decisiones y por enseñarme desde bien pequeño que todo trabajo tiene su recompensa.

* Tesis financiada parcialmente por el contrato predoctoral UCA/REC13VPCT/2020 asociado al proyecto PID2019-108991GB-I00 y por la ayuda del Plan Propio de la Universidad de Cádiz 2022-2023 con el fin de realizar una estancia para la Mención Doctorado Internacional.

Índice general

English Summary	XIII
Resumen	XVII
1. Contexto general de la tesis doctoral	1
1.1. Introducción y justificación del estudio	1
1.2. Hipótesis, objetivos y análisis crítico de antecedentes	7
1.3. Contribución y estructura del estudio	10
2. Preliminares	15
2.1. Teoría de conjuntos rugosos	15
2.1.1. Tablas de información y tablas de decisión	15
2.1.2. Reglas y algoritmos de decisión	24
2.2. Teoría de conjuntos rugosos difusos	28
3. Discusión conjunta de los resultados	35
3.1. Reductos, reductos valor y birreductos	36
3.1.1. Relación entre reductos y birreductos	38
3.1.2. Relación entre reductos valor y birreductos	43
3.2. Reglas y algoritmos de decisión en teoría de conjuntos rugosos difusos	47

3.2.1. Reglas de decisión	48
3.2.2. Algoritmos de decisión	52
3.3. Clasificación formal de nuevos objetos	58
3.3.1. Método de idoneidad	61
3.3.2. Método de ε -representatividad	64
3.3.3. Método de α -credibilidad	68
3.3.4. Método de α -credibilidad orientado por el soporte	70
4. Reductos valor y birreductos: Un estudio comparativo	73
5. Algoritmos de decisión en conjuntos rugosos para modelar con incertidumbre	95
6. Toma de decisiones en la teoría de conjuntos rugosos difusos	119
7. Conclusiones y trabajo futuro	149
7.1. Conclusiones alcanzadas	149
7.2. Perspectiva	153
Anexos. Otras publicaciones relacionadas	157
A. <i>Information Bireducts and Its Relationship with Reducts</i>	159
B. <i>Fuzzy Rough Set Decision Algorithms</i>	169
C. <i>Efficiency of fuzzy rough set decision algorithms</i>	185
D. <i>Relating decision rules and attribute implications</i>	195
Bibliografía	205

English Summary

Rough Set Theory (RST) has been widely studied since its introduction in the 1980s by Z. Pawlak. This theory is a mathematical tool for the analysis of relational datasets with imprecise or incomplete information. In this framework, datasets are interpreted as information tables composed of a set of objects and a set of attributes/properties that try to characterize every object. This thesis studies two important research topics in RST: attribute reduction and object classification. Although both topics can be studied independently, this work will show their close relationship.

This thesis is organized as follows. Chapter 1 presents the frameworks for the study and outlines the objectives to be achieved. Moreover, it also summarizes the contributions presented and details the structure of the work. Next, Chapter 2 gathers basic notions for the development of this thesis.

Chapter 3 is devoted to discuss the advances obtained in three of the most significant published papers. The first work focused on the relationship between reducts and value reducts with bireducts in RST. This environment allows to reduce the size of the analyzed dataset by eliminating irrelevant information. Reducts and value reducts handle the reduction of the set of attributes, while bireducts reduce the sets of attributes and objects simultaneously. However, the relationship between the previous reduction mechanisms has not been analyzed in the literature yet. The results of the comparative analysis have been published in:

- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Value reducts and bireducts: A comparative study. *Mathematical Methods in the Applied Sciences*, 46(2):1631–1650, 2023.

This paper shows how to obtain reducts and value reducts from bireducts, and vice versa, both in information and decision systems, as well as with the three existing types of bireducts. This relationship simplifies the computation of the sets of objects and attributes obtained from these notions in both environments. Furthermore, it provides a better understanding of bireducts and allows to study in detail the characteristics of value reducts, which have not been investigated as much as reducts and bireducts.

Next, the formal extension of several notions of RST has been carried out. This work is fundamental for modelling databases with imperfect information, such as real datasets, and will enable the classification of new objects to these databases in a further process. In RST, the studies of relational databases rely on the notion of decision algorithm, which is composed of a set of decision rules that satisfy meaningful properties in order to represent adequately the information contained in the given database and to make deductions. However, despite the relevance of this notion for data analysis, the corresponding studies have major limitations due to the rigidity of the classical setting. To address this issue, the paper

- 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.

is devoted to extend the previous studies to Fuzzy Rough Set Theory (FRST). This environment is an ideal setting to analyze databases due to the flexibility given by the fuzzy philosophy.

The aforementioned paper includes the generalization of numerous RST notions related to decision rules to the FRST framework, such as sup-

port, strength, certainty, etc., and especially highlighting the introduction of fuzzy decision algorithms. This fact is fundamental for studying datasets with imperfect and inconsistent information.

The use of decision rules in the FRST framework helps significantly to the classification of new data. Nowadays, this is a topic of great interest due to its relation with the design of intelligent systems, as well as its impact in many fields, such as social, clinical or financial. In

- F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Decision making in fuzzy rough set theory. *Mathematics*, 11(19), 2023.

the results of the previous paper are used in order to provide new definitions, properties, and mathematical mechanisms for decision-making, including four methods to classify new objects. Each approach faces the decision-making problem from a different perspective, so it is possible to apply them simultaneously for a more solid classification. This fact, together with the benefits of using the FRST environment, provides the user with significant flexibility in addressing any classification problem starting from databases with incomplete information.

It is important to emphasize that the papers mentioned above are published in high-impact scientific journals, specifically in three journals that are collected in the first quartile (Q1) of the “*Mathematics*” and “*Mathematics, Applied*” category of the *Journal Citation Reports*. These manuscripts have been included in Chapters 4, 5 y 6.

Finally, Chapter 7 presents the most relevant conclusions of this work and some lines of future work in the short, medium and long term.

Resumen

Desde su introducción en los años ochenta por Z. Pawlak, la teoría de conjuntos rugosos (RST, de sus siglas en inglés, *Rough Set Theory*) ha sido ampliamente estudiada. Esta teoría conforma una herramienta matemática que permite analizar conjuntos de datos relacionales que contienen información imprecisa o incompleta. En este marco de trabajo, los conjuntos de datos se interpretan como tablas de información, constituidas por un conjunto de objetos y un conjunto de atributos que definen a los primeros.

Esta tesis doctoral se enmarca en dos líneas de investigación importantes en RST: la reducción o selección de atributos y la clasificación de nuevos objetos. Aunque ambas líneas se pueden desarrollar de manera independiente, a lo largo de este trabajo, se mostrará que están estrechamente relacionadas.

Diferentes capítulos conforman la presente tesis doctoral. En el Capítulo 1 se presentan los marcos de trabajo en los que tiene lugar este estudio, así como los objetivos a lograr. Además, también se sintetizan las contribuciones presentadas y se detalla la estructura de esta memoria. A continuación, el Capítulo 2 recopila las nociones previas imprescindibles para el desarrollo de esta tesis doctoral.

En el Capítulo 3 se expone una discusión sobre los avances derivados de la investigación realizada en tres de los artículos publicados más significativos. En primer lugar, se muestra un estudio pormenorizado de la relación entre los reductos y reductos valor con los birreductos en RST.

Una de las grandes ventajas que proporciona este entorno es la reducción del tamaño de los conjuntos de datos mediante la eliminación de información irrelevante, tanto en el conjunto de atributos, considerando reductos y reductos valor, como en ambos conjuntos, el de atributos y el de objetos, usando birreductos. Sin embargo, en la literatura no se ha estudiado la relación entre ambos mecanismos de reducción. Los resultados correspondientes al análisis comparativo realizado durante el desarrollo de la presente tesis doctoral se han publicado en

- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Value reducts and bireducts: A comparative study. *Mathematical Methods in the Applied Sciences*, 46(2):1631–1650, 2023.

En dicho artículo se muestra la posibilidad de obtener reductos y reductos valor a partir de birreductos, y viceversa, tanto en sistemas de información como de decisión, así como con los tres tipos de birreductos que existen. Esta relación simplifica el cálculo de los conjuntos de objetos y atributos en ambos ambientes y proporciona un mayor conocimiento de los birreductos. Además, permite conocer mejor las características de los reductos valor, que no han sido estudiados con tanto detalle como los reductos y birreductos.

A continuación, se ha llevado a cabo la extensión formal de distintas nociones de RST, fundamentales para el modelado de bases de datos con información imperfecta, lo cual posibilitará en un siguiente proceso la clasificación de nuevos objetos en este tipo de bases de datos. Este hecho es especialmente relevante para el estudio de conjuntos de datos reales, debido a la inconsistencia e imprecisión presente en ellos. En RST, estos estudios se apoyan en la noción de algoritmo de decisión, que está formado por un conjunto de reglas de decisión satisfaciendo una serie de propiedades para representar adecuadamente la información presente en la base de datos y realizar deducciones. Sin embargo, a pesar de la relevancia de

esta noción para el análisis de datos, los estudios correspondientes tienen grandes limitaciones debido a la rigidez del entorno clásico. Con el fin de resolver este inconveniente, el artículo

- 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.

está dedicado a trasladar los estudios anteriores al marco de trabajo proporcionado por la teoría de conjuntos rugosos difusos (FRST, de sus siglas en inglés, *Fuzzy Rough Set Theory*). Este entorno constituye un escenario ideal para realizar dicha tarea debido a la flexibilidad propia de la filosofía difusa.

El artículo mencionado anteriormente recoge la generalización de numerosas nociones de RST relativas a las reglas de decisión al marco de FRST, como la de soporte, fuerza, certeza, etc., y destacando especialmente la introducción de los algoritmos de decisión difusos. Este hecho es fundamental para abordar el estudio de conjuntos de datos con información imperfecta y con inconsistencias.

El uso de las reglas de decisión en el entorno de FRST facilita considerablemente la clasificación de datos nuevos. Este es un tema de gran interés en la actualidad debido a su relación con el diseño de sistemas inteligentes, así como por su impacto en numerosos ámbitos, como el social, clínico o financiero. En

- F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Decision making in fuzzy rough set theory. *Mathematics*, 11(19), 2023.

se utilizan los resultados del artículo previo para proporcionar nuevas definiciones, propiedades y mecanismos matemáticos para la toma de decisiones. En particular, se incluyen cuatro métodos de clasificación de nue-

vos objetos. Cada uno de ellos aborda la problemática de la toma de decisiones desde un punto de vista diferente, permitiendo aplicarlos simultáneamente para obtener una clasificación más sólida. Este hecho, sumado a las ventajas derivadas del uso del marco de FRST, confiere al usuario de una gran libertad para afrontar cualquier problema de clasificación con una base de datos de partida con información imperfecta.

Cabe destacar que los artículos mencionados anteriormente se encuentran publicados en revistas científicas de alto impacto, concretamente en tres artículos que están recogidos en el *Journal Citation Reports*. Dichos artículos se han incluido en los Capítulos 4, 5 y 6.

Finalmente, el Capítulo 7 presenta las conclusiones más relevantes del trabajo realizado y algunas líneas de trabajo futuro a corto, medio y largo plazo.

Capítulo 1

Contexto general de la tesis doctoral

1.1. Introducción y justificación del estudio

La sociedad actual se encuentra inmersa en un proceso continuo de transformación tecnológica y digital, dando lugar a sistemas ágiles de interacción, participación y utilización de los servicios ofrecidos por la comunidad científica, sanitaria, educativa, empresarial, financiera, etc. Este proceso de transformación y su acelerada integración en las actividades de la vida cotidiana, han producido un notable incremento en la cantidad de datos generados, conteniendo de manera intrínseca información con incertidumbre, imprecisa, incompleta, contradictoria. Este tipo de información es usual hoy día por la forma de expresar el conocimiento, por los mecanismos de registro de los datos y por el mismo procesamiento de los datos, que no suelen tener protocolos que aseguren la “perfección” de los datos. Con el propósito de obtener el máximo rendimiento y expresar todo el potencial proporcionado por las bases de datos (reales) con información imperfecta, se hace necesario utilizar mecanismos innovadores para un

tratamiento adecuado y una gestión óptima de dicha información, con el objetivo indispensable en la era actual de mejorar procesos, pronosticar comportamientos futuros y elaborar estrategias para la toma de decisiones.

Las matemáticas y las ciencias de la computación son dos disciplinas íntimamente relacionadas de gran utilidad para afrontar este reto. Por un lado, las matemáticas ofrecen un marco de trabajo sólido con herramientas y técnicas formales que facilitan la representación y el manejo de la información con incertidumbre, favoreciendo el razonamiento sobre la misma. Por otro lado, las ciencias de la computación se encargan de la implementación de dicha teoría a través de la elaboración de algoritmos, programas y aplicaciones informáticas eficientes. Por consiguiente, las sinergias entre ambas disciplinas permiten el desarrollo y progreso científico en el estudio de conjuntos de datos con información imprecisa, acometiéndolo por medio de enfoques complementarios.

En el ámbito matemático se han desarrollado diferentes herramientas dedicadas al estudio de conjuntos de datos. Las contribuciones presentadas en la presente tesis doctoral se enmarcan fundamentalmente en dos de ellas. Por un lado, Lotfi A. Zadeh introdujo en la década de los 60 la teoría de conjuntos difusos [101] (FST, de sus siglas en inglés, *Fuzzy Set Theory*) como una generalización de la noción de conjunto clásica, en la que la pertenencia de los elementos a los conjuntos difusos queda determinada por un valor (grado de pertenencia) en el intervalo unidad. De aquí, cuanto mayor sea dicho valor, mayor será la pertenencia del elemento correspondiente al conjunto dado. Por ejemplo, si se considera el conjunto de “personas de avanzada edad”, una persona con 90 años pertenece a dicho conjunto en un mayor grado que una con 61 años. Además, el de esta última no debería ser muy diferente del grado de la de 60 años, evitando los grandes saltos entre la pertenencia completa (con valor 1) y la no pertenencia (con valor 0) que existen cuando se consideran conjuntos

“clásicos”. Los conjuntos difusos han sido ampliamente estudiados desde su introducción por su capacidad de modelar información imperfecta, conteniendo inconsistencias e imprecisiones. La relevancia de esta teoría radica especialmente en que ofrece un marco de trabajo muy flexible, contando con una gran capacidad de adaptación a la subjetividad del usuario y a los datos de partida. Por otro lado, Pawlak presentó en los años 80 la teoría de conjuntos rugosos [71] (RST, de sus siglas en inglés, *Rough Set Theory*) para el manejo de información incompleta e inconsistente, aunque desde un punto de vista distinto del que propuso Zadeh. Esta teoría matemática facilita también el tratamiento de datos y permite la extracción de conclusiones relevantes, por lo que constituye otro entorno ideal para el estudio de la información almacenada en conjuntos de datos. Además, se han combinado ambas filosofías, las de RST y FST, obteniendo como consecuencia la teoría de conjuntos rugosos difusos (FRST, de sus siglas en inglés, *Fuzzy Rough Set Theory*), logrando dotar a este marco de trabajo de una mayor flexibilidad. Dado que RST será el pilar fundamental para el desarrollo de esta tesis doctoral, a continuación, se presentará esta teoría formal con más detalle, junto con su extensión difusa.

Teoría de conjuntos rugosos

La teoría de conjuntos rugosos [70, 71] fue introducida como una herramienta matemática para el manejo de información imperfecta, incompleta, inconsistente, ..., contenida en las bases de datos relacionales. Su principal objetivo es proporcionar un marco de trabajo formal para traducir de manera automática los datos en información y de ahí a conocimiento, con el fin de interpretar los datos de partida y realizar, por ejemplo, tareas de clasificación de nuevos objetos, que puede interpretarse como el diagnosticar a un nuevo paciente en función de sus síntomas y de la teoría y experiencia del médico, pero en este caso de forma automática. A partir de

los datos originales se construye la estructura formal sobre la que definir las distintas nociones y funciones matemáticas. En un primer momento, se identifica un conjunto de objetos que cumplen una serie de propiedades o atributos, estas entradas forman lo que se llama en RST una tabla de información, también conocida como sistema de información. Esta interpretación matemática de las bases de datos relacionales posibilita llevar a cabo estudios detallados de las mismas, ya que facilita considerablemente su análisis.

El impacto de RST en otras teorías relacionadas es considerable, dando lugar a numerosos estudios que adaptan nociones de RST para su uso en otros marcos de trabajo. En particular, en [75] se combinó el algoritmo de los K -vecinos más cercanos con RST para analizar con más detalle aquellos objetos con mayor inconsistencia en la tabla estudiada, detectados por medio de las nociones de RST. Además, en [4] se adaptó la noción algebraica de congruencia para introducir una reducción de atributos óptima basada en RST en el marco de trabajo del análisis de conceptos formales [44, 95] (FCA, de sus siglas en inglés, *Formal Concept Analysis*). Recíprocamente, la filosofía de FCA se utilizó en [6, 7, 10, 63] para la reducción de atributos en extensiones de RST.

Debido al gran tamaño de las bases de datos en la actualidad, cada vez es más común encontrar problemas para el manejo y tratamiento de la información almacenada [13, 42, 43, 81]. Este hecho hace necesario llevar a cabo un preprocesamiento de los datos, con el fin de eliminar información redundante y realizar estudios de manera más eficiente [2, 3, 4, 10, 21]. En RST surge la noción de reducto, como un subconjunto minimal de atributos que preservan la información presente en el conjunto de datos dado. Debido a la importancia de esta noción, los reductos han sido ampliamente estudiados en la literatura [28, 49, 63, 100, 103].

Al margen de la noción de reducto, existen otras nociones centradas en la reducción de información irrelevante. En particular, Pawlak introdu-

jo los reductos valor [72] para proporcionar una mayor reducción de las bases de datos relacionales. Los reductos valor se encuentran asociados a cada objeto de la tabla, y conforman un subconjunto minimal de atributos que preserva la información dada por dicho objeto. De esta forma, es posible descartar algunos valores de atributos para cada objeto, proporcionando una mayor reducción que la dada por los reductos.

Además de reducir el conjunto de atributos, también se hace necesario reducir el conjunto de objetos, ya que algunos de ellos pueden proporcionar una información distorsionada, ser redundantes entre sí o incluso causar inconsistencias. Con este objetivo, Janusz, Ślęzak y Nguyen introdujeron la noción de birreducto [50], permitiendo reducir simultáneamente los conjuntos de objetos y atributos. Los birreductos ligan ambos conjuntos con el objetivo de evitar las redundancias en la tabla y mejorar su consistencia, descartando en cualquier caso información irrelevante o que puede distorsionar el estudio de la base de datos considerada. No obstante, hay que tener cuidado con los objetos no considerados por la eliminación de la consistencia, para estar seguros de que estos objetos no son determinantes en el estudio. Asimismo, cabe resaltar que los birreductos generalizan a los reductos y que han sido analizados en profundidad [8, 9, 11, 12, 54, 60, 88, 89].

Otro tema de especial interés en el marco de RST es el estudio de dependencias entre atributos, con el objetivo de abordar la problemática de la toma de decisiones e inferir sobre el conjunto de datos. Para ello surgen las tablas de decisión (o también llamados sistemas de decisión), las cuales son un caso particular de tablas de información en las que se destaca un atributo sobre el resto, llamado atributo de decisión. El objetivo principal de este enfoque es analizar cómo se ven afectados los valores del atributo de decisión por los valores del resto de atributos, llamados atributos de condición. Con el fin de facilitar esta tarea, a partir de las tablas de decisión se extraen reglas de decisión [5, 48, 58, 72, 89, 99], las cuales sinte-

tizan en términos lógicos la información contenida en la tabla de decisión. Las reglas de decisión van acompañadas de indicadores de relevancia, los cuales las describen desde diferentes puntos de vista y permiten extraer conclusiones significativas de las mismas. Algunos de estos indicadores de relevancia son el soporte, la certeza o la credibilidad.

Por otra parte, para llevar a cabo un estudio adecuado es imprescindible que el conjunto de reglas a analizar sea representativo para la tabla de decisión dada. Con este objetivo Pawlak también introdujo los algoritmos de decisión [73], como un conjunto de reglas de decisión independientes, con soporte no nulo, representando a todos los objetos y preservando la consistencia de la tabla, lo que asegura una correcta representación de las tablas de decisión. Como consecuencia, los algoritmos de decisión constituyen un pilar fundamental para el estudio de las tablas de decisión. Sin embargo, no se ha extendido su uso, como se argumentará cuando se realice el análisis crítico de los antecedentes.

Teoría de conjuntos rugosos difusos

La teoría de conjuntos rugosos difusos (FRST, de sus siglas en inglés, *Fuzzy Rough Set Theory*) [26, 27, 38, 52, 67, 77] surge como una generalización de RST a partir de la filosofía propia de los conjuntos difusos [101], a fin de proporcionar una mayor flexibilidad que RST. Como consecuencia, FRST permite al usuario adaptar mejor su estudio a sus necesidades y abordar una mayor gama de aplicaciones reales.

La filosofía de FRST también ha atraído el interés de otros marcos de trabajo. Por ejemplo, en [97] se estudiaron y combinaron conjuntos rugosos difusos y funciones estadísticas y de probabilidad y en [98] se utilizaron los conjuntos suaves difusos (del inglés, *fuzzy soft sets*) para proponer un método de pronóstico combinado, permitiendo así complementar una metodología basada en conjuntos rugosos.

Una de las principales aportaciones al entorno de FRST viene dada por la incorporación de la filosofía multiadjunta que se realizó en [29]. Por ejemplo, en este nuevo entorno, cada par objeto atributo puede tener asociado un par adjunto distinto, lo que permite determinar grados de preferencia sobre los objetos y atributos [65]. Otra extensión significativa de FRST, teniendo en cuenta los atributos como se hace en FCA, es la que define a los retículos de conceptos orientados a propiedades y los orientados a objetos [61, 62, 63]. Estos marcos verifican una serie de propiedades muy interesantes [22, 34, 35], como la eliminación de restricciones sobre los conjuntos tales como la conmutatividad y la asociatividad, y el uso de preferencias como se comentó anteriormente, que otorgan al entorno de FRST de una flexibilidad aun mayor.

Además, dada las ventajas que proporciona el marco multiadjunto, este entorno de trabajo también se ha considerado en otras teorías matemáticas centradas en el análisis de datos, como es el caso de FCA. Esta teoría cuenta con su generalización al caso multiadjunto [36, 64, 65], posibilitando el estudio de retículos de conceptos en gran detalle.

1.2. Hipótesis, objetivos y análisis crítico de antecedentes

Como se ha comentado anteriormente, la teoría de conjuntos rugosos ha tenido un gran impacto desde que la introdujera Pawlak en los ochenta, y han sido numerosos grupos de investigación de todo el mundo los que han adoptado y desarrollado esta teoría, tanto desde un punto de vista teórico [56, 69, 86, 93] como aplicado [1, 55, 57, 83]. Una de las líneas de trabajo más significativas es la selección de atributos, en la que destacan las nociones de reducto, reducto valor y birreducto. No obstante, la literatura que compara los reductos y los birreductos es realmente limita-

da [8, 89], y no tienen en cuenta los reductos valor. Dado el importante papel de todas estas nociones en el análisis de conjuntos de datos, se hace necesario llevar a cabo un estudio exhaustivo de dichas nociones con el fin de obtener relaciones entre ellas, de manera que sea posible, por ejemplo, calcular unas en función de otras y viceversa. Este hecho permitirá al usuario afrontar el estudio correspondiente desde el punto de vista más adecuado a sus intereses y, en cualquier momento, cambiar de enfoque para extraer conclusiones adicionales.

Por otra parte, como se expuso anteriormente, las reglas de decisión son de gran utilidad para estudiar las tablas de decisión, especialmente mediante el uso de los indicadores de relevancia. Además, para un correcto análisis de dichas tablas, es necesario la consideración de algoritmos de decisión. Estas nociones e indicadores no han llegado a estudiarse en profundidad en el marco de trabajo de FRST. Incluso los algoritmos de decisión no han tenido un impacto relevante en el caso clásico, no teniendo una clara continuidad desde su introducción en [73], sin embargo, son relevantes para representar la información de las tablas de decisión. Puesto que el entorno difuso proporciona numerosas ventajas, destacando por encima de todas la flexibilidad aportada por la estructura algebraica considerada, es de gran relevancia estudiar las medidas y nociones más representativas que aun no se han tratado con profundidad en este marco de trabajo, como son los algoritmos de decisión y la eficiencia de estos algoritmos. Para llevar a cabo este estudio será necesario la extensión de numerosas nociones relacionadas con las reglas de decisión en las que se apoyan dichos algoritmos, permitiendo un análisis detallado de las tablas de decisión en el entorno de FRST. En primer lugar, se ha introducido la noción de algoritmo de decisión en este marco de trabajo, y en un futuro próximo se realizará un análisis detallado de los mismos por medio de la generalización difusa de la noción de eficiencia.

Finalmente, la toma de decisiones es un tema de interés actual gracias

a su repercusión en numerosos ámbitos, como la psicología [92], la economía [30] o las ciencias políticas [51]. Un objetivo fundamental en los métodos para la toma de decisiones es el descubrimiento de patrones y la aplicación de este conocimiento a nuevos escenarios, que no han sido considerados en el conjunto de datos de partida, con el fin de determinar su comportamiento. Esta tarea ya se ha abordado en diferentes campos como en inteligencia artificial, clasificación dinámica [66], algoritmos basados en reglas [94], métodos para la toma de decisiones multicriterio [84, 96], práctica clínica [46] y psicoterapia [47].

La clasificación de datos nuevos también se ha estudiado en el marco de la teoría de los conjuntos rugosos debido a su potencial para tratar información imprecisa [41, 45, 68, 87, 90, 91]. Por ejemplo, en [68, 87] los nuevos objetos a clasificar son comparados con los antecedentes de las reglas de decisión consideradas. Esta filosofía es similar a la seguida por la técnica de clustering de razonamiento automático (en inglés, *machine learning*) [59], considerando cada regla como un cluster y estudiando la pertenencia del nuevo objeto a cada una de ellos. A continuación, en [68] se realiza una votación entre las diferentes decisiones posibles a tomar para el nuevo objeto mediante el soporte de las reglas cuyo antecedente coincide totalmente con el nuevo objeto, lo cual es una condición realmente exigente. En cambio, [87] presenta una alternativa mediante el uso de relaciones de tolerancia, las cuales proporcionan un valor en el intervalo unidad para mostrar la similitud entre el objeto nuevo a clasificar y el antecedente de la regla de decisión comparada. De esta forma, la mejor decisión a tomar viene dada por el mayor valor obtenido de la relación de tolerancia. Por otra parte, el método de clasificación propuesto en [45] es una combinación de los enfoques presentados en [68, 87]. Respecto a los atributos del antecedente con valores categóricos, se requiere que estos valores sean iguales a los tomados por el objeto en dichos atributos, mientras que para los atributos con valores numéricos se utiliza una distancia euclídea para calcular

la similitud del antecedente y del objeto nuevo. Dada la relevante caracterización obtenida por las relaciones de tolerancia separables [102], nos apoyaremos en esta noción para definir los métodos de clasificación propuestos, permitiendo obtener resultados más sólidos que en los métodos introducidos en [45, 68, 87, 90].

Por otro lado, en [41, 91] se usó el indicador de relevancia de la credibilidad para la clasificación de nuevos objetos en función de t-normas e implicaciones difusas. El principal objetivo fue complementar cada comparativa entre el objeto a clasificar y los antecedentes de las reglas de decisión consideradas con la credibilidad de dichas reglas. En consecuencia, la decisión a tomar para el nuevo objeto vendrá dada por una regla de decisión con credibilidad elevada. En este trabajo los estudios se enmarcarán en un marco multiadjunto, por lo que se extenderá tanto la noción de credibilidad como los métodos de clasificación basados en ella. Con ello, será posible definir métodos de clasificación más generales, proporcionando una mayor libertad al usuario encargado del estudio correspondiente.

1.3. Contribución y estructura del estudio

Esta sección está dedicada a presentar las principales contribuciones de la tesis doctoral, que se han publicado en distintos artículos científicos y capítulo de libro de editoriales de prestigio con impacto, destacando los tres artículos publicados en el primer cuartil (Q1) de las categorías de «Matemáticas» y «Matemática Aplicada» del *Journal Citation Reports*. En primer lugar, el Capítulo 2 recoge las nociones de RST y FRST que han sido fundamentales para el desarrollo de esta tesis doctoral, fijando de esta forma los marcos de trabajo correspondientes.

Por otro lado, el Capítulo 3 se encuentra dividido en tres partes para presentar las principales definiciones y resultados introducidos en cada

uno de los tres estudios, publicados en revistas científicas de alto impacto, que sustentan esta tesis doctoral. En la Sección 3.1, que se corresponde con el artículo

- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Value reducts and bireducts: A comparative study. *Mathematical Methods in the Applied Sciences*, 46(2):1631–1650, 2023.

se exponen las relaciones obtenidas entre los reductos y birreductos por un lado, y los reductos valor y birreductos por otro, tanto en tablas de información como de decisión. Estas relaciones son interesantes para el análisis de conjuntos de datos, ya que permiten obtener la reducción de información proporcionada por unas nociones a partir de las otras y, como consecuencia, estudiarlas en profundidad. El artículo mencionado se incluye íntegramente en el Capítulo 4. Por su parte, la Sección 3.2 recopila las contribuciones aportadas en el artículo

- 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.

En él, se han estudiado diversas propiedades de las nociones introducidas, relacionadas con las reglas y algoritmos de decisión, mostrando que dichas nociones generalizan a sus equivalentes en el marco de RST. En particular, los algoritmos de decisión se han presentado en función de cinco parámetros, los cuales garantizan la independencia de las reglas de decisión que los conforman, su admisibilidad, la representación de la información dada por todos los objetos y la preservación de la consistencia presente en el conjunto de datos analizados. El uso de estas nociones permite abordar el estudio de conjuntos de datos directamente en el marco de trabajo de FRST. Dicho artículo se incorpora en el Capítulo 5. Posteriormente, en la Sección 3.3 se presentan los cuatro métodos de clasificación de nuevos objetos introducidos en el artículo

- F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Decision making in fuzzy rough set theory. *Mathematics*, 11(19), 2023.

junto con algunas nociones previas indispensables para su aplicación. Cada uno de estos métodos se basa en la comparativa del nuevo objeto con cada regla de decisión considerada para realizar la clasificación correspondiente. La principal diferencia entre estos métodos radica en la filosofía seguida a posteriori. En particular, el primer método propuesto solo tiene en cuenta la semejanza del objeto nuevo con el antecedente de las reglas comparadas, mientras que el segundo método también considera la representatividad de dichas reglas por medio del soporte difuso. Por otro lado, los dos últimos métodos se apoyan en la credibilidad de la regla comparada y la semejanza mencionada anteriormente, y la diferencia entre ambos es la inclusión del soporte difuso de las reglas consideradas en el cuarto método. El hecho de afrontar la clasificación de nuevos objetos desde enfoques diferentes facilita el proceso de la toma de decisiones mediante el uso combinado de los métodos definidos. El artículo anterior se incluye en el Capítulo 6.

Además de los artículos mencionados anteriormente publicados en revistas de impacto (Q1), durante el desarrollo de la presente tesis doctoral se han publicado los siguientes capítulos de libro en editoriales de prestigio y con factor de impacto que han contribuido a la realización de la misma:

- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Information Bireducts and Its Relationship with Reducts, *Studies in Computational Intelligence* 1040, pp. 37–44, 2023.
- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Fuzzy rough set decision algorithms. In D. Ciucci, I. Couso, J. Medina, D. Slezak, D. Petturiti, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in*

Knowledge-Based Systems - 19th International Conference, IPMU 2022, Milan, Italy, July 11-15, 2022, Proceedings, Part I, volume 1601 of *Communications in Computer and Information Science*, pages 63-76. Springer, 2022.

- F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Efficiency of Fuzzy Rough Set Decision Algorithms, *Studies in Computational Intelligence* 1127:17–24, 2024.
- F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Relating decision rules and attribute implications. *In the 16th International Conference on Concept Lattices and Their Applications (CLA 2022), CEUR Workshop Proceedings*, 2022.

Estos trabajos quedan recogidos en los Anexos A, B, C y D, respectivamente. Es especialmente importante destacar el último de estos, ya que es el único enmarcado en FCA. Este estudio se basa en el análisis comparativo de las reglas de decisión de RST y las implicaciones de atributos de FCA, y constituye un pilar básico para los estudios futuros.

Además, también se han realizado estudios prácticos. En particular, se ha analizado en detalle la noción de eficiencia de algoritmo de decisión para la detección de anomalías y mantenimiento predictivo de instalaciones fotovoltaicas, apoyándonos en datos suministrados por el Grupo Energético de Puerto Real (Cádiz). Este trabajo se recoge en el siguiente artículo enviado para su revisión a una revista de impacto.

- R. G. Aragón, F. Chacón-Gómez, M. E. Cornejo, J. Medina and E. Ramírez-Poussa. Fault detection and predictive maintenance of photovoltaic facilities.

Finalmente, y fruto de la colaboración con Didier Dubois y Henri Prade, de la Universidad Paul Sabatier de Toulouse, Francia, que son investigadores sénior de elevado prestigio, se están realizando estudios en el

marco de FCA incorporando la teoría de posibilidad [39, 40]. Dicha combinación permite el análisis de información incompleta o con incertidumbre en el entorno de FCA en gran profundidad. Estos trabajos futuros, junto con otros planteados y las conclusiones extraídas tras la realización de la tesis doctoral, se presentan en el Capítulo 7.

Capítulo 2

Preliminares

En este capítulo se presentarán las nociones en las que se sustentan las definiciones y resultados obtenidos a lo largo del desarrollo de la presente tesis doctoral, a fin de favorecer su seguimiento y comprensión.

2.1. Teoría de conjuntos rugosos

En el marco de trabajo de RST las bases de datos relacionales son estudiadas por medio de tablas de información o de decisión en función de las necesidades del usuario. Aunque ambos tipos de estudio se centran en el análisis de información, las tablas de decisión destacan un atributo a analizar en función del resto. Las definiciones formales de estas nociones se introducirán a continuación.

2.1.1. Tablas de información y tablas de decisión

El uso de tablas de información y decisión [72, 73] es de gran utilidad para el estudio de bases de datos relacionales, ya que permiten representarlas de manera sintetizada. El propósito del estudio a realizar determina-

rá la consideración de las tablas de información o de decisión. En el caso de que se analicen posibles dependencias entre atributos se hará uso de las tablas de decisión, en caso contrario se considerarán las tablas de información. En primer lugar, se recuerdan las nociones relacionadas con estas últimas.

Definición 2.1. Sean U y $\mathcal{A}$ conjuntos no vacíos de objetos y atributos, respectivamente. Una tabla de información es una tupla $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ tal que $\mathcal{V}_{\mathcal{A}} = \{V_a \mid a \in \mathcal{A}\}$, donde V_a es el conjunto de valores asociados al atributo a sobre U , y $\overline{\mathcal{A}} = \{\bar{a}: U \rightarrow V_a \mid a \in \mathcal{A}\}$.

Las tablas de información se representan en forma de tabla, donde cada columna está asociada a un atributo $a \in \mathcal{A}$ y cada fila a un objeto $x \in U$, de forma que cada celda expone el valor del atributo a en el objeto x , denotado como $\bar{a}(x) \in V_a$. Este tipo de representación sintetiza sustancialmente la información conocida, por lo que es de especial utilidad para los estudios a realizar.

El análisis del conocimiento contenido en las tablas de información se lleva a cabo mediante la comparativa de los objetos en función de los valores tomados por los atributos. Para esta finalidad, resulta de gran utilidad la siguiente relación de equivalencia.

Definición 2.2. Sea $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ una tabla de información. La aplicación de indiscernibilidad $I: \mathcal{P}(\mathcal{A}) \rightarrow \mathcal{P}(U \times U)$ se define, para cada $B \subseteq \mathcal{A}$, como la relación de equivalencia

$$I(B) = \{(x, y) \in U \times U \mid \bar{a}(x) = \bar{a}(y), \text{ para todo } a \in B\}$$

llamada relación de B -indiscernibilidad. Escribimos cada clase de equivalencia de $I(B)$ como $[x]_{I(B)} = \{y \in U \mid (x, y) \in I(B)\}$. La partición en el conjunto de objetos U determinada por $I(B)$ se denota como $U/I(B) = \{[x]_{I(B)} \mid x \in U\}$.

Es conveniente destacar que esta relación de equivalencia permite agrupar en clases de equivalencia aquellos objetos que son indistinguibles de

acuerdo a los atributos de B , lo cual lleva a definir las nociones de objetos B -indiscernibles y B -discernibles.

Definición 2.3. *Dada una tabla de información $(U, \mathcal{A}, \mathcal{V}_\mathcal{A}, \overline{\mathcal{A}})$ y un subconjunto de atributos $B \subseteq \mathcal{A}$, se dice que $x, y \in U$ son B -indiscernibles si $y \in [x]_{I(B)}$. En caso contrario, se dice que x, y son B -discernibles.*

Reducir el tamaño de las bases de datos relacionales es una tarea crucial que facilita el análisis y la extracción de información, simplificando notablemente el tratamiento de los datos. Con este propósito surge la noción de reducto, la cual se centra en la eliminación de variables redundantes o superfluas para expresar el conocimiento contenido en la tabla sin pérdida de información.

Definición 2.4. *Dada una tabla de información $(U, \mathcal{A}, \mathcal{V}_\mathcal{A}, \overline{\mathcal{A}})$ y un subconjunto de atributos $B \subseteq \mathcal{A}$, se dice que:*

- $a \in B$ es dispensable en B si $I(B) = I(B \setminus \{a\})$. En caso contrario, a es indispensable en B .
- B es independiente si todos sus atributos son indispensables en B .
- $B' \subseteq B$ es un reducto de B si B' es independiente e $I(B) = I(B')$.

Conviene observar que los reductos son subconjuntos de atributos minimales que preservan la información contenida en la tabla. Por tanto, los atributos no pertenecientes a un reducto son innecesarios para estudiar la tabla de información considerada.

La siguiente noción permite reducir aun más el tamaño de la base de datos, a partir de la eliminación de algunos valores de ciertos atributos que no se han podido eliminar previamente.

Definición 2.5. *Sean $(U, \mathcal{A}, \mathcal{V}_\mathcal{A}, \overline{\mathcal{A}})$ una tabla de información, $B \subseteq \mathcal{A}$ y $x \in U$. Se dice que:*

- El valor de un atributo $a \in B$ es dispensable en B para x , si se verifica la igualdad $[x]_{I(B)} = [x]_{I(B \setminus \{a\})}$. En caso contrario, el valor de a es indispensable en B para x .
- B es independiente para x , si el valor de a es indispensable en B para x , para cada atributo $a \in B$.
- $B' \subseteq B$ es un reducto valor de B para x , si B' es independiente para x y $[x]_{I(B)} = [x]_{I(B')}$.

Nótese que los reductos valor son subconjuntos de atributos minimales que preservan la clase de equivalencia del objeto considerado. Por consiguiente, en la tabla se pueden descartar para este objeto, y para los objetos $\mathcal{A}$ -indiscernibles a él, los atributos que no pertenezcan al reducto valor.

A continuación, se presenta la noción de birreducto de información, la cual posibilita la reducción simultánea de los conjuntos de objetos y atributos sin perder información relevante.

Definición 2.6. Sean $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ una tabla de información, $X \subseteq U$ y $B \subseteq \mathcal{A}$. Se dice que:

- B es X -irreducible si no existe $B' \subset B$ tal que todos los pares de objetos $x, y \in X$ son B' -discernibles.
- X es B -inextensible si no existe $X' \subset U$ con $X \subset X'$ tal que todos los pares de objetos $x, y \in X'$ son B -discernibles.
- El par (X, B) es un birreducto de información si todos los pares de objetos $x, y \in X$ son B -discernibles, B es X -irreducible y X es B -inextensible.

Los birreductos de información proporcionan subtablas de información. Estas se encuentran conformadas por los subconjuntos maximales de objetos X y minimales de atributos B , los cuales están asociados entre sí como indica el par (X, B) en la Definición 2.6. Es fundamental observar

que en dichas subtablas se elimina toda la información redundante, desde la filosofía RST, ya que solo se consideran objetos B -discernibles. Por lo tanto, solo un objeto de cada clase de equivalencia de $I(B)$ es expuesto en dicha subtabla, reduciendo significativamente el conjunto de objetos. Como consecuencia, y dado que todas las conclusiones extraídas sobre estos objetos son extrapolables al resto de objetos de las correspondientes clases de equivalencia, los birreductos de información son de especial utilidad para el estudio de la tabla de información original.

Una vez presentadas todas las nociones enmarcadas en las tablas de información se procederá a estudiar las bases de datos relacionales analizando dependencias de atributos. Para ello, se considera un caso particular de tablas de información en el que se destaca un atributo, llamado atributo de decisión, con el fin de analizar cómo afectan los valores del resto de atributos en los valores tomados por este atributo.

Definición 2.7. 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 $\overline{\mathcal{A}_d} = \{\bar{a}: U \rightarrow V_a \mid a \in \mathcal{A}_d\}$. En este caso, a los atributos pertenecientes a $\mathcal{A}$ se les denomina atributos de condición y al atributo d atributo de decisión.

Las tablas de decisión son consideradas cuando se desea estudiar cómo los atributos de condición afectan al atributo de decisión, permitiendo extraer relaciones de dependencia. Por lo tanto, este tipo de tablas es muy útil para analizar la toma de decisiones.

A continuación, se trasladarán todos los conceptos introducidos al marco de trabajo de las tablas de decisión. En primer lugar, se puede observar que las Definiciones 2.2 y 2.3 se formalizan de manera completamente análoga y su traducción es idéntica en el entorno de las tablas de decisión, por lo que se omitirán. Para presentar la noción equivalente a la Definición 2.4

en el entorno de las tablas de decisión, es necesario introducir la noción de región positiva.

Definición 2.8. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}$ y $x \in U$. La región positiva de la partición $U/I(\{d\})$ con respecto a B se define como:

$$POS_B(\{d\}) = \bigcup_{X \in U/I(\{d\})} B_*(X)$$

donde $B_*(X) = \{x \in U \mid [x]_{I(B)} \subseteq X\}$ es la aproximación inferior de X mediante B . De aquí,

$$POS_B(\{d\}) = \{x \in U \mid [x]_{I(B)} \subseteq [x]_{I(\{d\})}\}$$

Si $x \in POS_B(\{d\})$ entonces se dice que x es un objeto consistente. Si $POS_B(\{d\}) = \mathcal{A}$ se dirá que la tabla de decisión es consistente.

La región positiva es de suma importancia en las tablas de decisión ya que, fijado el subconjunto de atributos B respecto al que se estudia la tabla, descarta los objetos B -indiscernibles que son d -discernibles. Consecuentemente, los objetos proporcionados por la región positiva no presentan contradicciones con ningún otro objeto de la tabla de decisión.

La reducción del conjunto de atributos en las tablas de decisión se realiza haciendo uso de los conjuntos d -reductos, los cuales se introducen a continuación.

Definición 2.9. Dada una tabla de decisión $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ y un subconjunto de atributos $B \subseteq \mathcal{A}$, se dice que:

- $a \in B$ es d -dispensable en B si $POS_B(\{d\}) = POS_{B \setminus \{a\}}(\{d\})$. En caso contrario, el atributo a es d -indispensable en B .
- B es d -independiente si todos sus atributos son d -indispensables en B .
- $B' \subseteq B$ es un d -reducto de B si B' es d -independiente y se verifica que $POS_B(\{d\}) = POS_{B'}(\{d\})$.

Dado que la región positiva proporciona el conjunto de objetos consistentes, los d -reductos son subconjuntos de atributos minimales que preservan la consistencia de la tabla original. En consecuencia, los atributos no pertenecientes a un d -reducto son innecesarios, ya que no son determinantes para conservar la consistencia presente en la tabla de decisión.

La siguiente definición permite estudiar la dispensabilidad de los valores de atributos en cada objeto en las tablas de decisión.

Definición 2.10. *Dada una tabla de decisión $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, un subconjunto de atributos $B \subseteq \mathcal{A}$ y $x \in U$, se dice que:*

- *El valor de un atributo $a \in B$ es d -dispensable en B para x , si*

$$[x]_{I(B)} \subseteq [x]_{I(\{d\})} \quad \text{y} \quad [x]_{I(B \setminus \{a\})} \subseteq [x]_{I(\{d\})}$$

En caso contrario, el valor del atributo a es d -indispensable en B para x .

- *B es d -independiente para x si el valor de a es d -indispensable en B para x , para cada atributo $a \in B$.*
- *$B' \subseteq B$ es un d -reducto valor de B para x si B' es d -independiente para x y la inclusión $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$ implica que $[x]_{I(B')} \subseteq [x]_{I(\{d\})}$.*

Se deduce que los d -reductos valor son subconjuntos minimales de atributos que preservan la consistencia del objeto estudiado. Además, es fácil ver que objetos $\mathcal{A}$ -indiscernibles comparten d -reductos valor.

Para la reducción simultánea del conjunto de objetos y atributos en tablas de decisión se consideran dos definiciones diferentes de birreductos, las cuales también eliminan las inconsistencias de la tabla de decisión original. La diferencia entre ambas definiciones radica en los objetos comparados. Por un lado, en los birreductos de decisión se comparan los objetos pertenecientes a dichos birreductos. Por otro, en los U -birreductos de decisión se compara cada objeto perteneciente al U -birreducto con cada objeto

perteneciente a la tabla de decisión. En primer lugar, se presentan los birreductos de decisión.

Definición 2.11. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $X \subseteq U$ y $B \subseteq \mathcal{A}$. Se dice que:

- B es X -irreducible si no existe $B' \subset B$ tal que todos los pares de objetos $x, y \in X$ d -discernibles son B' -discernibles.
- X es B -inextensible si no existe $X' \subset U$ con $X \subset X'$ tal que todos los pares de objetos $x, y \in X'$ d -discernibles son B -discernibles.
- El par (X, B) es un birreducto de decisión si todos los pares de objetos $x, y \in X$ d -discernibles son B -discernibles, B es X -irreducible y X es B -inextensible.

Mediante los birreductos de decisión es posible obtener subtablas de decisión, conformadas por los subconjuntos de objetos y atributos de condición determinados por el birreducto en cuestión, junto con el atributo de decisión de partida. Además, por la condición de discernibilidad impuesta sobre los objetos, estas subtablas son consistentes, de manera que la información contenida en ellas no da lugar a contradicciones.

Nótese que los birreductos de decisión eliminan las inconsistencias a nivel local, es decir, en las subtablas de decisión obtenidas a partir de dichos birreductos. Como consecuencia, objetos inconsistentes en la tabla de decisión original pueden pertenecer a birreductos de decisión, aunque claramente estos deben ser distintos. Sin embargo, este hecho puede resultar un inconveniente, pues en ese caso se trabajaría con objetos cuya información aporta inconsistencia en la tabla original. Por tanto, una alternativa es no considerar estos objetos en ningún birreducto. Esta es la filosofía de los U -birreductos de decisión.

Definición 2.12. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $X \subseteq U$ y $B \subseteq \mathcal{A}$. Se dice que:

- *B es X -irreducible si no existe $B' \subset B$ tal que todos los pares de objetos $x \in X, y \in U$ d -discernibles son B' -discernibles.*
- *X es B -inextensible si no existe $X' \subset U$ con $X \subset X'$ tal que todos los pares de objetos $x \in X', y \in U$ d -discernibles son B -discernibles.*
- *El par (X, B) es un U -birreducto de decisión si todos los pares de objetos $x \in X, y \in U$ d -discernibles son B -discernibles, B es X -irreducible y X es B -inextensible.*

Se puede comprobar que solo los objetos consistentes en la tabla de decisión pueden pertenecer a un U -birreducto de decisión, debido a que cada objeto perteneciente a dicho U -birreducto es comparado con todos los objetos de la tabla de decisión. Este aspecto es el que marca la diferencia entre la Definición 2.11 y la Definición 2.12, lo que nos lleva a afirmar que el número de objetos en los U -birreductos de decisión es menor que en los birreductos de decisión.

Antes de finalizar esta sección, es conveniente remarcar que en esta tesis doctoral estudiaremos tanto reductos y reductos valor en las tablas de información, como d -reductos y d -reductos valor en las tablas de decisión, con respecto al conjunto de atributos completo $\mathcal{A}$, pues el principal objetivo es preservar toda la información relevante presente en las tablas correspondientes. Además, dado que todas estas nociones se centran en la reducción de información, mediante la combinación de reductos y reductos valor en las tablas de información, y de d -reductos y d -reductos valor en las tablas de decisión, es posible realizar numerosas reducciones diferentes. Este hecho proporciona una gran libertad al usuario encargado de los estudios a realizar, permitiendo elegir aquellos atributos y valores de atributos más convenientes para su análisis.

2.1.2. Reglas y algoritmos de decisión

Las reglas de decisión surgen para analizar las tablas de decisión en términos lógicos [73], lo cual facilita la interpretación de las tablas de decisión, describiendo las decisiones tomadas en términos de condiciones que deben satisfacerse. Antes de presentar la noción de regla de decisión, es necesario introducir el lenguaje formal que utilizarán.

Definición 2.13. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $B \subseteq \mathcal{A}_d$. 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 de los conectivos lógicos de conjunción $\wedge$ y disyunción $\vee$.

Para cada $\Phi \in For(B)$, con $\Phi = (a, v)$, se denotará $\|\Phi\|_S$ al conjunto de objetos que satisface Φ en S , esto es:

$$\|\Phi\|_S = \|(a, v)\|_S = \{x \in U \mid \bar{a}(x) = v\}$$

Inductivamente, dadas $\Phi, \Psi \in For(B)$, el conjunto de objetos que satisface la conjunción de fórmulas $\Phi \wedge \Psi$ y la disyunción de fórmulas $\Phi \vee \Psi$ en S se define como:

$$\begin{aligned} \|\Phi \wedge \Psi\|_S &= \|\Phi\|_S \cap \|\Psi\|_S \\ \|\Phi \vee \Psi\|_S &= \|\Phi\|_S \cup \|\Psi\|_S \end{aligned}$$

A continuación, se introduce la noción de regla de decisión. Como se ha destacado anteriormente, las reglas permiten extraer información de las tablas de decisión y analizarlas en gran detalle, siendo capaces de inferir conclusiones a partir de unas condiciones dadas. Por estos motivos, las reglas de decisión constituyen un pilar fundamental para la síntesis de la información contenida en las tablas de decisión.

Definición 2.14. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $B \subseteq \mathcal{A}$. Una regla de decisión en S es una expresión $\Phi \rightarrow \Psi$, donde $\Phi \in For(B)$ es el antecedente de la regla y $\Psi \in For(\{d\})$ es el consecuente de la regla.

Además, se dice que un objeto $x \in U$ satisface la regla de decisión $\Phi \rightarrow \Psi$ si $x \in \|\Phi \wedge \Psi\|_S$. De esta forma, se dice que el objeto x induce la regla de decisión $\Phi \rightarrow \Psi$, que puede denotarse como $\Phi_x \rightarrow \Psi_x$.

Para extraer conclusiones relevantes de la tabla de decisión es necesario analizar la información proporcionada por las reglas de decisión. Con este fin se utilizan los indicadores de relevancia, los cuales describen a las reglas de decisión desde distintos puntos de vista. En particular, el *soporte* proporciona el número de objetos que satisfacen la regla de decisión en cuestión, la *fuerza* determina la proporción de objetos que satisfacen dicha regla, es decir, la representatividad de dicha regla en la tabla, la *certeza* considera la proporción de objetos que satisfacen la regla respecto a los que satisfacen el antecedente de la misma, y la *cobertura* representa la proporción de objetos que satisfacen la regla respecto a los que satisfacen el consecuente. Estos indicadores se definen formalmente como sigue.

Definición 2.15. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $\Phi \rightarrow \Psi$ una regla de decisión en S . Se denomina:

- soporte (del inglés *support*) de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$supp_S(\Phi, \Psi) = \text{card}(\|\Phi \wedge \Psi\|_S)$$

donde $\text{card}(\cdot)$ denota el cardinal del conjunto considerado.

- fuerza (del inglés *strength*) de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$\sigma_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{\text{card}(U)}$$

- certeza (del inglés *certainty*) de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$cer_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{\text{card}(\|\Phi\|_S)}$$

cuando $\text{card}(\|\Phi\|_S) \neq 0$.

- cobertura (del inglés *coverage*) de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$cov_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{card(\|\Psi\|_S)}$$

cuando $card(\|\Psi\|_S) \neq 0$.

A partir de la noción de certeza, se dirá que $\Phi \rightarrow \Psi$ es una *regla de decisión cierta* si $cer_S(\Phi, \Psi) = 1$. En este caso, también se dice que la regla es *consistente*. Si $cer_S(\Phi, \Psi) = 0$, se dirá que es una *regla de decisión falsa*. Por el contrario, si $cer_S(\Phi, \Psi) \in (0, 1)$ se dirá que es una *regla de decisión no completamente cierta*.

Por otro lado, es importante que un conjunto de reglas de decisión extraídas a partir de una tabla de decisión proporcione la misma información que la contenida en dicha tabla, o al menos que preserve la consistencia y la información más relevante de los objetos de la tabla. Los algoritmos de decisión son conjuntos de reglas que persiguen este hecho exigiendo una serie de propiedades a cumplir por estas reglas.

Definición 2.16. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $Dec(S) = \{\Phi_i \rightarrow \Psi_i \mid i \in \{1, 2, \dots, m\}\}$, con $m \geq 2$ un conjunto de reglas de decisión en S . Se dice que:

1. $Dec(S)$ es un conjunto de reglas de decisión mutuamente excluyentes por parejas (independientes), si cada par de reglas de decisión $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisface que $\Phi = \Phi'$ o $\|\Phi \wedge \Phi'\|_S = \emptyset$, y $\Psi = \Psi'$ o $\|\Psi \wedge \Psi'\|_S = \emptyset$.
2. $Dec(S)$ cubre U si $\|\bigvee_{i=1}^m \Phi_i\|_S = \|\bigvee_{i=1}^m \Psi_i\|_S = U$.
3. La regla $\Phi \rightarrow \Psi \in Dec(S)$ es admisible en S si $supp_S(\Phi, \Psi) \neq 0$.
4. $Dec(S)$ preserva la consistencia de S si

$$POS_{\mathcal{A}}(\{d\}) = \bigvee_{\Phi \rightarrow \Psi \in Dec^+(S)} \|\Phi\|_S$$

donde $Dec^+(S)$ es el conjunto de reglas de decisión ciertas de $Dec(S)$.

A un conjunto de reglas de decisión $Dec(S)$ satisfaciendo las propiedades anteriores se le denomina algoritmo de decisión en S , y se denota como $DA(S)$.

Es fundamental destacar el papel que desempeña cada propiedad presentada en la Definición 2.16. La primera de ellas garantiza que no existan redundancias tanto en los antecedentes como en los consecuentes de las reglas de decisión del algoritmo, es decir, que dos antecedentes o dos consecuentes diferentes no sean satisfechos por un mismo objeto. La segunda propiedad asegura que todos los objetos de la tabla están representados en algún antecedente y en algún consecuente, por lo que no se pierde información de la tabla de decisión. La tercera propiedad exige que todas las reglas sean satisfechas por al menos un objeto, descartando así reglas que no aportan información relevante. Por último, la cuarta propiedad garantiza que las reglas del algoritmo preserven la consistencia presente en la tabla de decisión, la cual está proporcionada por la región positiva.

Por otro lado, dado que todas las reglas de un algoritmo de decisión son admisibles, de acuerdo a la Definición 2.14, todas ellas son inducidas por objetos pertenecientes a la tabla de decisión. En el caso de que cada objeto induzca alguna regla del algoritmo se obtendrán los denominados algoritmos de decisión completos.

Definición 2.17. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $DA(S)$ un algoritmo de decisión. Si para cada $x \in U$ existe $\Phi \rightarrow \Psi \in DA(S)$ tal que $x \in \|\Phi \wedge \Psi\|_S$ entonces se dirá que $DA(S)$ es un algoritmo de decisión completo.

Es importante destacar que el conjunto de reglas inducidas por cada objeto de la tabla de decisión dada, y en las que se consideren todos los atributos de condición de dicha tabla o un d -reducto, conforma un algoritmo de decisión completo. En efecto, respecto al primer punto de la Definición 2.16, si existieran dos antecedentes (respectivamente consecuentes)

que se satisfacen por un mismo objeto entonces, por la Definición 2.13 y teniendo en cuenta que se consideran los mismos atributos en todas las reglas de decisión, cada uno de ellos toma el mismo valor en ambos antecedentes (respectivamente consecuentes). Como consecuencia, ambos antecedentes (respectivamente consecuentes) son iguales, lo que implica que las reglas de decisión sean mutuamente excluyentes por parejas (independientes). Por otro lado, los puntos dos y tres son triviales al considerar las reglas de decisión inducidas por los objetos de la tabla de decisión. Finalmente, el cuarto punto también se verifica puesto que los objetos consistentes, dados por la región positiva, son aquellos que satisfacen los antecedentes de las reglas ciertas, ya que estas preservan la consistencia de la tabla de decisión al considerar todos los atributos de dicha tabla o los presentes en un d -reducto. En definitiva, cuando el tamaño de la tabla de decisión es reducido, es de gran interés tomar en consideración estos algoritmos, ya que proporcionan una información totalmente precisa y exhaustiva de la correspondiente tabla de decisión, al describirla mediante reglas inducidas por los objetos y usando todos los atributos, o simplemente los dados por un d -reducto.

2.2. Teoría de conjuntos rugosos difusos

En esta sección se introducirán las nociones más relevantes del marco de trabajo de FRST, con el propósito de estudiar tablas de decisión en este entorno. En primer lugar, se presenta la noción de cardinal de un conjunto difuso, que generaliza al cardinal de conjuntos clásicos [31].

Definición 2.18. *Dado un universo G y $H \subseteq G$ un conjunto difuso caracterizado por la función de pertenencia $h: G \rightarrow [0, 1]$, el cardinal de H se define como*

$$\text{card}_F(H) = \sum_{g \in G} h(g)$$

A continuación, se introduce un tipo particular de relación de tolerancia que se utilizará más adelante debido a su interesante propiedad [102].

Definición 2.19. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $a \in \mathcal{A}$ y $T_a: V_a \times V_a \rightarrow [0, 1]$ una relación de tolerancia. Si $T_a(v, w) = 1$ implica que $v = w$ entonces a T_a se le denomina relación de tolerancia separable.

Nótese que, toda relación de tolerancia T_a satisface que $T_a(v, v) = 1$ para todo $v \in V_a$. Como consecuencia, las relaciones de tolerancia separables aseguran que dos valores están completamente relacionados si y solo si son iguales. Teniendo en cuenta que este no es un requisito exigente, en algunos casos se necesitará considerar esta condición para obtener ciertas propiedades relevantes.

La siguiente noción permite comparar dos objetos mediante los valores que toman en un atributo.

Definición 2.20. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $a \in \mathcal{A}_d$. Una relación de a -indiscernibilidad es una $[0, 1]$ -relación de tolerancia difusa $R_a: U \times U \rightarrow [0, 1]$, es decir, una relación difusa reflexiva y simétrica.

A partir de esta definición, se puede afirmar que cuanto mayor sea el valor dado por R_a , más similares serán los objetos comparados de acuerdo al atributo a . La siguiente definición extiende la anterior a un subconjunto de atributos arbitrario.

Definición 2.21. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}$ y $R_a: U \times U \rightarrow [0, 1]$ una relación de a -indiscernibilidad para todo $a \in \mathcal{A}$. La relación de B -indiscernibilidad $R_B: U \times U \rightarrow [0, 1]$ se define para cada par de objetos $x, y \in U$ como

$$R_B(x, y) = @(\mathcal{R}_B^{x,y}(a_1), \dots, \mathcal{R}_B^{x,y}(a_m))$$

donde $\mathcal{R}_B^{x,y}: \mathcal{A} \rightarrow [0, 1]$ se define para cada $a \in \mathcal{A}$ como

$$\mathcal{R}_B^{x,y}(a) = \begin{cases} R_a(x, y) & \text{si } a \in B \\ 1 & \text{en caso contrario} \end{cases}$$

$y @: [0, 1]^m \rightarrow [0, 1]$ es un operador agregador, es decir, un operador creciente en cada argumento satisfaciendo $@(1, \dots, 1) = 1$ y $@(0, \dots, 0) = 0$.

Nótese que R_B es una $[0, 1]$ -relación de tolerancia difusa para todo $B \subseteq \mathcal{A}$. Además, la interpretación del valor dado por R_B es análoga a la dada por R_a en la Definición 2.20.

En lo que sigue se introducen algunas nociones necesarias para analizar tablas de decisión en un marco de trabajo multiadjunto, las cuales fueron dadas en [29]. Este entorno proporciona una gran flexibilidad, otorgando mucha libertad al usuario encargado de realizar los estudios correspondientes de acuerdo a sus necesidades. En primer lugar, introducimos la noción de triple adjunto.

Definición 2.22. Sean $(P_1, \leq_1)$, $(P_2, \leq_2)$ y $(P_3, \leq_3)$ conjuntos parcialmente ordenados y las aplicaciones $\&: P_1 \times P_2 \rightarrow P_3$, $\swarrow: P_3 \times P_2 \rightarrow P_1$ y $\nwarrow: P_3 \times P_1 \rightarrow P_2$. Entonces $(\&, \swarrow, \nwarrow)$ es un triple adjunto con respecto a P_1, P_2 y P_3 si:

- $\&$ es creciente en ambos argumentos.
- $\swarrow$ y $\nwarrow$ son crecientes en el primer argumento y decrecientes en el segundo.
- $\&, \swarrow, \nwarrow$ satisfacen la propiedad de adjunción, es decir,

$$x \leq_1 z \swarrow y \quad \text{si y solo si} \quad x \& y \leq_3 z \quad \text{si y solo si} \quad y \leq_2 z \nwarrow x$$

para todo $x \in P_1, y \in P_2, z \in P_3$.

La aplicación $\&$ es el conjuntor y $\swarrow$ (respectivamente $\nwarrow$) la implicación residuada por la derecha (respectivamente por la izquierda) de $\&$.

Otra noción de especial relevancia en el marco multiadjunto es la de par adjunto [25], la cual es similar a la noción de triple adjunto, pero considerando solo dos aplicaciones.

Definición 2.23. Sean $(P_1, \leq_1)$, $(P_2, \leq_2)$ y $(P_3, \leq_3)$ conjuntos parcialmente ordenados y las aplicaciones $\&: P_1 \times P_2 \rightarrow P_3$ y $I: P_2 \times P_3 \rightarrow P_1$. Entonces $(\&, I)$ es un par adjunto con respecto a P_1, P_2 y P_3 si:

- $\&$ es creciente en ambos argumentos.
- I es creciente en el segundo argumento y decreciente en el primero.
- $\&, I$ satisfacen que

$$\&(x, y) \leq_3 z \quad \text{si y solo si} \quad x \leq_1 I(y, z)$$

para todo $x \in P_1, y \in P_2, z \in P_3$.

Cabe destacar que, se ha intercambiado el orden de los argumentos tomados por la aplicación I con respecto al orden en las aplicaciones $\swarrow$ y $\nwarrow$ de la Definición 2.22. Intercambiando nuevamente el orden de dichos argumentos, se deduce que si $(\&, \swarrow, \nwarrow)$ es un triple adjunto con respecto a P_1, P_2 y P_3 , entonces $(\&, \swarrow)$ es un par adjunto con respecto a P_1, P_2 y P_3 . Además, si el conjuntor $\&$ es conmutativo, entonces $(\&, \nwarrow)$ también es un par adjunto con respecto a P_1, P_2 y P_3 . En caso de ambigüedad, para diferenciar entre triples y pares adjuntos se usará una notación infija cuando consideremos triples adjuntos y una notación prefija para los pares adjuntos, como se evidencia en las Definiciones 2.22 y 2.23, respectivamente.

Ahora se presenta la noción de marco multiadjunto orientado a propiedades.

Definición 2.24. Dado un conjunto parcialmente ordenado $(P_1, \leq_1)$, dos retículos completos $(L_2, \leq_2)$ y $(L_3, \leq_3)$, y triples adjuntos $(\&_i, \swarrow^i, \nwarrow_i)$, con respecto a P_1, L_2, L_3 , para todo $i \in \{1, \dots, n\}$, la tupla $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ es un marco multiadjunto orientado a propiedades.

Los marcos multiadjuntos permiten considerar diferentes triples adjuntos en función del estudio a realizar. De esta forma, el usuario puede

seleccionar un triple adjunto para cada par objeto atributo, dependiendo de las características de estos con respecto al resto. Por ejemplo, esta asignación se puede realizar para determinar distintos grados de preferencia entre los objetos o atributos, como se observó en [62, 65]. A continuación, se introducen las nociones de aproximación superior e inferior en el marco de FRST. En dicha definición, dados dos conjuntos Y y Z , se denota como Y^Z al conjunto de aplicaciones de Z a Y , es decir, $Y^Z = \{h \mid h: Z \rightarrow Y\}$.

Definición 2.25. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}$, R_B una relación de B -indiscernibilidad, $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ un marco multiadjunto orientado a propiedades y $\tau: U \times U \rightarrow \{1, \dots, n\}$. Dados $g \in L_2^U$ y $f \in L_3^U$, se definen las aproximaciones superior $\uparrow^\pi: L_2^U \rightarrow L_3^U$ e inferior $\downarrow^N: L_3^U \rightarrow L_2^U$, respectivamente, como:

$$\begin{aligned} g^{\uparrow^\pi}(x) &= \sup\{R_B(x, y) \&_{\tau(x, y)} g(y) \mid y \in U\} \\ f^{\downarrow^N}(y) &= \inf\{f(x) \frown_{\tau(x, y)} R_B(x, y) \mid x \in U\} \end{aligned}$$

Así, $g^{\uparrow^\pi}$ es la aproximación superior de g y $f^{\downarrow^N}$ es la aproximación inferior de f .

Nótese que, en el caso general de los marcos multiadjuntos orientados a propiedades y objetos, estos operadores son un caso particular de los operadores posibilidad y necesidad [29]. Una vez establecidos el marco multiadjunto, es posible estudiar tablas de decisión en este entorno. Por otro lado, dado que la presente tesis se centrará en tablas de decisión cuyo atributo de decisión es booleano, se introduce la noción de región positiva en el entorno multiadjunto cuando R_d solo toma valores booleanos, la cual se basa en la noción de aproximación inferior.

Definición 2.26. Sean $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $B \subseteq \mathcal{A}$. Dado un marco multiadjunto orientado a propiedades $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$, $\tau: U \times U \rightarrow \{1, \dots, n\}$, R_B una relación de B -indiscernibilidad y R_d una relación de d -discernibilidad booleana, la región B -positiva difusa multiadjunta

se define, para cada $y \in U$, como:

$$POS_B^f(y) = (R_d y)^{\downarrow^N}(y) = \inf\{(R_d y)(x) \frown_{\tau(x,y)} R_B(x, y) \mid x \in U\}$$

donde $R_d y: U \rightarrow \{0, 1\}$ se define como $(R_d y)(x) = R_d(y, x)$ y $\frown_{\tau(x,y)}$ es la implicación difusa residuada por la izquierda de $\&_{\tau(x,y)}$ asociada con el par de objetos x, y .

Es importante destacar que el valor dado por la región B -positiva difusa multiadjunta muestra la dependencia del atributo de decisión d con respecto a los atributos de condición pertenecientes a B en el objeto evaluado. De esta forma, dado un objeto $y \in U$, si $POS_B^f(y)$ es próximo a 1, entonces los valores de los atributos de B determinan en gran medida el valor del atributo d en el objeto y . En consecuencia, el objeto y tendrá una gran consistencia en la tabla de decisión y la información dada por dicho objeto será de gran interés. Además, nótese que la Definición 2.26 generaliza a la definición clásica de región positiva (Definición 2.8), tal y como se aseguró en [29] y probó en [20].

Al igual que en el entorno clásico, la reducción de información es de vital importancia para simplificar su análisis en el marco de trabajo multiadjunto. Antes de presentar la noción de reducto en este entorno, es fundamental introducir la siguiente definición.

Definición 2.27. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}$ y $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$ un marco multiadjunto orientado a propiedades. Una aplicación monótona $m: \mathcal{P}(\mathcal{A}) \rightarrow [0, 1]$ es una medida en S si la condición

$$R_d(x, y) \frown_{\tau(x,y)} R_B(x, y) = R_d(x, y) \frown_{\tau(x,y)} R_{\mathcal{A}}(x, y), \text{ para todo } x, y \in U$$

implica que $m(B) = 1$.

La medida presentada en la definición anterior es esencial para calcular reductos difusos, pues permite calcular el grado de representatividad

del subconjunto de atributos analizado en la tabla de decisión, con respecto al atributo de decisión. La noción de reducto en el marco de trabajo multiadjunto es presentada a continuación.

Definición 2.28. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}$, m una medida en S y $\alpha \in (0, 1]$. Al conjunto B se le denomina **superreducto difuso de m -decisión de grado α** si $\alpha \leq m(B)$. Además, si $\alpha \not\leq m(B')$ para cada $B' \subset B$ entonces a B se le denomina **reducto difuso de m -decisión de grado α** .

Para eliminar atributos que no aporten información relevante en el entorno multiadjunto es necesario disponer de reductos difusos de m -decisión de grado 1, pudiendo descartar los atributos no pertenecientes a dicho reducto. No obstante, obtener reductos de grado 1 en un marco de trabajo difuso es realmente complejo, dada la gran variedad de nociones que intervienen en su cálculo. Como consecuencia, el usuario encargado del estudio deberá establecer un umbral de forma que, si el grado del reducto considerado supera el umbral dado, entonces se podrá eliminar los atributos no pertenecientes a dicho reducto.

Capítulo 3

Discusión conjunta de los resultados

En este capítulo se muestran los resultados alcanzados tras la investigación realizada durante el desarrollo de la tesis doctoral. Cada sección se focaliza en uno de los tres artículos que se han seleccionado para la presente tesis. En primer lugar, surgió la necesidad de estudiar las tablas de información y decisión en el marco de trabajo de RST. En particular, nos centramos en estudiar y comparar diferentes nociones relacionadas con la reducción de información. En la Sección 3.1 se expone la relación obtenida en [18] entre algunas nociones de RST definidas originariamente de manera independiente, como son los reductos, reductos valor y birreductos. Posteriormente, dada la complejidad de los conjuntos de datos reales, llevamos a cabo un estudio de dichos datos desde el punto de vista difuso y multiadjunto. La Sección 3.2 se centra en los avances obtenidos en [20], donde se han introducido numerosas nociones relacionadas con las reglas de decisión en el entorno de FRST, las cuales generalizan a las correspondientes nociones de RST y permiten el estudio de bases de datos relacionales bajo el marco de trabajo difuso. Por último, realizamos el

estudio de la toma de decisiones desde el punto de vista de FRST, dada la importancia de esta temática en numerosos campos como se destacó en la introducción. En la Sección 3.3 se muestran los métodos definidos en [15] para la clasificación de nuevos datos. Para llevar a cabo dicha clasificación es imprescindible realizar un análisis previo de la información disponible. En los artículos citados se destaca la importancia de cada una de las contribuciones, ya sea por la aportación de resultados novedosos, o por la generalización y mejora de nociones y resultados previos.

3.1. Reductos, reductos valor y birreductos

Esta sección incluye la discusión de los resultados obtenidos en [18], donde se llevó a cabo un estudio comparativo de las nociones de reducto, reducto valor y birreducto. En dichas contribuciones, se muestra la existencia de diferentes relaciones entre ellas, permitiendo la obtención de unos a partir de otros y viceversa.

Por un lado, los reductos y reductos valor fueron introducidos por Pawlak [70, 71, 72] en la década de los ochenta para reducir el tamaño de las bases de datos sin perder información relevante. En particular, los reductos se centran en la reducción del conjunto de atributos para todos los objetos manteniendo la discernibilidad entre objetos, mientras que los reductos valor se enfocan en la reducción de atributos en cada objeto de manera independiente, proporcionando un mayor nivel de simplificación de la información. Por otro lado, Janusz et al. presentaron en 2012 la noción de birreducto [50], cuyo objetivo es la reducción simultánea de los conjuntos de objetos y de atributos. Dicha reducción se basa en la eliminación de redundancias o inconsistencias presentes en la base de datos original, permitiendo la obtención de subtablas sin información redundante ni inconsistente.

Existen algunos trabajos en los que se relacionan las nociones de reducto y birreducto. Por ejemplo, en [8] se expone la obtención de reductos a partir de birreductos de información, mientras que en [89] se caracterizan birreductos de decisión en función de d -reductos. Sin embargo, se pueden obtener más relaciones entre las nociones de reducto y birreducto. Además, en la literatura no se encuentran estudios relacionando las nociones de reducto valor y birreducto, los cuales son incorporados en [17, 18] y se destacarán en esta sección.

En lo que sigue, se fija una tabla de información $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ y una tabla de decisión $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$. Como se comentó anteriormente, por medio de los birreductos de información es posible obtener subtablas de información sin redundancias, es decir, todos los objetos de dichas subtablas son discernibles entre sí. De esta forma, los birreductos de información recogen un subconjunto minimal de objetos que permite describir de manera exhaustiva la tabla de información de acuerdo al subconjunto de atributos de dicho birreducto. Además, fijado un subconjunto de atributos $B \subseteq \mathcal{A}$, los birreductos con B como subconjunto de atributos contienen la información de la tabla original con respecto a B , puesto que todos los objetos pertenecen a alguno de estos birreductos, como pone de manifiesto el siguiente resultado.

Proposición 3.1. *Sean $B \subseteq \mathcal{A}$ un subconjunto de atributos y $X_1, \dots, X_n \subseteq U$ subconjuntos de objetos. Si los pares $(X_1, B), \dots, (X_n, B)$ son todos los birreductos de información con B como subconjunto de atributos, entonces*

$$U = \bigcup_{i=1}^n X_i$$

Demostración. Véase la demostración de la Proposición 12 en [18]. $\square$

De aquí, se puede recuperar la información relevante de la tabla original a partir de los birreductos de información. Por ejemplo, los atributos

dispensables no se recuperarían. La introducción de la Proposición 3.1 es de especial relevancia para los objetos B -indiscernibles, ya que por la Definición 2.6, los objetos B -discernibles en la tabla original pertenecen a todos los birreductos de información con B como subconjunto de atributos.

En la siguiente sección se mostrará las relaciones obtenidas entre los reductos y birreductos de información en el marco de las tablas de información, y la relación entre los d -reductos y, tanto los birreductos de decisión como los U -birreductos de decisión, en el entorno de las tablas de decisión.

3.1.1. Relación entre reductos y birreductos

En esta sección se tendrán en cuenta los reductos del conjunto de atributos completo $\mathcal{A}$ (Definición 2.4), por lo que también nos referiremos a ellos como reductos, sin hacer referencia al conjunto $\mathcal{A}$. En primer lugar, se presenta el estudio llevado a cabo en tablas de información. La obtención de birreductos de información a partir de reductos ya fue estudiada en [8]. Sin embargo, el siguiente resultado propone la situación inversa, es decir, la obtención de birreductos a partir de un reducto conocido.

Proposición 3.2. *Dado $B \subseteq \mathcal{A}$, si B es un reducto de $\mathcal{A}$ entonces existe $X \subseteq U$ tal que (X, B) es un birreducto de información.*

Demostración. Se halla en la demostración de la Proposición 16 en [18]. $\square$

Como consecuencia, dado un reducto es posible construir una subtabla cuyos objetos son discernibles mediante los atributos del reducto. Además, puesto que un reducto es un subconjunto minimal de atributos que preserva la información de la tabla original, la subtabla mencionada contendrá dicha información de manera minimal en atributos y en objetos no

redundantes. Si el reducto es el conjunto de atributos completo $\mathcal{A}$ se deduce una doble equivalencia, puesto que, si $(X, \mathcal{A})$ es un birreducto de información, entonces $\mathcal{A}$ es irreducible para preservar la $\mathcal{A}$ -indiscernibilidad de los objetos de la tabla. Como resultado, se deduce inmediatamente que $\mathcal{A}$ es reducto.

Corolario 3.1. *Existe $X \subseteq U$ tal que $(X, \mathcal{A})$ es un birreducto de información si y solo si $\mathcal{A}$ es reducto de $\mathcal{A}$.*

Demostración. Se encuentra en la demostración del Corolario 17 en [18]. $\square$

Por otro lado, la Definición 2.6 presenta un aspecto a resaltar. Si existen dos objetos $\mathcal{A}$ -indiscernibles entonces para todo $B \subseteq \mathcal{A}$ también serán B -indiscernibles. Es por ello que dichos objetos no pueden pertenecer a un mismo birreducto de información, haciendo necesario el cálculo de dos birreductos diferentes para tener en consideración a ambos objetos. Sin embargo, consideramos innecesario el cálculo de ambos birreductos, puesto que ambos objetos contienen la misma información de partida. Para resolver esta situación, se propone agrupar todos los objetos $\mathcal{A}$ -indiscernibles en clases de equivalencia y trabajar en este marco, obteniendo las denominadas tablas de información reducidas por clases.

Definición 3.1. *Sea $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ una tabla de información. Una tabla de información reducida por clases es una tupla $(U^*, \mathcal{A}_s, \mathcal{V}_{\mathcal{A}_s}, \overline{\mathcal{A}}_s)$ tal que $U^* = \{[x]_{I(\mathcal{A})} \mid x \in U\}$, $\mathcal{A}_s = \mathcal{A} \cup \{s\}$ con $s \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_s} = \{V_a \mid a \in \mathcal{A}_s\}$ siendo V_a el conjunto de valores asociados al atributo a sobre U^* , y $\overline{\mathcal{A}}_s$ es el conjunto que agrupa las aplicaciones asociadas a cada $a \in \mathcal{A}$ definidas como*

$$\begin{aligned} \overline{a}^*: \quad U^* &\longrightarrow V_a \\ [x]_{I(\mathcal{A})} &\longmapsto \overline{a}(x) \end{aligned}$$

y la aplicación asociada con el atributo s definida como

$$\begin{aligned} \overline{s}: \quad U^* &\longrightarrow V_s \\ [x]_{I(\mathcal{A})} &\longmapsto \text{card}([x]_{I(\mathcal{A})}) \end{aligned}$$

De manera inmediata se comprueba que las aplicaciones $\bar{a}^*$ para cada $a \in \mathcal{A}$ y la aplicación $\bar{s}$ están bien definidas. Como se observa, se agrupan todos los objetos $\mathcal{A}$ -indiscernibles en una misma clase y se toma un registro de cuántos objetos pertenecen a cada clase de equivalencia. Con esta reinterpretación de las tablas de información, además de solventar el hecho descrito anteriormente, es posible obtener el siguiente resultado.

Proposición 3.3. *Sean $(U^*, \mathcal{A}_s, \mathcal{V}_{\mathcal{A}_s}, \overline{\mathcal{A}_s})$ una tabla de información reducida por clases y $B \subseteq \mathcal{A}$. El par (U^*, B) es un birreducto de información si y solo si B es un reducto de $\mathcal{A}$.*

Demostración. Véase la demostración de la Proposición 20 en [18]. $\square$

En definitiva, en esta ocasión se obtiene una doble equivalencia entre reductos y birreductos de información considerando un subconjunto de atributos B distinto de $\mathcal{A}$.

A continuación se presenta los resultados obtenidos para tablas de decisión. Es importante señalar que no todos los resultados previos de tablas de información son extensibles a las tablas de decisión, como es el caso de la Proposición 3.1, véase el Ejemplo 23 en [18]. El siguiente resultado aporta una condición necesaria para obtener un birreducto de decisión.

Proposición 3.4. *Sean $X \subseteq U$ y $B \subseteq \mathcal{A}$. Si el par (X, B) es un birreducto de decisión, entonces $POS_B(\{d\}) \subseteq X$.*

Demostración. Consúltese la demostración de la Proposición 24 en [18]. $\square$

Por consiguiente, la región positiva facilita el cálculo de los birreductos de decisión proporcionando un punto de partida para obtenerlos. Así, el conjunto de objetos formado por la región positiva debe ser completado mientras se satisfacen las propiedades de la Definición 2.11 para obtener un birreducto de decisión. El siguiente resultado permite la obtención directa de un d -reducto a partir de un birreducto de decisión.

Proposición 3.5. *Sea $B \subseteq \mathcal{A}$ un subconjunto de atributos. Si el par (U, B) es un birreducto de decisión, entonces B es un d -reducto de $\mathcal{A}$.*

Demostración. Se halla en la demostración de la Proposición 25 en [18]. $\square$

De esta forma, la subtabla asociada al birreducto de decisión (U, B) no presenta inconsistencias, al igual que la tabla de decisión original ya que todos los objetos d -discernibles son B -discernibles. Además, contiene la misma información que la tabla de partida, siendo así un marco de trabajo ideal para su análisis.

Por otro lado, en el Ejemplo 26 en [18] se muestra que el recíproco de la Proposición 3.5 no es cierto en general, y que no es posible extender la Proposición 3.2 ni el Corolario 3.1 a tablas de decisión. Además, la consideración de tablas de decisión reducidas por clases no da lugar al cálculo de birreductos de decisión, ya que estas tablas tienen como objetivo la eliminación de redundancias, mientras que los birreductos de decisión permiten que objetos redundantes, es decir, objetos B -indiscernibles y d -indiscernibles, pertenezcan a un mismo birreducto. Por otro lado, imponiendo que $\text{POS}_{\mathcal{A}}(\{d\}) = U$, tal y como se muestra en [89], existe una total equivalencia entre d -reductos y birreductos de decisión. Sin embargo, esta condición sobre la región positiva es bastante restrictiva, especialmente cuando se trata de conjuntos de datos reales. Por consiguiente, se concluye que en el ámbito de las tablas de decisión la relación entre los d -reductos y los birreductos de decisión es más débil.

Por último, se exponen las distintas relaciones obtenidas entre los d -reductos y los U -birreductos de decisión. En primer lugar, se presenta un resultado que caracteriza a los objetos de los U -birreductos de decisión.

Proposición 3.6. *Dado (X, B) un U -birreducto de decisión, se satisface la siguiente propiedad:*

$$x \in X \quad \text{si y solo si} \quad [x]_{I(B)} \subseteq [x]_{I(\{d\})}$$

Demostración. Véase la demostración de la Proposición 32 en [18]. $\square$

Cabe destacar que, mediante la Proposición 3.6, se puede concluir que, para cada $B \subseteq \mathcal{A}$ existe a lo sumo un U -birreducto de decisión con B como subconjunto de atributos, a diferencia de los birreductos de información y decisión. Por ello, solo se obtiene una única subtabla de decisión, constituida por aquellos objetos consistentes de acuerdo al conjunto de atributos B , es decir, o bien objetos d -indiscernibles o bien B -discernibles y d -discernibles. Además, por la Definición 2.8 y la Proposición 3.6, el subconjunto de objetos de los U -birreductos de decisión es exactamente la región positiva con respecto al subconjunto de atributos de dicho U -birreducto, lo cual es un hecho más específico que la conclusión dada en la Proposición 3.4 y facilita sustancialmente la obtención de U -birreductos de decisión.

Puesto que la Definición 2.12 es más general que la Definición 2.11, al comparar cada objeto del birreducto con todos los presentes en la tabla, se han obtenido más resultados relacionando d -reductos y U -birreductos de decisión. A continuación, se presenta el resultado equivalente a la Proposición 3.5.

Proposición 3.7. *Dado $B \subseteq \mathcal{A}$, si el par (U, B) es un U -birreducto de decisión, entonces B es un d -reducto de $\mathcal{A}$.*

Demostración. Se obtiene inmediatamente de la Proposición 3.5, ya que es fácil ver que (U, B) es un birreducto de decisión si y solo si (U, B) es un U -birreducto de decisión. $\square$

Es necesario mencionar que el recíproco de la Proposición 3.7 es falso en general, y se debe exigir que $\text{POS}_{\mathcal{A}}(\{d\}) = U$ para su verificación, como se muestra en [89]. A pesar de esto, es posible encontrar una relación equivalente a las presentadas en la Proposición 3.2 y en el Corolario 3.1 entre ambas nociones.

Proposición 3.8. *Dado $B \subseteq \mathcal{A}$, si B es un d -reducto de $\mathcal{A}$, entonces existe $X \subseteq U$ tal que (X, B) es un U -birreducto de decisión.*

Demostración. Se deduce de la Proposición 3.6. □

Corolario 3.2. *Existe $X \subseteq U$ tal que $(X, \mathcal{A})$ es un U -birreducto de decisión si y solo si $\mathcal{A}$ es un d -reducto de $\mathcal{A}$.*

Demostración. Se concluye a partir de la Proposición 3.6. □

Conviene destacar que la subtabla dada por el birreducto obtenido por la Proposición 3.8 es consistente. De aquí, en dicha subtabla se verifica el recíproco de la Proposición 3.7.

En resumen, se han aportado resultados que relacionan los reductos y los birreductos tanto en tablas de información como de decisión. De esta forma, el usuario puede calcular o bien reductos o bien birreductos en función de sus necesidades, y extraer los otros de manera prácticamente inmediata. En la siguiente sección se muestran los resultados obtenidos relacionando los reductos valor y los birreductos.

3.1.2. Relación entre reductos valor y birreductos

Como se expuso anteriormente, los reductos valor complementan a los reductos proporcionando una mayor simplificación de la información relevante, dado que esta reducción es realizada objeto a objeto y se puede llevar a cabo tras el cálculo de los reductos. Sin embargo, estos no han sido estudiados con tanto detalle como los reductos a pesar de sus ventajas relevantes para el análisis de tablas de información y decisión. Por ello, consideramos imprescindible compararlos con los birreductos. Dado que todos los resultados a presentar hacen referencia a reductos valor de $\mathcal{A}$, nos referiremos a ellos también como reductos valor, sin hacer referencia al conjunto $\mathcal{A}$. Al igual que en la Sección 3.1.1, se comienza mostrando

el estudio realizado cuando se consideran tablas de información. El primer resultado facilita el cálculo de birreductos de información a partir de reductos valor conocidos.

Proposición 3.9. *Sean $x \in U$ y $B \subseteq \mathcal{A}$ un reducto valor de $\mathcal{A}$ para x . Entonces, para cada $y_0 \in [x]_{I(\mathcal{A})}$ existe un birreducto de información (X_{y_0}, B) tal que $y_0 \in X_{y_0}$. De hecho, todos los birreductos de información (X, B) satisfacen que $[x]_{I(\mathcal{A})} \cap X \neq \emptyset$.*

Demostración. Véase la demostración de la Proposición 39 en [18]. □

Es conveniente destacar dos consecuencias importantes derivadas de la Proposición 3.9. En primer lugar, asegura la existencia de birreductos de información cuyo subconjunto de atributos sea un reducto valor. De aquí, para el cálculo de los mismos solo es necesario comprobar que en el subconjunto de objetos se verifican las condiciones de discernibilidad e inestabilidad de la Definición 2.6. En segundo lugar, garantiza que algún objeto $\mathcal{A}$ -indiscernible a x pertenezca a dicho birreducto, por lo que proporciona información relevante para su cálculo. Por último, cabe mencionar que si además de las hipótesis anteriores se verifica que $[x]_{I(\mathcal{A})} = \{x\}$, entonces el objeto x pertenece a todos los birreductos de información cuyo subconjunto de atributos es su reducto valor, como prueba el Corolario 40 en [18].

Por otro lado, para la obtención de reductos valor a partir de birreductos de información dados es necesario debilitar la hipótesis sobre el reducto valor dada en la Proposición 3.9. No obstante, este hecho permite la extracción de una doble equivalencia entre ellos, tal y como se manifiesta en el siguiente resultado.

Proposición 3.10. *Sean $B \subseteq \mathcal{A}$, $X_1, \dots, X_n \subseteq U$ y $x \in U$ con $(X_1, B), \dots, (X_n, B)$ todos los birreductos de información con B como subconjunto de atributos. Entonces, para cada $i \in \{1, \dots, n\}$ existe $y_i \in [x]_{I(\mathcal{A})}$ tal que $y_i \in X_i$ si y solo si existe $B' \subseteq B$ tal que B' es un reducto valor de $\mathcal{A}$ para x .*

Demostración. Se encuentra en la prueba de la Proposición 42 en [18]. $\square$

Por tanto, conocidos todos los birreductos de información con un subconjunto de atributos dado, es posible determinar la existencia de reductos valor para los objetos pertenecientes a los birreductos correspondientes. Además, el subconjunto de atributos de estos birreductos constituye un punto de partida para el cálculo de los reductos valor, de manera que solo es necesario estudiar su irreducibilidad para el objeto considerado. Por otro lado, si $[x]_{I(\mathcal{A})} = \{x\}$, entonces el objeto x pertenece a todos los birreductos de información con B como subconjunto de atributos, si y solo si existe un reducto valor B' de $\mathcal{A}$ para x con $B' \subseteq B$, como expone el Corolario 43 en [18]. Además, el Ejemplo 44 en [18] refleja que esta condición de contención es imprescindible para que la Proposición 3.10 se verifique, puesto que no es una igualdad en general.

Una vez expuesto el estudio llevado a cabo en las tablas de información, se muestra la relación obtenida entre los d -reductos valor y birreductos de decisión.

Proposición 3.11. *Sea $B \subseteq \mathcal{A}$ un d -reducto valor de $\mathcal{A}$ para $x \in U$ tal que $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Entonces, existe un birreducto de decisión (X_0, B) tal que $x \in X_0$. De hecho, todos los birreductos de decisión (X, B) satisfacen que $[x]_{I(B)} \subseteq X$.*

Demostración. Se halla en la demostración de la Proposición 46 en [18]. $\square$

Obsérvese que, por la Definición 2.10, la condición $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$ es necesaria para que exista algún d -reducto valor para el objeto x . Sin embargo, este requisito no es demasiado exigente, ya que en general es difícil que un par de objetos sean $\mathcal{A}$ -indiscernibles y d -discernibles, pues se tienen en consideración todos los atributos de la tabla de decisión para su comparativa.

Por otro lado, y al igual que en la Proposición 3.9, se garantiza la existencia de birreductos de decisión que tienen como subconjunto de atribu-

tos al d -reducto valor dado. En consecuencia, a partir de los d -reductos valor es posible obtener subtablas de decisión consistentes, a las que pertenecerán todos los objetos indiscernibles al objeto dado de acuerdo al d -reducto valor considerado. No obstante, debemos destacar que el resultado equivalente a la Proposición 3.10 no es cierto en las tablas de decisión, como se detalla después del Ejemplo 47 en [18]. Como resultado, el conocimiento de un conjunto de birreductos de decisión no asegura la obtención de d -reductos valor a partir de ellos.

Finalmente, se presentan los resultados que estudian los d -reductos valor y los U -birreductos de decisión. El primero de ellos es el equivalente a la Proposición 3.11 para U -birreductos de decisión.

Proposición 3.12. *Sea $B \subseteq \mathcal{A}$ un d -reducto valor de $\mathcal{A}$ para $x \in U$ tal que $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Entonces, existe $X \subseteq U$, tal que el par (X, B) es un U -birreducto de decisión satisfaciendo que $[x]_{I(B)} \subseteq X$.*

Demostración. Consúltese la demostración de la Proposición 49 en [18]. $\square$

Las conclusiones derivadas de la Proposición 3.12 son muy similares a las de la Proposición 3.11, con la particularidad de que el U -birreducto de decisión a obtener es único. Por lo tanto, cada d -reducto valor está asociado a un único U -birreducto de decisión, el cual permite la extracción de toda la información consistente de acuerdo a los atributos de dicho d -reducto valor en una única subtabla de decisión.

Por último, es reseñable destacar que, dado un U -birreducto de decisión, es posible conocer la existencia de d -reductos valor.

Proposición 3.13. *Sean (X, B) un U -birreducto de decisión y $x \in U$ tal que $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Entonces, $[x]_{I(B)} \subseteq X$ si y solo si existe $B' \subseteq B$ tal que B' es un d -reducto valor de $\mathcal{A}$ para x .*

Demostración. Se encuentra en la prueba de la Proposición 50 en [18]. $\square$

En consecuencia, y teniendo en cuenta la Proposición 3.6, para todo objeto perteneciente al U -birreducto de decisión es posible obtener un d -reducto valor. Además, el cálculo de este se simplifica notablemente, ya que el subconjunto de atributos del U -birreducto supone un punto de partida. De esta forma, se deduce una mayor relación entre los d -reductos valor y los U -birreductos de decisión que con los birreductos de decisión, como ocurrió con los d -reductos en la Sección 3.1.1. Además, tal y como ocurre en la Proposición 3.10, la condición $B' \subseteq B$ es fundamental para que se satisfaga la Proposición 3.13, como se muestra en el Ejemplo 51 en [18].

Se concluye que todos los resultados presentados en esta sección son de interés para el análisis de información en conjuntos de datos desde un punto de vista clásico, puesto que se centran en nociones importantes para la reducción de información. En particular, estos resultados aportan numerosas relaciones entre los reductos, reductos valor y birreductos tanto en tablas de información como de decisión, las cuales permiten la obtención de unos a partir de otros y viceversa, a pesar de que todas estas nociones fueron introducidas de manera independiente.

3.2. Reglas y algoritmos de decisión en teoría de conjuntos rugosos difusos

Las reglas de decisión en el entorno de RST fueron introducidas por Pawlak [72, 73] para manejar la información contenida en tablas de decisión en términos lógicos. De esta forma, el análisis de las bases de datos se puede sistematizar mejor, por lo que las reglas de decisión facilitan la obtención de información y la extracción de conclusiones. Para describir a las reglas, Pawlak definió algunos indicadores de relevancia. Estos muestran la importancia de las reglas de decisión, permitiendo al usuario seleccio-

nar aquellas más significativas. Además, también introdujo los algoritmos de decisión, que son un conjunto de reglas satisfaciendo ciertas propiedades que garantizan que no se pierda información relevante al describir la tabla de decisión considerada por medio de las reglas de decisión de dicho algoritmo.

Como se comentó en el Capítulo 1, el marco de trabajo de RST puede ser insuficiente para el análisis de la información contenida en conjuntos de datos reales. Esto se debe a la necesidad de considerar algún grado de flexibilidad para el manejo de dicha información, ya sea por poseer imprecisión o imperfección en los datos de partida, o por ser descrita mediante cantidades numéricas, cuyo manejo como valores booleanos podría provocar una pérdida de información o introducir un nivel de complejidad innecesario. Aunque en algunos trabajos el análisis de los datos se aborda desde un punto de vista clásico mediante una discretización de los datos, como en [76], es más natural y óptimo realizarlo en un marco de trabajo difuso, aprovechando las ventajas del mismo. En esta sección, se presenta la extensión de numerosas nociones de RST [72, 73] al marco de trabajo difuso, lo cual fue realizado en [20]. Estas nociones se definen interpretando la base de datos como una tabla de decisión, puesto que el estudio girará en torno a las reglas de decisión, que son el punto de partida para clasificar nuevos objetos, como veremos en la Sección 3.3.

3.2.1. Reglas de decisión

Al igual que en el entorno clásico, en el marco de trabajo difuso las reglas de decisión también se apoyan en la noción de fórmula. Aunque estas se basan en pares atributo valor como en la Definición 2.13, su significado es diferente debido al cambio de entorno.

Definición 3.2. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $B \subseteq \mathcal{A}_d$. 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 de los conectivos lógicos de conjunción $\wedge$ y disyunción $\vee$.

Dada una familia de $[0, 1]$ -relaciones de tolerancia difusas separables $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$, la aplicación $\|\cdot\|_S^T : For(B) \rightarrow [0, 1]^U$ 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 , por medio de la relación entre los valores de los atributos en el objeto x y los valores de los atributos en la fórmula Φ .

Por último, para cada $\Phi, \Psi \in For(B)$, la conjunción y disyunción de fórmulas se define, para todo $x \in U$, como sigue:

$$\begin{aligned} \|\Phi \wedge \Psi\|_S^T(x) &= \min\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\} \\ \|\Phi \vee \Psi\|_S^T(x) &= \max\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\} \end{aligned}$$

En consecuencia, $\|\Phi\|_S^T$ es un conjunto difuso, donde el grado de pertenencia de cada objeto a dicho conjunto se corresponde con el grado de satisfacibilidad de la fórmula Φ por parte del objeto considerado. En el siguiente resultado, se muestra que la Definición 3.2 generaliza la Definición 2.13, siendo esta última introducida por Pawlak.

Proposición 3.14. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}_d$, $\Phi \in For(B)$ y $x \in U$. Si $T = \{T_a : V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ es una familia de relaciones de tolerancia separables booleanas entonces:

$$x \in \|\Phi\|_S \quad \text{si y solo si} \quad \|\Phi\|_S^T(x) = 1$$

Demostración. Se halla en la demostración de la Proposición 24 en [20]. $\square$

Por consiguiente, y dado que la intersección y unión de conjuntos se generalizan con el mínimo y máximo de conjuntos difusos respectivamente, la conjunción y disyunción de fórmulas de la Definición 3.2 generalizan

a la conjunción y disyunción de fórmulas de la Definición 2.13. Aunque la definición de regla de decisión en el entorno difuso es análoga a la Definición 2.14, esta es incluida para hacer esta sección autocontenida.

Definición 3.3. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $B \subseteq \mathcal{A}$. Una regla de decisión en S es una expresión $\Phi \rightarrow \Psi$, donde $\Phi \in \text{For}(B)$, $\Psi \in \text{For}(\{d\})$ son el antecedente y consecuente de la regla decisión, respectivamente.

Además, dada una familia de $[0, 1]$ -relaciones de tolerancia difusas separables $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$, diremos que un objeto $x \in U$ satisface la regla de decisión $\Phi \rightarrow \Psi$ si $\|\Phi \wedge \Psi\|_S^T(x) = 1$. De esta forma, podemos decir que el objeto x induce la regla de decisión $\Phi \rightarrow \Psi$ y esta puede denotarse como $\Phi_x \rightarrow \Psi_x$.

Las reglas de decisión son esenciales para sintetizar la información contenida en tablas de decisión y sobre ellas se definirán varias nociones que contribuyan a analizar estas tablas. Por otro lado, en el Ejemplo 25 de [20] se detalla el proceso de obtención de un conjunto de reglas para el análisis de una tabla de decisión. En primer lugar, se analiza la dispensabilidad de cada atributo de condición mediante la noción de reducto difuso de m -decisión, detectándose la posibilidad de prescindir de un atributo sin perder información relevante. Posteriormente, se extrae la regla inducida por cada objeto de la tabla considerando todos los atributos de condición excepto el descartado. Finalmente, se calculan los grados de satisfacibilidad a cada antecedente y consecuente de dichas reglas por parte de cada objeto de la tabla, ya que serán fundamentales en lo que sigue.

Para la extracción de conclusiones significativas a partir de las reglas de decisión, es necesario su descripción por medio de indicadores de relevancia. A continuación, se presentan los indicadores equivalentes a los dados en la Definición 2.15 en el entorno difuso.

Definición 3.4. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $\Phi \rightarrow \Psi$ una regla de decisión en S y $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ una familia de

$[0, 1]$ -relaciones de tolerancia difusas separables. Se denomina:

- T -soporte de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$supp_S^T(\Phi, \Psi) = card_F(\|\Phi \wedge \Psi\|_S^T)$$

- T -fuerza de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$\sigma_S^T(\Phi, \Psi) = \frac{supp_S^T(\Phi, \Psi)}{card(U)}$$

- T -certeza de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$cer_S^T(\Phi, \Psi) = \frac{supp_S^T(\Phi, \Psi)}{card_F(\|\Phi\|_S^T)}$$

cuando $card_F(\|\Phi\|_S^T) \neq 0$.

- T -cobertura de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$cov_S^T(\Phi, \Psi) = \frac{supp_S^T(\Phi, \Psi)}{card_F(\|\Psi\|_S^T)}$$

cuando $card_F(\|\Psi\|_S^T) \neq 0$.

Mediante los indicadores introducidos en esta definición es posible interpretar cada una de las reglas de decisión obtenidas a partir de una tabla de decisión en el entorno difuso. De esta forma, es posible compararlas entre ellas y destacar aquellas más significativas para describir el conjunto de datos analizado. Se consideran reglas significativas a aquellas que tienen mayor T -soporte, y en consecuencia mayor T -fuerza, pues muestran la representatividad de dicha regla en la tabla, y las de mayor T -certeza, ya que este indicador expone el impacto de la satisfacción del antecedente en la satisfacción del consecuente. Como consecuencia, a mayor T -certeza, más fiable será la regla. Por otro lado, la T -cobertura es muy útil para detectar el impacto de la satisfacción del consecuente en la satisfacción del antecedente.

En el Ejemplo 29 de [20] se continúa el Ejemplo 25 calculando el valor de cada uno de los indicadores anteriores para cada regla considerada. Dichos valores se recogen en la Tabla 11 del mencionado ejemplo, donde destacan la regla r_5 por su elevado T -soporte, y las reglas r_2 y r_4 por sus altas T -cobertura y T -certeza, respectivamente, a pesar de sus bajos T -soportes.

El siguiente resultado muestra que, efectivamente, la Definición 3.4 extiende la Definición 2.15 al marco de trabajo difuso.

Proposición 3.15. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $\Phi \rightarrow \Psi$ una regla de decisión en S . Si $T = \{T_a: V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ es una familia de relaciones de tolerancia separables booleanas entonces:*

$$\begin{aligned} \text{supp}_S^T(\Phi, \Psi) &= \text{supp}_S(\Phi, \Psi) \\ \sigma_S^T(\Phi, \Psi) &= \sigma_S(\Phi, \Psi) \\ \text{cer}_S^T(\Phi, \Psi) &= \text{cer}_S(\Phi, \Psi) \\ \text{cov}_S^T(\Phi, \Psi) &= \text{cov}_S(\Phi, \Psi) \end{aligned}$$

Demostración. Obsérvese la demostración de la Proposición 28 en [20]. $\square$

Podemos concluir que esta sección recoge la extensión de varias nociones relacionadas con las reglas de decisión al entorno difuso, permitiendo analizar tablas de decisión bajo este marco de trabajo. En la siguiente sección se continúa generalizando nociones clásicas, esta vez relacionadas con los algoritmos de decisión.

3.2.2. Algoritmos de decisión

En esta sección se presenta nuestro estudio sobre los algoritmos de decisión, ya que son especialmente importantes para el análisis de la información presente en las tablas de decisión. Esto se debe a que están formados por un conjunto de reglas de decisión que describen a dichas tablas sin

redundancias, sin perder información y de manera precisa. Antes de presentar los algoritmos en el ambiente difuso, es necesario introducir un par de nociones previas. La primera de ellas es esencial para comparar pares de fórmulas.

Definición 3.5. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ una familia de $[0, 1]$ -relaciones de tolerancia difusas separables y $\Phi, \Phi' \in For(\mathcal{A}_d)$, con $\Phi = (a_1, v_1) \wedge \dots \wedge (a_n, v_n)$ y $\Phi' = (a'_1, w_1) \wedge \dots \wedge (a'_m, w_m)$. La relación de F -indiscernibilidad es una $[0, 1]$ -relación de tolerancia difusa separable $R_{Fd} : For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$ definida como

$$R_{Fd}(\Phi, \Phi') = \begin{cases} \bigwedge_{i \in \{1, \dots, n\}} T_{a_i}(v_i, w_i) & \text{si } n = m \text{ y } a_i = a'_i \text{ para todo } i \in \{1, \dots, n\} \\ 0 & \text{en caso contrario} \end{cases}$$

Dado $\alpha \in [0, 1]$, el R_{Fd} - α -bloque se define para cada $\Phi \in For(\mathcal{A}_d)$ como:

$$[\Phi]_\alpha = \{\Phi' \in For(\mathcal{A}_d) \mid \alpha \leq R_{Fd}(\Phi, \Phi')\}$$

Si $\Phi' \in [\Phi]_\alpha$ entonces se dice que Φ y Φ' están R_{Fd} - α -relacionados.

Obsérvese que, para comparar dos fórmulas, ambas deben estar compuestas por los mismos atributos. Además, se puede ver que esta relación no es de equivalencia. Como resultado, se obtienen bloques de tolerancia en lugar de clases de equivalencia. En el Ejemplo 31 de [20], se comparan los antecedentes de todas las reglas de decisión, obteniéndose los resultados presentes en la Tabla 12. Por último, el umbral α debe ser establecido por el usuario cuando se lleve a cabo el estudio de acuerdo a sus preferencias. De aquí, si desea imponer que las fórmulas sean muy similares entre sí para que se consideren relacionadas entonces debe tomar un valor de α elevado. En caso contrario, deberá permitir una mayor tolerancia disminuyendo el valor de α . Este umbral es indispensable para determinar la consistencia de una regla dada.

Definición 3.6. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $\alpha \in (0, 1]$, $R_{Fd}: For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$ una relación de F -indiscernibilidad y $Dec(S)$ un conjunto de reglas de decisión en S . El conjunto de reglas de decisión α -consistentes se define como:

$$Dec_{\alpha}^{+}(S) = \{\Phi \rightarrow \Psi \in Dec(S) \mid \text{si para cada } \Phi' \rightarrow \Psi' \in Dec(S) \text{ tal} \\ \text{que } \Phi' \in [\Phi]_{\alpha} \text{ entonces } \Psi' \in [\Psi]_{\alpha}\}$$

Nótese que, el caso $\alpha = 0$ no se ha considerado en esta definición ya que, para cualquier conjunto de reglas de decisión $Dec(S)$ se obtiene por la Definición 3.5 que $Dec_0^{+}(S) = Dec(S)$. Como consecuencia, cualquier regla es 0-consistente y por ello, el conjunto $Dec_0^{+}(S)$ no aporta información de interés. Cuando $\alpha \neq 0$, el conjunto de reglas α -consistentes resulta de vital importancia para la síntesis de la información presente en una tabla de decisión, ya que contiene a las reglas de decisión que no presentan contradicciones, dado el umbral α .

Ahora, y tras haber introducido todas las nociones previas necesarias, se presenta la noción de algoritmo de decisión en el marco difuso.

Definición 3.7. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $\alpha \in (0, 1]$, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1)$, $\alpha_3 \in (0, card(U)]$, $Dec(S) = \{\Phi_i \rightarrow \Psi_i \mid i \in \{1, 2, \dots, m\}\}$, con $m \geq 2$ un conjunto de reglas de decisión en S y $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ una familia de $[0, 1]$ -relaciones de tolerancia difusas separables. Se dice que:

1. $Dec(S)$ es un conjunto de reglas de decisión $\alpha_1\alpha_2$ -mutuamente excluyentes por parejas (independientes), si cada par de reglas de decisión $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisface que $\Phi = \Phi'$ o $\|\Phi \wedge \Phi'\|_S^T(x) \leq \alpha_1$ y $\Psi = \Psi'$ o $\|\Psi \wedge \Psi'\|_S^T(x) \leq \alpha_2$, para todo $x \in U$.
2. $Dec(S)$ cubre U si $card_F(\|\bigvee_{i=1}^m \Phi_i\|_S^T) = card_F(\|\bigvee_{i=1}^m \Psi_i\|_S^T) = card(U)$.

3. Una regla de decisión $\Phi \rightarrow \Psi \in Dec(S)$ es α_3 -admisible en S si $\alpha_3 \leq \text{supp}_S^T(\Phi, \Psi)$.
4. $Dec(S)$ preserva la α -consistencia de S con un grado α_4 si se verifica la siguiente igualdad para todo $x \in U$:

$$|POS_{\mathcal{A}}^f(x) - \bigvee_{\Phi \rightarrow \Psi \in Dec_{\alpha}^+(S)} \|\Phi\|_S^T(x)| \leq \alpha_4$$

A un conjunto de reglas de decisión $Dec(S)$ satisfaciendo las propiedades anteriores para los valores $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ se le denomina $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -algoritmo de decisión en S y se denota como $DA_T(S)$.

Además del umbral α introducido anteriormente, la Definición 3.7 presenta otros cuatro umbrales. Los dos primeros, α_1 y α_2 , se centran en comparar cada par de antecedentes y consecuentes, respectivamente, para valorar cuán diferentes son entre ellos. A menor umbral, más diferencia habrá entre ellos y por tanto más estricta será esta propiedad. Por otro lado, el umbral α_3 establece el mínimo de los T -soportes de todas las reglas. De esta forma, cuanto mayor sea este umbral, más representativas serán las reglas en la tabla. Con respecto al umbral α_4 , tiene como objetivo comparar dos nociones centradas en el análisis de la consistencia en bases de datos. La primera es la región $\mathcal{A}$ -positiva difusa multiadjunta, la segunda las reglas α -consistentes. Así, cuanto menor sea el umbral α_4 , más equivalente será la consistencia proporcionada por el conjunto de reglas con la presente en la tabla de decisión. En particular, el objetivo perseguido determinando los umbrales $\alpha_1, \alpha_2, \alpha_4$ bajos y el umbral α_3 elevado es el de obtener un algoritmo lo más estricto posible, ya que estos umbrales son representativos de la capacidad de clasificación del algoritmo por lo comentado anteriormente. Por último, nótese que los valores $\alpha_1 = \alpha_2 = \alpha_4 = 1$ no son tenidos en cuenta en esta definición, ya que estos verifican las desigualdades correspondientes en cualquier circunstancia, y por tanto, no aportan información relevante para el estudio.

Es importante destacar que, eligiendo adecuadamente los umbrales $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ a partir de α , el conjunto de reglas inducidas por cada objeto de la tabla de decisión estudiada conforma un $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmo de decisión, como se muestra en el Ejemplo 38 de [20]. Bajo dichas hipótesis, estos algoritmos son un caso particular de los $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmos de decisión completos.

Definición 3.8. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $DA_T(S)$ un $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmo de decisión. Si para cada $x \in U$ existe una regla $\Phi \rightarrow \Psi \in DA_T(S)$ inducida por x entonces a $DA_T(S)$ se le denomina $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmo de decisión completo.

Como consecuencia, para describir una tabla de decisión con mayor precisión en términos de un algoritmo de decisión, se recomienda que este sea completo para garantizar que cada objeto se encuentre totalmente representado en una regla de dicho algoritmo.

El siguiente resultado muestra que el conjunto de reglas α -consistentes de la Definición 3.6 generaliza al conjunto de reglas ciertas de la Definición 2.16.

Proposición 3.16. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión y $T = \{T_a: V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ una familia de relaciones de tolerancia separables booleanas. Si $\alpha > 0$ y $DA_T(S)$ es un $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmo de decisión completo entonces

$$Dec_\alpha^+(S) = Dec^+(S)$$

Demostración. Consúltese la prueba de la Proposición 35 en [20]. $\square$

Para finalizar, es esencial remarcar que los $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmos de decisión completos generalizan a los clásicos dados en la Definición 2.16, como se expone en el siguiente resultado.

Proposición 3.17. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $T = \{T_a : V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ una familia de relaciones de tolerancia separables booleanas, $R_{\mathcal{A}}$ y R_a relaciones de tolerancia separables booleanas para todo $a \in \mathcal{A}_d$. Dado $DA_T(S)$, $\alpha \in (0, 1]$, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1)$ y $0 < \alpha_3 \leq \min\{\text{supp}_S^T(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in DA_T(S)\}$, entonces $DA_T(S)$ es un $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -algoritmo de decisión completo si y solo si es un algoritmo de decisión completo.

Demostración. Véase la demostración de la Proposición 37 en [20]. $\square$

Por tanto, la noción propuesta de algoritmo de decisión para el entorno difuso extiende la noción clásica introducida por Pawlak y, dada su importancia en el marco de trabajo clásico, se concluye que la Definición 3.8 es de especial interés para el análisis de información desde el punto de vista difuso.

Por último, en la Sección 5 de [20] también se incluyó un ejemplo de un conjunto de datos reales para analizarlo con las nociones presentadas. Se eligió dicho ejemplo porque fue analizado desde un punto de vista clásico en [76]. Utilizando una amplia variedad de nociones introducidas en [20] se pudo realizar un análisis bastante profundo, extrayendo así numerosas conclusiones acerca de las reglas de decisión, mostrando aquellas más representativas y el porqué de este hecho, etc. Todas estas conclusiones se pueden consultar en [20].

En definitiva, esta sección se ha centrado en la introducción de numerosas nociones de RST en el entorno difuso para el análisis de información en este marco de trabajo. Se ha demostrado que todas ellas generalizan a las correspondientes nociones clásicas, por lo que serán de especial importancia para estudiar conjuntos de datos con información imperfecta, imprecisa o incompleta.

3.3. Clasificación formal de nuevos objetos

Otro tema de gran interés en el análisis de la información contenida en conjuntos de datos es el uso de la misma para predecir el comportamiento de nuevos datos. Un ejemplo de este hecho es la selección de tratamientos de nuevos pacientes basándose en el historial clínico de pacientes anteriores [46], pero también es aplicable a otros ámbitos, como la economía o las ciencias sociales. Este problema también se ha abordado desde el marco de trabajo de RST [41, 68], pero sigue habiendo terreno por explorar. En esta sección, se mostrarán las propuestas introducidas en [15], donde se presentaron cuatro métodos diferentes para la clasificación de nuevos objetos que se integran en una tabla de decisión dada. Cada uno de estos métodos se basará en una nueva noción, que será determinante para dar un grado de verdad para cada posible decisión a tomar por el método en cuestión. Además, estas nociones se apoyan en un conjunto de reglas de decisión dado, por lo que dependiendo de dicho conjunto es posible obtener diferentes decisiones para el nuevo objeto. Este estudio se apoyará en las reglas inducidas por los objetos de la tabla de decisión considerando todos los atributos, para así usar toda la información disponible en la tabla.

Por otro lado, es importante destacar que los valores de los atributos de condición en el nuevo objeto se asumen conocidos y el de decisión desconocido. De esta forma, a través de los atributos de condición, se asignará un valor al atributo de decisión, clasificando así al nuevo objeto. Además, el universo de objetos queda dividido en dos conjuntos, $\mathcal{U} = U \cup U^c$, donde U son los objetos de la tabla de decisión dada sobre los que se dispone de total información y U^c los objetos cuyo valor del atributo de decisión es desconocido.

Antes de presentar los métodos de clasificación propuestos, es necesario introducir algunas nociones que serán útiles en lo que sigue. En primer

lugar, dado que las fórmulas introducidas en la Definición 3.2 son conjuntos difusos, es posible definir α -cortes sobre ellas.

Definición 3.9. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $B \subseteq \mathcal{A}_d$ y $\Phi \in For(B)$. Dado $\alpha \in [0, 1)$, el α -corte estricto de $\|\Phi\|_S^T$ se define como

$$(\|\Phi\|_S^T)_\alpha = \{x \in U \mid \|\Phi\|_S^T(x) > \alpha\}$$

Obsérvese que, a diferencia de la definición usual de α -corte de un conjunto difuso [101], se ha considerado una desigualdad estricta [104]. Este hecho será importante en lo que sigue, ya que permitirá demostrar que uno de los cuatro métodos de clasificación propuestos en este trabajo generaliza a uno introducido previamente en la literatura. Por otra parte, los α -cortes son de especial importancia para proporcionar los objetos que satisfacen la fórmula en un mayor grado, permitiendo seleccionar aquellos objetos que son más representativos de la fórmula en función de α .

A continuación, se presentan un par de indicadores de relevancia que permitirán continuar analizando las reglas de decisión en el entorno de FRST. El primero de ellos es una ligera variante del T -soporte.

Definición 3.10. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión, $\Phi \rightarrow \Psi \in Dec(S)$ y $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ una familia de $[0, 1]$ -relaciones de tolerancia difusas separables. Se denomina T -soporte normalizado (NT -soporte) de la regla de decisión $\Phi \rightarrow \Psi$ al valor:

$$supp_S^{NT}(\Phi, \Psi) = \frac{supp_S^T(\Phi, \Psi)}{\max\{supp_S^T(\Phi', \Psi') \mid \Phi' \rightarrow \Psi' \in Dec(S)\}}$$

Así, el NT -soporte de una regla es una normalización de su T -soporte respecto al máximo de estos últimos, en lugar de respecto al cardinal de objetos como lo considera la T -fuerza.

Por otro lado, la credibilidad de una regla de decisión ya ha sido definida en FRST anteriormente [41, 90, 91]. Esta muestra el grado en el que,

si un objeto de la tabla de decisión satisface el antecedente de la regla estudiada, o antecedentes de reglas similares, entonces también satisface el consecuente de la regla de decisión dada. Por tanto, a mayor credibilidad, más veraz será la descripción dada por dicha regla. A continuación, se introduce la extensión de la credibilidad de una regla de decisión al marco de trabajo multiadjunto.

Definición 3.11. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión, $\Phi_x \rightarrow \Psi_x \in Dec(S)$ la regla de decisión inducida por un objeto $x \in U$, $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ una familia de $[0, 1]$ -relaciones de tolerancia difusas separables y $\alpha \in [0, 1]$. Dado el marco multiadjunto orientado a propiedades $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$, donde $\&_i$ es un conjuntor asociativo para todo $i \in \{1, \dots, n\}$; se denomina α - T -credibilidad de la regla de decisión $\Phi_x \rightarrow \Psi_x$, denotada como $\mu_{S,\tau}^\alpha(\Phi_x, \Psi_x)$, al valor

$$\mu_{S,\tau}^\alpha(\Phi_x, \Psi_x) = \&_i \{ I_{\tau(x,y)}(\|\Phi_x\|_S^T(y), \mu_{\Psi_{x,S,\tau}}(y)) \mid y \in (\|\Phi_x\|_S^T)_\alpha \}$$

con

$$\mu_{\Psi_{x,S,\tau}}(y) = \&_{j_y} \{ I_{\tau(y,y')}(\|\Phi_y\|_S^T(y'), \|\Psi_x\|_S^T(y')) \mid y' \in (\|\Phi_y\|_S^T)_\alpha \text{ y } \Phi_y \rightarrow \Psi_y \in Dec(S) \}$$

donde $i, j_y \in \{1, \dots, n\}$ y $\Phi_y \rightarrow \Psi_y \in Dec(S)$ es una regla de decisión inducida por $y \in (\|\Phi_x\|_S^T)_\alpha$.

Dada la complejidad de la noción anterior, se expone una explicación detallada de la misma. En primer lugar, sea $\Phi_x \rightarrow \Psi_x$ una regla de decisión de la cual se quiere calcular su α - T -credibilidad. Esta es una regla inducida por el objeto x , lo cual no supone una restricción exigente ya que es natural considerar reglas de decisión inducidas por objetos para el estudio de tablas de decisión. A continuación, para cada objeto $y \in (\|\Phi_x\|_S^T)_\alpha$, es decir, para cada objeto representativo del antecedente Φ_x dado el umbral α , se calcula el valor $\mu_{\Psi_{x,S,\tau}}(y)$, el cual representa la credibilidad de y con respecto a x . Este cálculo se realiza mediante la comparativa de los grados de satisfacibilidad de objetos representativos de Φ_y con respecto a

esta fórmula y los grados de satisfacibilidad de dichos objetos con respecto al consecuente Ψ_x . De esta forma, cuanto mayor sea el valor $\mu_{\Psi_x, S, \tau}(y)$, más fiable será la información dada por el objeto y . Por otro lado, el valor $\mu_{S, \tau}^\alpha(\Phi_x, \Psi_x)$ estudia la transitividad de los valores $\|\Phi_x\|_S^T(y)$ a $\|\Psi_x\|_S^T(y')$ por medio de $\|\Phi_y\|_S^T(y')$ y la definición de $\mu_{\Psi_x, S, \tau}(y)$. Como consecuencia, si el valor $\mu_{\Psi_x, S, \tau}(y)$ es elevado, y dado que el objeto y es representativo del antecedente Φ_x , la α - T -credibilidad de la regla de decisión $\Phi_x \rightarrow \Psi_x$ también será elevada. Esta definición será fundamental para los métodos de clasificación que serán presentados en las Secciones 3.3.3 y 3.3.4.

Antes de presentar cada uno de los métodos, es conveniente destacar que todos siguen una estructura similar, de manera que se refuerza el procedimiento empleado. En ella, se comienza definiendo una nueva noción que permite calcular cuán recomendable es la toma de cada una de las posibles decisiones para cada nuevo objeto a clasificar. De la noción correspondiente se obtendrá un valor perteneciente al intervalo unidad, de manera que, a mayor valor obtenido más recomendable será la toma de la decisión especificada. Respecto a los métodos, el primer paso siempre se centra en el cómputo de estos valores para cada posible decisión a tomar. En el segundo, todos estos valores son comparados y si existe alguno elevado y suficientemente diferente del resto de valores entonces es posible tomar una decisión. En caso contrario, cada método proporciona pasos alternativos para recalcular valores que nos permitan tomar una decisión de manera clara y precisa.

3.3.1. Método de idoneidad

Para el primer método se considerará el grado de satisfacibilidad dado en la Definición 3.2. De esta forma, se calculará el grado de satisfacibilidad del nuevo objeto estudiado con respecto a cada antecedente de las reglas de decisión consideradas. Cuanto mayor sea este grado, más se parecerá el

nuevo objeto al antecedente correspondiente, por lo que más recomendable será tomar la decisión del consecuente asociado a dicho antecedente. La noción en la que se basa el primer método propuesto es el grado de idoneidad y es presentado a continuación.

Definición 3.12. 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 $\sigma_{Dec(S)}: U^c \times V_d \rightarrow [0, 1]$, definida como

$$\sigma_{Dec(S)}(x, v) = \max\{\|\Phi\|_S^T(x) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

Es importante destacar que la Definición 3.12 solo se centra en el máximo de los grados de satisfacibilidad para toda posible decisión, en una línea similar a la seguida en [45, 68, 87], con el objetivo de considerar solo la regla más “cercana” al objeto. Sin embargo, este operador agregador podrá ser modificado si es necesario tomar una decisión más precisa. Además, es fácil ver que $0 \leq \sigma_{Dec(S)}(x, v) \leq 1$ para todo nuevo objeto x , decisión $\bar{d}(x) = v$, y conjunto de regla de decisión $Dec(S)$ de acuerdo a la Definición 3.2, por lo que el grado de idoneidad es un grado de verdad.

A continuación, se presenta el primer método de clasificación de nuevos objetos propuesto en [15]. Este método se denomina *método de idoneidad* y se abrevia como *S-método*, del inglés *suitability method*. Los pasos de este método son los siguientes:

- S1. Dado un nuevo objeto a estudiar $x \in U^c$, se calculan los grados de idoneidad $\sigma_{Dec(S)}(x, v)$ de cada decisión $\bar{d}(x) = v$, con $v \in V_d$.
- S2. Se comparan todos los grados de idoneidad obtenidos. Si el grado de idoneidad de una decisión $\bar{d}(x) = v$ es elevado y suficientemente diferente de los grados del resto de decisiones posibles, entonces se toma la decisión $\bar{d}(x) = v$ para el objeto x .

- S3. Si existen algunos valores del atributo de decisión $v \in V \subset V_d$ con grados de idoneidad similares de manera que no es posible tomar una decisión para el objeto x , entonces es necesario considerar más grados de satisfacibilidad para estas decisiones dependiendo de un umbral $\beta \in [0, 1]$ en lugar de solo el máximo. En particular, considerando el conjunto $\mathcal{F}_x^\beta(v) = \{\Phi \in For(\mathcal{A}) \mid \Phi \rightarrow \Psi \in Dec(S), \beta \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}$, se define para cada decisión $\bar{d}(x) = v$, con $v \in V$, el valor

$$\mathcal{M}_{\mathcal{F}_x^\beta(v)} = \sum_{\Phi \in \mathcal{F}_x^\beta(v)} \frac{\|\Phi\|_S^T(x)}{\text{card}(\mathcal{F}_x^\beta(v))}$$

Como consecuencia, se calcula una media aritmética de los grados de satisfacibilidad que superan el umbral β . Además, obsérvese que, si $\text{card}(\mathcal{F}_x^\beta(v)) = 1$, entonces $\mathcal{M}_{\mathcal{F}_x^\beta(v)} = \sigma_{Dec(S)}(x, v)$. Una vez realizado este cálculo para cada $v \in V$, se comparan los valores obtenidos, tomando la decisión $\bar{d}(x) = w$ si $\mathcal{M}_{\mathcal{F}_x^\beta(w)}$ es elevado y suficientemente diferente de $\mathcal{M}_{\mathcal{F}_x^\beta(v)}$ para todo $v \in V \setminus \{w\}$.

- S4. Si no es posible tomar una decisión de acuerdo al tercer paso, se modifica el conjunto $\mathcal{F}_x^\beta(v)$, ya sea cambiando el umbral β o considerando otro operador de agregación en lugar de la media aritmética que pueda adaptarse mejor a los datos.

El S-método es aplicado en el Ejemplo 2 de [15], donde se ha estudiado la clasificación de tres nuevos objetos de acuerdo a la tabla de decisión del Ejemplo 1 de [15]. Los dos primeros objetos han sido clasificados usando solamente los dos primeros pasos del S-método, ya que había una gran diferencia entre los grados de idoneidad obtenidos. Sin embargo, para clasificar el último nuevo objeto ha sido necesaria la ejecución de los cuatro pasos presentados. En cualquier caso, se ha podido tomar una decisión robusta, mostrando así las ventajas del S-método.

Finalmente, en [15] también se concluye que el S-método presenta una mejora sobre el método propuesto en [87]. Esto se debe principalmente al uso de relaciones de tolerancia separables para realizar comparativas más precisas, y a la inclusión de los pasos S3 y S4, los cuales ofrecen muchas posibilidades al usuario para llevar a cabo la toma de la decisión más adecuada.

3.3.2. Método de ε -representatividad

Puesto que en el ejemplo mencionado anteriormente se necesitó usar los cuatro pasos del S-método para clasificar el tercer nuevo objeto, se hace palpable la necesidad de proporcionar un método de clasificación alternativo para facilitar la toma de decisiones en estas situaciones. Además, en el S-método no se tiene en cuenta el T -soporte de las reglas, el cual es un indicador de especial relevancia ya que proporciona la representatividad de la reglas en la tabla de decisión. El segundo método de clasificación propuesto en [15], el cual proporcionará una toma de decisiones más sólida y tendrá en consideración el T -soporte de las reglas de decisión comparadas con el nuevo objeto, se basa en la siguiente definición.

Definición 3.13. 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 $\rho_{Dec(S)}^\varepsilon: U^c \times V_d \rightarrow [0, 1]$ definida como

$$\rho_{Dec(S)}^\varepsilon(x, v) = \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | \varepsilon \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}$$

Si $\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\} = \emptyset$, entonces el objeto x es ε -indecidible.

Nuevamente se obtiene un valor de verdad, ya que es fácil ver que $0 \leq \rho_{Dec(S)}^\varepsilon(x, v) \leq 1$, independientemente del objeto x a clasificar, de la decisión $\bar{d}(x) = v$ estudiada, del conjunto de reglas de decisión $Dec(S)$ considerado y del umbral ε .

Obsérvese que, para el cálculo del grado de ε -representatividad de una decisión $\bar{d}(x) = v$ se consideran las reglas de decisión cuyos antecedentes sean satisfechos por el nuevo objeto x con un grado de satisfacibilidad mayor o igual que ε . Por tanto, el umbral ε permite adaptar el cálculo a las necesidades del usuario. Si este prefiere realizar un estudio más estricto, puede considerar un mayor valor de ε de manera que solo se tengan en cuenta las reglas cuyos antecedentes sean más similares al nuevo objeto. En caso contrario, se considerarán más reglas en la comparativa, pudiendo conducir a una clasificación más imprecisa. Además, los T -soportes de las reglas han sido ponderados por los grados de satisfacibilidad correspondientes, de manera que a igualdad de T -soporte entre dos reglas, tendrá un mayor peso aquella regla cuyo antecedente sea más similar al objeto estudiado.

Antes de introducir el método de clasificación basado en la Definición 3.13 se presentan un par de resultados a tener en cuenta.

Proposición 3.18. *Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión, $x \in U^c$ un nuevo objeto y $\varepsilon \in [0, 1]$. Entonces*

- $\sum_{v \in V_d} \rho_{Dec(S)}^\varepsilon(x, v) = 1$.
- $\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d\} < \varepsilon$ si y solo si el objeto x es ε -indecidible.

Demostración. Véase la prueba de la Proposición 1 de [15]. □

Del primer ítem de la proposición anterior se concluye que en general los grados de ε -representatividad de diferentes decisiones serán más similares entre sí que los grados de idoneidad, por lo que se deberá ser más

flexible en la comparativa de los grados de ε -representatividad. Por otro lado, el segundo ítem relaciona las Definiciones 3.12 y 3.13 por medio del umbral ε cuando no sea posible tomar una decisión de acuerdo al grado de ε -representatividad.

El siguiente resultado se centra en aquellos casos en los que es posible destacar un valor del atributo de decisión con respecto al resto.

Lema 3.1. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión y $x \in U^c$ un nuevo objeto. Si existe $v_k \in V_d$ y $\varepsilon \in [0, 1]$ satisfaciendo

$$\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d \setminus \{v_k\}\} < \varepsilon \leq \sigma_{Dec(S)}(x, v_k)$$

entonces

$$\begin{aligned} \rho_{Dec(S)}^\varepsilon(x, v_k) &= 1 \\ \rho_{Dec(S)}^\varepsilon(x, v_j) &= 0 \end{aligned}$$

para todo $v_j \in V_d \setminus \{v_k\}$.

Demostración. Se encuentra en la demostración del Lema 1 en [15]. $\square$

A partir del Lema 3.1, es posible tomar decisiones de manera completamente precisa, pues una decisión tendrá grado de ε -representatividad igual a 1 y el resto 0. Por lo tanto, este resultado ayuda a situaciones en las que es bastante difícil tomar una decisión, como la clasificación del tercer objeto nuevo en el Ejemplo 2 de [15]. Un posible inconveniente del Lema 3.1 es que el valor del umbral puede resultar demasiado elevado para el usuario, el cual desearía realizar un estudio con mayor flexibilidad. Por ello, se presenta el segundo método de clasificación introducido en [15], el cual se basa en la Definición 3.13 y se denomina *método de ε -representatividad*, abreviado como *R-método*, que proviene del inglés *ε -representativeness method*. Los pasos del R-método se presentan a continuación:

- R1. Dado un nuevo objeto a clasificar $x \in U^c$ y un umbral ε , se calculan los grados de ε -representatividad $\rho_{Dec(S)}^\varepsilon(x, v)$ para todo $v \in V_d$.
- R2. Se comparan todos los grados de ε -representatividad obtenidos. Si el grado de ε -representatividad de una decisión $\bar{d}(x) = v$ es elevado y suficientemente diferente del resto de decisiones posible, entonces se tomará la decisión $\bar{d}(x) = v$ para el objeto x .
- R3. Si existen algunos valores del atributo de decisión con grados de ε -representatividad similares que impiden clasificar al objeto en cuestión, es necesario cambiar el umbral ε para realizar un estudio más o menos flexible, bajando o subiendo dicho umbral respectivamente.

El R-método fue empleado en el Ejemplo 3 de [15], afrontando el mismo estudio de clasificación que en el ejemplo anterior. En primer lugar, se aplica el método considerando el mismo umbral $\varepsilon = 0.7$ para cada uno de los tres nuevos objetos. Para los dos primeros se obtiene una clasificación totalmente precisa, ya que el umbral dado verifica la hipótesis del Lema 3.1. Este hecho no se da para el tercer objeto, pero aun así es posible tomar una decisión en el paso R2, sin necesidad de cambiar el umbral. Además, la decisión dada para cada nuevo objeto coincide con las obtenidas en el ejemplo anterior, lo cual refuerza ambos métodos. También se estudia el grado de ε -representatividad para cada posible decisión y umbral $\varepsilon \in [0, 1]$ en cada nuevo objeto. Se concluye que la mejor decisión a tomar para el primer objeto no cambia independientemente del umbral elegido, mientras que en los dos últimos sí se produce un cambio en los umbrales 0.54 y 0.69, respectivamente. Por último, es importante destacar que el rango de posibles umbrales verifican el Lema 3.1 para el último objeto nuevo es reducido, exponiendo así la mayor dificultad existente para clasificar al mismo.

En conclusión, el R-método mejora significativamente al S-método, ya que resuelve los problemas que sufre este último y proporciona una ma-

yor libertad al usuario, adaptando el estudio a sus necesidades. Además, como se comprueba en el Ejemplo 3 de [15], el R-método proporciona más información que el S-método, permitiendo extraer conclusiones significativas sobre los nuevos objetos a clasificar.

3.3.3. Método de α -credibilidad

Algunos autores [41, 91] incluyeron el indicador de la credibilidad de las reglas de decisión en su estudio de clasificación de nuevos objetos, dada la importancia del mismo en el análisis de las reglas. Puesto que la Definición 3.11 extiende a las dadas en [41, 91] mediante el uso de los pares adjuntos y del umbral α , se proporcionará un método de clasificación más adaptable a las necesidades del usuario. Dicho método se basa en la siguiente noción.

Definición 3.14. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión, $\alpha \in [0, 1)$ y $(\&, I)$ un par adjunto. Para cada nuevo objeto $x \in U^c$ y $v \in V_d$, el grado de α -credibilidad de la decisión $\bar{d}(x) = v$ de acuerdo a $Dec(S)$ viene dado por la aplicación $\theta_{Dec(S)}^\alpha: U^c \times V_d \rightarrow [0, 1]$, definida como

$$\theta_{Dec(S)}^\alpha(x, v) = \text{máx}\{\&(\|\Phi\|_S^T(x), \mu_{S, \tau}^\alpha(\Phi, \Psi)) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

Se obtiene que $0 \leq \theta_{Dec(S)}^\alpha(x, v) \leq 1$ para todo nuevo objeto x , decisión $\bar{d}(x) = v$, y conjunto de regla de decisión $Dec(S)$. En consecuencia, el grado de α -credibilidad es un grado de verdad. El *método de α -credibilidad*, también abreviado como *C-método*, del inglés *α -credibility method*, está compuesto por los siguientes pasos:

- C1. Dado un nuevo objeto a clasificar $x \in U^c$, se calculan los grados de α -credibilidad $\theta_{Dec(S)}^\alpha(x, v)$ para todo $v \in V_d$.

- C2. Se comparan todos los grados de α -credibilidad. Si existe una decisión $w \in V_d$ tal que $\theta_{Dec(S)}^\alpha(x, w)$ es elevado y suficientemente diferente de $\theta_{Dec(S)}^\alpha(x, v)$, para todo $v \in V_d \setminus \{w\}$, entonces se tomará la decisión $\bar{d}(x) = w$ para el objeto x .
- C3. Si existen algunos valores del atributo de decisión con grados de α -credibilidad similares, entonces se debe reexaminar la relación asumida entre las variables involucradas en el cálculo. Así, es posible seleccionar otro par adjunto $(\&, I)$ del marco multiadjunto orientado a propiedades considerado, para agregar las variables correspondientes de una mejor forma.

En el Ejemplo 4 de [15] se aplicó el C-método para clasificar los mismos objetos ya estudiados con el S-método y el R-método. Se obtuvo la misma clasificación para cada nuevo objeto que en los estudios precedentes, detectando nuevamente una mayor dificultad para clasificar el último nuevo objeto. Sin embargo, la clasificación proporcionada por el C-método es más sólida que la dada por el S-método. Debido a esto, se concluye que la inclusión de la α -credibilidad ayuda a la toma de decisiones. Además, la α - T -credibilidad de la regla de decisión en la que se apoya la clasificación del nuevo objeto debe ser elevada, ya que en caso contrario se obtendría un grado de α -credibilidad bajo y el C-método descartaría tomar dicha decisión. Como consecuencia, la decisión tomada estará respaldada por una regla en la que podemos confiar en un alto grado. De esta forma, la decisión proporcionada por el C-método queda reforzada.

Por otro lado, este último método presentado no tiene en cuenta el T -soporte de las reglas comparadas con el nuevo objeto, lo cual puede inducir a la toma de una decisión basada en una regla con poca representatividad en la tabla. El último método propuesto también tomará en consideración este indicador de relevancia.

3.3.4. Método de α -credibilidad orientado por el soporte

Finalmente, el cuarto método introducido en [15] es una ligera modificación del presentado en la Sección 3.3.3, incluyendo ahora la representatividad de las reglas comparadas mediante el uso del NT -soporte. Este método se apoya en el grado de α -credibilidad orientado por el soporte.

Definición 3.15. Sean $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ una tabla de decisión, $Dec(S)$ un conjunto de reglas de decisión, $\alpha \in [0, 1)$ y $(\&, I)$ un par adjunto, con $\&$ un conjuntor asociativo. Para cada nuevo objeto $x \in U^c$ y $v \in V_d$, el grado de α -credibilidad orientado por el soporte de la decisión $\bar{d}(x) = v$ de acuerdo a $Dec(S)$ viene dado por la aplicación $\delta_{Dec(S)}^\alpha: U^c \times V_d \rightarrow [0, 1]$ definida como

$$\delta_{Dec(S)}^\alpha(x, v) = \max\{\&(\|\Phi\|_S^T(x), \mu_{S, \tau}^\alpha(\Phi, \Psi), \text{supp}_S^{NT}(\Phi, \Psi)) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

Una vez más, la noción en la que se basa el método a presentar es un grado de verdad, ya que $0 \leq \delta_{Dec(S)}^\alpha(x, v) \leq 1$ para todo nuevo objeto x , decisión $\bar{d}(x) = v$, $\alpha \in [0, 1)$ y conjunto de reglas de decisión $Dec(S)$. El denominado *método de α -credibilidad orientado por el soporte* y abreviado como *D-método*, del inglés *support-driven- α -credibility method*, es presentado a continuación:

- D1. Dado un nuevo objeto para clasificar $x \in U^c$, se calculan los grados de α -credibilidad orientado por el soporte $\delta_{Dec(S)}^\alpha(x, v)$ para todo $v \in V_d$.
- D2. Se comparan los grados obtenidos de α -credibilidad orientado por el soporte. Si existe una decisión $w \in V_d$ tal que $\delta_{Dec(S)}^\alpha(x, w)$ es elevado y suficientemente diferente de $\delta_{Dec(S)}^\alpha(x, v)$, para todo $v \in V_d \setminus \{w\}$, entonces se elige la decisión $\bar{d}(x) = w$ para el objeto x .
- D3. Si existen algunos valores del atributo de decisión con grados de α -credibilidad orientado por el soporte similares, entonces se comprobará la relación entre las variables involucradas en el cálculo. Así, se

puede elegir otro par adjunto ($\&, I$) del marco multiadjunto orientado a propiedades considerado para representar más apropiadamente dicha relación y para obtener una mejor agregación de los valores correspondientes.

El D-método se usó en el Ejemplo 5 de [15] para clasificar los mismos objetos de los ejemplos anteriores y así, comprobar su eficacia. En dicho ejemplo, se muestra la obtención de los mismos resultados, por lo que la clasificación dada por todos los métodos propuestos coincide. Sin embargo, es importante destacar una mayor complejidad para tomar la decisión del segundo objeto. Esto se debe a que las reglas con la decisión tomada tienen un menor *NT*-soporte. Aun así, se aplicó el paso D3 del D-método considerando tres pares adjuntos diferentes, reinterpretando la relación entre el grado de satisfacibilidad, el grado de α -credibilidad y el *NT*-soporte, obteniendo siempre la misma recomendación, por lo que se tomó la decisión obtenida.

La necesidad de aplicar los últimos pasos de los métodos presentados muestran que la base de datos estudiada tiene un alto nivel de inconsistencia o indefinición y que se requiere, independientemente de la decisión tomada, más información sobre los casos estudiados. Por ejemplo, es altamente recomendable la disposición de ejemplos completos que satisfagan en un alto grado los atributos del objeto a clasificar, permitiendo obtener una clasificación inmediata y precisa para el objeto analizado.

En conclusión, cada uno de los métodos de clasificación introducidos afronta el problema de la toma de decisiones desde un punto de vista diferente. Por tanto, todos ellos deben ser tenidos en cuenta cuando se quiera predecir el comportamiento de un nuevo objeto, especialmente cuando alguno de los métodos no proporcione una clasificación suficientemente satisfactoria.

Por otro lado, en la Sección 4 de [15] se llevó a cabo una comparativa

detallada entre el C-método y los métodos introducidos en [41, 91]. En ella, se mostró que todas las nociones usadas en el C-método generalizan a las empleadas en la literatura, ya que estas consideran de manera implícita un umbral y un par adjunto en particular, mientras que en nuestra propuesta se puede usar cualquiera, aportando una gran flexibilidad al estudio. Esto permite también la definición del paso C3, con el que el usuario puede elegir diferentes pares adjuntos para lograr una clasificación satisfactoria. Esta mayor flexibilidad también es palpable en el paso C2, donde el usuario debe valorar si la diferencia entre los diferentes grados de α -credibilidad es suficiente para tomar una decisión.

Finalmente, en la Sección 5 de [15] se llevó a cabo un estudio de clasificación de un nuevo objeto en un conjunto de datos reales. En primer lugar, se definieron las reglas de decisión usadas para clasificar el nuevo objeto y se analizaron mediante los indicadores de relevancia introducidos anteriormente. Tras aplicar cada uno de los métodos de clasificación, se concluye que el R-método proporciona una decisión más concluyente, derivada del detallado estudio que es posible realizar usando dicho método y de la Proposición 3.18 y el Lema 3.1. Además, el D-método es el único que proporciona una clasificación diferente a la dada por el resto de métodos, debido a la inclusión del NT -soporte. Este hecho muestra una gran inconsistencia en la base de datos, ya que dependiendo del punto de vista considerado para clasificar un nuevo objeto se obtienen decisiones diferentes. Por tanto, para llevar a cabo una clasificación precisa es necesario disponer de más objetos, cuyas decisiones se conozcan con total seguridad, similares al que se quiere clasificar.

Capítulo 4

Reductos valor y birreductos: Un estudio comparativo

Título:	Value reducts and bireducts: A comparative study
Autores:	Fernando Chacón-Gómez, M. Eugenia Cornejo, Jesús Medina, Eloísa Ramírez-Poussa
Revista:	Mathematical Methods in the Applied Sciences
Factor de impacto JCR:	2,9. Posición 29 de 267 (Q1)
Año:	2023
Categoría:	Mathematics, Applied
Publicación:	30 enero 2023
DOI:	10.1002/mma.8597

RESEARCH ARTICLE

Value reducts and bireducts: A comparative study

Fernando Chacón-Gómez¹ | M. Eugenia Cornejo² | Jesús Medina³ |
Eloísa Ramírez-Poussa⁴

Department of Mathematics, University of
Cádiz, Cádiz, Spain

Correspondence

M. Eugenia Cornejo, Department of
Mathematics, University of Cádiz, 11510
Cádiz, Spain.
Email: mariaeugenia.cornejo@uca.es

Communicated by: J. Banasiak

Funding information

Department of Economy, Knowledge,
Business and University of the Regional
Government of Andalusia, Grant/Award
Number: FEDER-UCA18-108612;
European Cooperation in Science and
Technology, Grant/Award Number:
CA17124; European Regional
Development Fund, Grant/Award
Number: PID2019-108991GB-I00; State
Research Agency (AEI)

In Rough Set Theory, the notion of bireduct allows to simultaneously reduce the sets of objects and attributes contained in a dataset. In addition, value reducts are used to remove some unnecessary values of certain attributes for a specific object. Therefore, the combination of both notions provides a higher reduction of unnecessary data. This paper is focused on the study of bireducts and value reducts of information and decision tables. We present theoretical results capturing different aspects about the relationship between bireducts and reducts, offering new insights at a conceptual level. We also analyze the relationship between bireducts and value reducts. The studied connections among these notions provide important profits for the efficient information analysis, as well as for the detection of unnecessary or redundant information.

KEYWORDS

bireducts, reducts, rough set theory, value reducts

MSC CLASSIFICATION

06A06, 06A15, 18A32, 03B60

1 | INTRODUCTION

Introduced by Pawlak in the 1980s, Rough Set Theory (RST) is a formal tool whose purpose is modeling and processing incomplete information contained in relational databases.¹ Due to the huge amount of data collected in databases, it is increasingly common to find problems for the management and treatment of the stored information.^{2–6} This fact makes necessary to carry out a preprocessing of data in order to remove redundant information.^{7–11}

Pawlak¹² provided a three-step method for analyzing information tables. The three-step method of rough set analysis addresses three different types of reductions that have to be carried out for the simplification of an information table.¹³ The first step deals with the search of a minimal subset of attributes capable of containing the information of the complete set of attributes. The second step studies the dependencies of the attribute values to simplify decision rules. Finally, the third step analyzes the dependencies between the decision rules in order to eliminate redundancies. All these steps show the relevance of the notion of reduct for the reduction mechanisms of information and decision tables in RST. Indeed, a generic definition of reduct was provided in Yao and Fu.¹³ This generic definition collects the sufficiency and non-redundancy conditions that must be satisfied by a reduct and is capable of accommodating the different interpretations of reduct, provided by each of the steps included in the Pawlak's method.

The notion of reduct, as a minimal set of attributes keeping the original information (associated with the first step mentioned above), has been widely studied within this mathematical theory.^{14–18} In addition, Pawlak¹² also defined value

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Mathematical Methods in the Applied Sciences published by John Wiley & Sons, Ltd.

reducts in RST in order to keep reducing the size of relational databases (associated with the second step). The idea of this notion is to consider the indispensable attributes for each object instead of indispensable attributes for all the objects. Hence, it is possible to disregard also certain values of attributes for each object, while the information of the original table is preserved. However, the notion of value reduct has not been so intensively studied so far. On the other hand, the notion of bireduct goes further and pursues the simultaneous removal of both, objects and attributes, by avoiding incompatibilities, redundancies and removing the existing noise.^{19–25}

This paper aims to know better all the previously mentioned notions at a conceptual level, that is, the notion of reduct, bireduct and value reduct, in order to take advantages the potentiality of these notions for the information analysis. We will see that some bireducts are related to reducts, showing that these bireducts offer the reduction of the first and the third steps commented above. Moreover, we will prove that other bireducts include the value reduct, and so they also provide the second step of the usual reduction of redundant information on a given dataset. Therefore, the computation and analysis of bireducts and value reducts provide relevant knowledge of information tables. We will also analyze bireducts whose subset of objects contains all objects, bireducts whose subset of attributes contains all attributes and bireducts whose subsets of objects or attributes are the empty set. Moreover, we have shown that, given a subset of attributes B of a bireduct, each object in the universe belongs at least to an information bireduct associated with B . The connection that has been established between the notions of reduct and bireduct provides important profits. For example, only from the computation of reducts, we can build sub-tables (bireducts) for the efficient information analysis, removing redundancies in the attributes, as well as in the objects. Furthermore, the study carried out about decision tables also provides important results. One of these results lets us simplify the procedure to obtain decision bireducts, since it allows us to dispense with all the objects belonging to the positive region during the computation process of bireducts, which significantly reduces the number of objects to be considered.

In addition, we will take special attention to the comparison of bireducts and value reducts in order to discover the existing relationships between these notions. This second part will also be very interesting because, although originally both notions were defined independently, we will show remarkable links between them. In particular, we will prove that every value reduct B for an object x is the subset of attributes of a bireduct (X, B) in which x belongs to the subset of objects of the bireduct, $x \in X$. All the notions and results will appear accompanied with examples to make easier their understanding.

Finally, it is convenient to mention that some of the notions considered in this paper are those introduced by Pawlak in his book.¹² However, this paper takes into account a slightly different nomenclature. In particular, the names of “information table” and “decision table” are considered instead of “knowledge representation system,” the name “reduct” refers in this paper to “reduct of knowledge” in Pawlak,¹² and “value reduct” and “ d -value reduct” are used instead of “reduction of categories” and “relative reduct of categories,” respectively.

The paper is organized as follows: preliminary notions of RST, together with an example, are recalled in Section 2. Afterwards, Section 3 presents the notion of bireduct together with different technical properties, as well as its relationship with reducts. Section 4 introduces the relationship between bireducts and value reducts through some results and examples. Finally, conclusions and future works are presented in Section 5.

2 | ROUGH SET THEORY

This section presents the basic definitions of RST which will play a crucial role in our study. First of all, it is convenient to recall that databases are represented as information tables in this framework. Formally, information tables are defined below.

Definition 1. Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. An information table is a tuple $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ such that $\mathcal{V}_{\mathcal{A}} = \{V_a | a \in \mathcal{A}\}$, where V_a is the set of values associated with the attribute a over U , and $\overline{\mathcal{A}} = \{\bar{a} : U \rightarrow V_a | a \in \mathcal{A}\}$.

Next, an equivalence relation is defined on the set of objects of an information table. This equivalence relation will be useful to compare objects according to a given subset of attributes.

Definition 2. Let $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ be an information table. The indiscernibility mapping $I : \mathcal{P}(\mathcal{A}) \rightarrow \mathcal{P}(U \times U)$ is defined, for each $B \subseteq \mathcal{A}$, as the equivalence relation

$$I(B) = \{(x, y) \in U \times U | \bar{a}(x) = \bar{a}(y), \text{ for all } a \in B\}$$

which is called B -indiscernibility relation. Each class of $I(B)$ can be written as $[x]_{I(B)} = \{y \in U \mid (x, y) \in I(B)\}$. The partition determined by $I(B)$ on the set of objects U is denoted as $U/I(B) = \{[x]_{I(B)} \mid x \in U\}$.

Notice that in the particular case of $B = \emptyset$, we have that $I(\emptyset) = U \times U$. Based on an indiscernibility relation, the notions of indiscernible and discernible objects are introduced as follows.

Definition 3. Given an information table $(U, \mathcal{A}, V_{\mathcal{A}}, \overline{\mathcal{A}})$ and a subset of attributes $B \subseteq \mathcal{A}$, we say that $x, y \in U$ are B -indiscernible if $y \in [x]_{I(B)}$. Otherwise, we say x, y are B -discernible.

Now, we define a fundamental notion in the analysis and extraction of information from databases, that is, the notion of reduct. Reducts are introduced to cover the need to reduce information tables, removing their redundant and/or superfluous variables without losing information.

Definition 4. Given an information table $(U, \mathcal{A}, V_{\mathcal{A}}, \overline{\mathcal{A}})$ and a subset of attributes $B \subseteq \mathcal{A}$, we say that:

- $a \in B$ is *dispensable* in B if $I(B) = I(B \setminus \{a\})$. Otherwise, a is *indispensable* in B .
- B is *independent* if all its attributes are indispensable in B .
- $B' \subseteq B$ is a *reduct* of B if B' is independent and $I(B) = I(B')$.

Notice that, a reduct of $\mathcal{A}$ is a subset of attributes that preserves the partition and therefore, it classifies objects in the same way as if we consider the whole set of attributes. Consequently, attributes not belonging to a reduct are unnecessary.

The following notion allows to remove some values of certain attributes, when it is not possible to delete such attributes without altering the information contained in the considered database.

Definition 5. Let $(U, \mathcal{A}, V_{\mathcal{A}}, \overline{\mathcal{A}})$ be an information table, $B \subseteq \mathcal{A}$ and $x \in U$. We say that

- The value of an attribute $a \in B$ is *dispensable* in B for x , if the equality $[x]_{I(B)} = [x]_{I(B \setminus \{a\})}$ holds. Otherwise, the value of a is *indispensable* in B for x .
- B is *independent* for x , if the value of a is indispensable in B for x , for each attribute $a \in B$.
- $B' \subseteq B$ is a *value reduct* of B for x , if B' is independent for x and $[x]_{I(B)} = [x]_{I(B')}$.

A helpful particular case of information table arises when an attribute is highlighted, which is defined below.

Definition 6. Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A decision table is a tuple $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with attribute a over U , and $\overline{\mathcal{A}_d} = \{\tilde{a} : U \rightarrow V_a \mid a \in \mathcal{A}_d\}$. In this case, the attributes of $\mathcal{A}$ are called condition attributes and d is called decision attribute.

In what follows, the notions of dispensable, independent and reduct are translated to the decision table framework. These concepts are based on the well-known notion of positive region which is given below.

Definition 7. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. The positive region of the partition $U/I(\{d\})$ with respect to B is defined as

$$POS_B(\{d\}) = \bigcup_{X \in U/I(\{d\})} B_*(X)$$

where $B_*(X) = \{x \in U \mid [x]_{I(B)} \subseteq X\}$.

Definition 8. Given a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ and a subset of attributes $B \subseteq \mathcal{A}$, we say that

- $a \in B$ is d -dispensable in B , if $POS_B(\{d\}) = POS_{(B \setminus \{a\})}(\{d\})$. Otherwise, attribute a is d -indispensable in B .
- B is d -independent if all its attributes are d -indispensable in B .
- $B' \subseteq B$ is a d -reduct of B if B' is d -independent and the equality $POS_B(\{d\}) = POS_{B'}(\{d\})$ holds.

The dispensable nature of the values of the attributes included in a decision table is studied by using the following definition.

Definition 9. Given a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, a subset of attributes $B \subseteq \mathcal{A}$ and $x \in U$, we say that

- The value of an attribute $a \in B$ is d -dispensable in B for x , if

$$[x]_{I(B)} \subseteq [x]_{I(\{d\})} \quad \text{and} \quad [x]_{I(B \setminus \{a\})} \subseteq [x]_{I(\{d\})}$$

Otherwise, the value of attribute a is d -indispensable in B for x .

- B is d -independent for x , if the value of a is d -indispensable in B for x , for each attribute $a \in B$.
- $B' \subseteq B$ is a d -value reduct of B for x , if B' is d -independent for x and the inclusion $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$ implies that $[x]_{I(B')} \subseteq [x]_{I(\{d\})}$.

This section finishes with an illustrative example in order to clarify the definitions presented previously. Specifically, the following example was introduced in Stawicki et al²⁵ for determining the suitability of a particular day for playing sport according to meteorological conditions.

Example 10. Consider an information table $(U, \mathcal{A}, \mathcal{V}_\mathcal{A}, \overline{\mathcal{A}})$ where the set of objects $U = \{1, 2, \dots, 14\}$ represents days of a certain month and the set of attributes $\mathcal{A} = \{Outlook(O), Temperature(T), Humidity(H), Wind(W)\}$ indicates meteorological conditions. This information table is represented in Table 1.

To begin with, we will see that $\mathcal{A}$ is actually the only reduct of the information table. It is easy to verify that there are no objects with the same values for the attributes of $\mathcal{A}$ and therefore, $[x]_{I(\mathcal{A})} = \{x\}$ for all $x \in U$. However, if we dispense with an attribute, for example we consider day 1 and the subset of attributes $\{O, T, H\}$, we obtain

$$\begin{aligned} [1]_{I(\{O,T,H\})} &= \{y \in U \mid y \text{ are } \{O, T, H\}\text{-indiscernible}\} \\ &= \{y \in U \mid \overline{O}(y) = \overline{O}(1) = \text{sunny}, \\ &\quad \overline{T}(y) = \overline{T}(1) = \text{hot}, \\ &\quad \overline{H}(y) = \overline{H}(1) = \text{high}\} \\ &= \{1, 2\} \end{aligned}$$

Since $[1]_{I(\mathcal{A})} = \{1\} \neq \{1, 2\} = [1]_{I(\{O,T,H\})}$, we can conclude that $I(\mathcal{A}) \neq I(\{O, T, H\})$. Hence, by Definition 4, attribute W is indispensable in $\mathcal{A}$. Following an analogous reasoning, we obtain $[1]_{I(\{O,H,W\})} = \{1, 8\}$ and therefore, attribute T is also indispensable in $\mathcal{A}$. In addition, O is indispensable in $\mathcal{A}$ because $[1]_{I(\{T,H,W\})} = \{1, 3\}$. Notice that, from object 1, we cannot deduce that attribute H be indispensable since $[1]_{I(\mathcal{A})} = \{1\} = [1]_{I(\{O,T,W\})}$. However, from object 3, we have that $[3]_{I(\mathcal{A})} = \{3\} \neq \{3, 13\} = [3]_{I(\{O,T,W\})}$. Then, by Definition 4, attribute H is indispensable in $\mathcal{A}$.

Consequently, each attribute in $\mathcal{A}$ is indispensable for the information processing and we can ensure that $\mathcal{A}$ is the only reduct of the information table. Therefore, we need all the attributes to manage information presented in Table 1. In other words, if we omit any attribute, we lose information from the RST point of view.

Now, we will compute the value reducts corresponding to $\mathcal{A}$ by using Definition 5. For that, we need to eliminate the values of attributes which are dispensable in $\mathcal{A}$ for each object.

For instance, we calculate a value reduct of $\mathcal{A}$ for the object 1. We have that the equality $[1]_{I(\mathcal{A})} = [1]_{I(\{O,T,W\})}$ holds. In addition, $[1]_{I(\{O,T,W\})} \neq [1]_{I(\{O,T,W\} \setminus \{a\})}$ for all $a \in \{O, T, W\}$, since $[1]_{I(\{O,T\})} = \{1, 2\}$, $[1]_{I(\{O,W\})} = \{1, 8, 9\}$ and $[1]_{I(\{T,W\})} = \{1, 3, 13\}$. By Definition 5, the value of the attributes O, T, W is indispensable for the object 1 and $\{O, T, W\}$ is independent for the object 1. Then, $\{O, T, W\}$ is a value reduct of $\mathcal{A}$ for 1. In fact, it is the only value reduct of $\mathcal{A}$ for 1.

TABLE 1 Table associated with the information table $(U, \mathcal{A}, \mathcal{V}_\mathcal{A}, \overline{\mathcal{A}})$ given in Example 10

Day	Outlook	Temperature	Humidity	Wind
1	Sunny	Hot	High	Weak
2	Sunny	Hot	High	Strong
3	Overcast	Hot	High	Weak
4	Rain	Mild	High	Weak
5	Rain	Cool	Normal	Weak
6	Rain	Cool	Normal	Strong
7	Overcast	Cool	Normal	Strong
8	Sunny	Mild	High	Weak
9	Sunny	Cool	Normal	Weak
10	Rain	Mild	Normal	Weak
11	Sunny	Mild	Normal	Strong
12	Overcast	Mild	High	Strong
13	Overcast	Hot	Normal	Weak
14	Rain	Mild	High	Strong

Following an analogous procedure with the rest of objects, we obtain Table 2. The information collected in Table 2 indicates the value reducts corresponding to $\mathcal{A}$ for the considered information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$. From Table 2, we can conclude that O and T are the most important attributes of the information table because its values are indispensable for the most of objects.

Now, we will include a new column in Table 1 in order to obtain a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ whose decision attribute is $\{d\} = \{Play\}$. The decision table is represented in Table 3 and it is studied below.

As it was shown previously, each pair of objects are $\mathcal{A}$ -discernible and so, we deduce the equality $POS_{\mathcal{A}}(\{d\}) = U$. Now, we will prove that $\{O, H, W\}$ is a d -reduct of $\mathcal{A}$. First of all, we compute $U/I(\{O, H, W\})$, obtaining $[1]_{I(\{O, H, W\})} = \{1, 8\}$, $[5]_{I(\{O, H, W\})} = \{5, 10\}$ and the rest of equivalence classes are the trivial classes. On the other hand, the set $U/I(\{d\})$ is composed of the classes $[1]_{I(\{d\})} = \{1, 2, 6, 8, 14\}$ and $[3]_{I(\{d\})} = \{3, 4, 5, 7, 9, 10, 11, 12, 13\}$, that is,

$$U/I(\{d\}) = \{[1]_{I(\{d\})}, [3]_{I(\{d\})}\}$$

As a consequence, we obtain

$$\begin{aligned} POS_{\{O, H, W\}}(\{d\}) &= \{x \in U \mid [x]_{I(\{O, H, W\})} \subseteq [x]_{I(\{d\})}\} = U \\ POS_{\{O, H\}}(\{d\}) &= [1]_{I(\{O, H\})} \cup [3]_{I(\{O, H\})} \cup [7]_{I(\{O, H\})} \cup [9]_{I(\{O, H\})} \\ &= \{1, 2, 3, 7, 8, 9, 11, 12, 13\} \\ POS_{\{O, W\}}(\{d\}) &= [3]_{I(\{O, W\})} \cup [4]_{I(\{O, W\})} \cup [6]_{I(\{O, W\})} \cup [7]_{I(\{O, W\})} \\ &= \{3, 4, 5, 6, 7, 10, 12, 13, 14\} \\ POS_{\{H, W\}}(\{d\}) &= [5]_{I(\{H, W\})} \\ &= \{5, 9, 10, 13\} \end{aligned}$$

Day	Outlook	Temperature	Humidity	Wind
1	Sunny	Hot	-	Weak
2	-	Hot	-	Strong
3	Overcast	Hot	High	-
4	Rain	-	High	Weak
5	Rain	Cool	-	Weak
6	Rain	Cool	-	Strong
7	Overcast	Cool	-	-
8	Sunny	Mild	High	-
9	Sunny	Cool	-	-
10	Rain	Mild	Normal	-
11	Sunny	Mild	Normal	-
12	Overcast	Mild	-	-
13	-	Hot	Normal	-
14	Rain	Mild	-	Strong

TABLE 2 Value reducts corresponding to $\mathcal{A}$ of the information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ given in Example 10

Day	Outlook	Temperature	Humidity	Wind	Play
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

TABLE 3 Table associated with the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 10

TABLE 4 d -value reducts corresponding to $\{O, H, W\}$ of the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 10

Day	Outlook	Temperature	Humidity	Wind	Play
1	Sunny	-	High	-	No
2	Sunny	-	High	-	No
3	Overcast	-	-	-	Yes
4	Rain	-	-	Weak	Yes
5	Rain	-	-	Weak	Yes
6	Rain	-	-	Strong	No
7	Overcast	-	-	-	Yes
8	Sunny	-	High	-	No
9	Sunny	-	Normal	-	Yes
10	Rain	-	-	Weak	Yes
11	Sunny	-	Normal	-	Yes
12	Overcast	-	-	-	Yes
13	Overcast	-	-	-	Yes
14	Rain	-	-	Strong	No

By Definition 8, each attribute in $\{O, H, W\}$ is d -indispensable and $\{O, H, W\}$ is d -independent. Consequently, $\{O, H, W\}$ is a d -reduct of $\mathcal{A}$. This fact implies that we can eliminate all values of *Temperature* for each object without losing information. Hence, we can decide if a day is suitable to play only by using *Outlook*, *Humidity* and *Wind*.

Next step is studying d -value reducts of $\{O, H, W\}$, since it is a d -reduct of $\mathcal{A}$, for all $x \in U$ by using Definition 9, which clearly also are d -value reducts of $\mathcal{A}$. For instance, we will show $\{O, H\}$ is a d -value reduct of $\{O, H, W\}$ for the object 1. According to the information collected in Table 3, we obtain

- $[1]_{I(\{O, H, W\})} = \{1, 8\} \subseteq [1]_{I(\{d\})} = \{1, 2, 6, 8, 14\}$.
- $[1]_{I(\{O, H\})} = \{1, 2, 8\} \subseteq [1]_{I(\{d\})}$.
- $[1]_{I(\{O\})} = \{1, 2, 8, 9, 11\} \not\subseteq [1]_{I(\{d\})}$.
- $[1]_{I(\{H\})} = \{1, 2, 3, 4, 8, 12, 14\} \not\subseteq [1]_{I(\{d\})}$.

Applying Definition 9, we have that the value of attributes O, H is d -indispensable and $\{O, H\}$ is d -independent. Therefore, $\{O, H\}$ is a d -value reduct of $\{O, H, W\}$ for the object 1. Hence, we can decide if a day is suitable to play by means of the values of *Outlook* and *Humidity* for the object 1. Following an analogous procedure with the rest of objects, we obtain Table 4, which indicates some d -value reducts corresponding to $\{O, H, W\}$ for the considered decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$. In addition, it is convenient to emphasize that there may exist more than one d -value reduct for an object. For instance, as Table 4 shows, $\{O, H\}$ is a d -value reduct of $\{O, H, W\}$ for the object 9, and it is easy to see that $\{H, W\}$ is a d -value reduct of $\{O, H, W\}$ for this object too.

Therefore, value reducts offer a different point of view of the relationships among the attributes, which is very useful to obtain an extra level of reduction and only consider the strictly necessary attributes and values in the computation, such as in the determination of decision rules.

3 | BIREDUCTS. PROPERTIES AND THE RELATIONSHIP WITH REDUCTS

Bireducts arise as an extension of the classical notions of reducts in RST providing an extra flexibility for operating with subsets of attributes and subsets of objects from which those attributes can be portrayed efficiently.^{19,22–25} Specifically, bireducts are used to reduce the number of attributes by avoiding the occurrence of incompatibilities and removing the existing noise in tabular databases.

This section introduces diverse properties of the different notions of bireduct: information bireduct, decision bireduct and U -decision bireduct, and the relationship with reducts, which will be illustrated with examples. These results complement the ones given in the initial study presented in Stawicki et al.²⁵

3.1 | Information bireducts

Information bireducts generalize information reducts by operating with subsets of attributes and subsets of objects.^{19,23,25} Specifically, information bireducts allow to distinguish objects by the subset of considered attributes, so they are good subrepresentations of the information table. In this section, an information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ will be fixed. The notion of information bireduct is formally defined as follows.

Definition 11. Given $X \subseteq U$ and $B \subseteq \mathcal{A}$, we say that

- B is X -irreducible if there is no $B' \subset B$ such that all pairs $x, y \in X$ are B' -discernible.
- X is B -inextensible if there is no $X' \subset U$ with $X \subset X'$ such that all pairs $x, y \in X'$ are B -discernible.
- The pair (X, B) is an information bireduct if all pairs $x, y \in X$ are B -discernible, B is X -irreducible and X is B -inextensible.

Notice that information bireducts generate information subtables with only different rows (all objects are B -discernible). The following technical property guarantees that, given $B \subseteq \mathcal{A}$, if there exist information bireducts in the information table with B as subset of attributes, then each object of the universe belongs to some information bireduct.

Proposition 12. Given a subset of attributes $B \subseteq \mathcal{A}$ and the subsets of objects $X_1, \dots, X_n \subseteq U$ such that the pairs $(X_1, B), \dots, (X_n, B)$ are all the information bireducts, then $U = \bigcup_{i=1}^n X_i$.

Proof. It is clear that $\bigcup_{i=1}^n X_i \subseteq U$. We will prove the other inclusion by reductio ad absurdum. Given $y \in U$, we will suppose that $y \notin \bigcup_{i=1}^n X_i$. Therefore, $y \notin X_i$ for all $i \in \{1, \dots, n\}$. Consequently, without loss of generality, we can suppose that there exists $x_1 \in X_1$ such that x_1, y are B -indiscernible. Hence, $(X_1 \setminus \{x_1\} \cup \{y\}, B)$ is an information bireduct which leads us to a contradiction, since by hypothesis $(X_1, B), \dots, (X_n, B)$ are all information bireducts with B as subset of attributes and $X_i \neq X_1 \setminus \{x_1\} \cup \{y\}$ for all $i \in \{1, \dots, n\}$ because $y \notin \bigcup_{i=1}^n X_i$. As a consequence, $U \subseteq \bigcup_{i=1}^n X_i$.

In conclusion, $U = \bigcup_{i=1}^n X_i$. □

The notion of information bireduct as well as the last property will be applied to a particular case next.

Example 13. Consider the information table given in Example 10 and the subset of attributes $B = \{O, T, W\} \subseteq \mathcal{A}$ in order to compute all information bireducts. First of all, we will calculate what objects are B -indiscernible. Taking into account the information collected in Table 1 and applying Definition 2, we obtain

$$\begin{aligned} U/I(B) = \{ & [1]_{I(B)}, [2]_{I(B)}, [3]_{I(B)}, [4]_{I(B)}, [5]_{I(B)}, \\ & [6]_{I(B)}, [7]_{I(B)}, [8]_{I(B)}, [9]_{I(B)}, [11]_{I(B)}, \\ & [12]_{I(B)}, [14]_{I(B)} \} \end{aligned}$$

where $[3]_{I(B)} = \{3, 13\}$, $[4]_{I(B)} = \{4, 10\}$ and the other classes are the trivial classes.

Therefore, according to Definition 3, the pairs of objects 3, 13 and 4, 10 are B -indiscernible and the other pairs of objects are B -discernible. Notice that, removing from U an object of each pair 3, 13 and 4, 10, we obtain the following sets $U \setminus \{10, 13\}$, $U \setminus \{3, 10\}$, $U \setminus \{4, 13\}$ and $U \setminus \{3, 4\}$, which are all B -inextensible subsets of objects by construction. Consequently, by Definition 11, the candidates to information bireducts are the pairs $(U \setminus \{10, 13\}, B)$, $(U \setminus \{3, 10\}, B)$, $(U \setminus \{4, 13\}, B)$ and $(U \setminus \{3, 4\}, B)$.

On the other hand, B must be X -irreducible, with $X \in \{U \setminus \{10, 13\}, U \setminus \{3, 10\}, U \setminus \{4, 13\}, U \setminus \{3, 4\}\}$, in order to satisfy the conditions required in the notion of information bireduct. It can be seen immediately that the objects 1, 2 are $\{O, T\}$ -indiscernible, the objects 1, 8 are $\{O, W\}$ -indiscernible and the objects 5, 9 are $\{T, W\}$ -indiscernible. Hence, for every of the four possibilities for X , there is no $B' \subset B$ such that all pairs $x, y \in X$ are B' -discernible, that is, B is X -irreducible. Therefore, we can ensure that $(U \setminus \{10, 13\}, B)$, $(U \setminus \{3, 10\}, B)$, $(U \setminus \{4, 13\}, B)$ and $(U \setminus \{3, 4\}, B)$ are all the information bireducts.

Finally, according to Proposition 12, it is straightforwardly obtained that each object belongs to some information bireduct, that is,

$$U = U \setminus \{10, 13\} \cup U \setminus \{3, 10\} \cup U \setminus \{4, 13\} \cup U \setminus \{3, 4\}$$

Therefore, the set of information bireducts with a fixed subset of attributes B allows the division of information tables into subtables, which are non-redundant and each object is represented in at least one of these subtables, so no information is lost. In this way, it is possible to rebuild the original table by using the generated subtables. Moreover, an important

TABLE 5 Table associated with $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ given in Example 15

Person	Height	Age	Gender
1	Medium	Young	Woman
2	Medium	Young	Woman
3	Medium	Young	Man
4	Short	Old	Woman

particular case is when the subset of attributes B is a reduct of $\mathcal{A}$. In this case, each information subtable contains the same information that the original table, providing a great representation of the database.

Next, two particular cases of information bireducts are introduced. The first one was already proved in Benítez-Caballero et al,²⁶ and it relates the notion of reduct to the notion of information bireduct.

Proposition 14 (Benítez-Caballero et al.²⁶). *Given $B \subseteq \mathcal{A}$, if the pair (U, B) is an information bireduct, then B is a reduct of $\mathcal{A}$.*

Since (U, B) is an information bireduct, we can conclude that B distinguishes all the objects of U and for each subset $B' \subseteq B$ there exists $x_{B'}, y_{B'} \in U$ such that they are B' -indiscernible. As a consequence, B is a reduct of $\mathcal{A}$. However, if B is a reduct, there may have objects that B is not able to distinguish because $\mathcal{A}$ do not distinguish them. Therefore, the reciprocal of the last property is not true in general, as it is illustrated in the following example.

Example 15. Let $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ be the information table represented in Table 5, where the set of objects $U = \{1, 2, 3, 4\}$ represents patients, the set of attributes is given by $\mathcal{A} = \{\text{Height}(h), \text{Age}(a), \text{Gender}(g)\}$.

From Table 5 it is easy to see that the equality $I(\{h, g\}) = I(\mathcal{A})$ holds, since $U/I(\mathcal{A}) = U/I(\{h, g\}) = \{[1]_{I(\mathcal{A})}, [3]_{I(\mathcal{A})}, [4]_{I(\mathcal{A})}\}$ where $[1]_{I(\mathcal{A})} = \{1, 2\}$ and the other classes are the trivial classes. Notice that $[1]_{I(\{h\})} = \{1, 2, 3\} \neq [1]_{I(\{h, g\})}$ and $[1]_{I(\{g\})} = \{1, 2, 4\} \neq [1]_{I(\{h, g\})}$, then attributes h and g are indispensable. Therefore, by using Definition 4, we conclude that $\{h, g\}$ is a reduct of $\mathcal{A}$. Now, we will show that $(U, \{h, g\})$ is not an information bireduct, that is, the counterpart of Proposition 14 is not satisfied.

The pair of objects 1,2 are $\{h, g\}$ -indiscernible whereas the other pairs of objects are $\{h, g\}$ -discernible. As a consequence, all information bireducts with $\{h, g\}$ as subset of attributes are $(\{1, 3, 4\}, \{h, g\})$ and $(\{2, 3, 4\}, \{h, g\})$. Therefore, $(U, \{h, g\})$ is not an information bireduct.

As it is showed in the previous example, reducts play an important role in the calculus of information bireducts. Now, we present the result that summarizes this property.

Proposition 16. *Given $B \subseteq \mathcal{A}$, if B is a reduct of $\mathcal{A}$ then there exists $X \subseteq U$ such that (X, B) is an information bireduct.*

Proof. Suppose that B is a reduct of $\mathcal{A}$. Consider a subset of objects $X \subseteq U$ such that each pair of objects $x, y \in X$ are B -discernible and X is B -inextensible. On the other hand, since B is a reduct of $\mathcal{A}$, for each $B' \subset B$ there exist $x_{B'}, y_{B'} \in X$ B' -discernible and B' -indiscernible. As a consequence, B is X -irreducible and (X, B) is an information bireduct.

Thanks to this result, it is possible to relate the notion of information bireduct to the notion of reduct. Therefore, if a reduct B of an information table is known, a non-redundant subtable can be built by using this subset of attributes and which contains the same information as the original table. Thus, the reduction given by a bireduct associated with a reduct is equivalent to the first and the third steps of Pawlak three-step rough set analysis.

In the next result, we study information bireducts whose subset of attributes is the total set $\mathcal{A}$, obtaining a double equivalence between information bireducts and reducts of $\mathcal{A}$. $\square$

Corollary 17. *There exists $X \subseteq U$ such that $(X, \mathcal{A})$ is an information bireduct if and only if $\mathcal{A}$ is a reduct of $\mathcal{A}$.*

Proof. Suppose that $(X, \mathcal{A})$ is an information bireduct. Hence, each pair $x, y \in X$ are $\mathcal{A}$ -discernible. We will prove that $\mathcal{A}$ is a reduct of $\mathcal{A}$ by reductio ad absurdum. Suppose that there exists $B \subset \mathcal{A}$ being B a reduct of $\mathcal{A}$. Then, $I(B) = I(\mathcal{A})$. Therefore, each pair $x, y \in X$ are B -discernible, obtaining a contradiction because $\mathcal{A}$ is X -irreducible.

The counterpart is obtained from Proposition 16. $\square$

From Corollary 17, we can conclude that the reciprocal of Proposition 16 is only satisfied when $B = \mathcal{A}$.

Next result studies information bireducts whose subset of attributes is the empty set, that is, there are no attribute to distinguish the elements in the subset of objects.

Proposition 18. *Let (X, B) be an information bireduct, with $X \subseteq U$ and $B \subseteq \mathcal{A}$. Then $B = \emptyset$ if and only if $X = \{x\}$ for all $x \in U$.*

Proof. Suppose $B = \emptyset$ and there exist $x, y \in X$ such that $x \neq y$. By hypothesis $(X, \emptyset)$ is an information bireduct, then we obtain that x, y are $\emptyset$ -discernible. This fact lead us to a contradiction, since each pair of objects in U are $\emptyset$ -indiscernible. Consequently, $X = \{x\}$ for all $x \in U$.

Now, we prove the counterpart supposing that $X = \{x\}$ for all $x \in U$. By hypothesis (X, B) is an information bireduct, then X is B -inextensible and we obtain that x, y are B -indiscernible for all $y \in U$. Therefore, for every $B' \subseteq B$, we have that x, y are B' -indiscernible for all $y \in U$. Since B is X -irreducible, it implies that $B = \emptyset$. $\square$

Consequently, when there exists a subset of attributes $B \subseteq \mathcal{A}$ that cannot discern any pair of objects, we cannot find any information bireducts with any subset $B' \subseteq B$ as subset of attributes, with the exception of the empty set. If the subset of attributes considered is the empty set, which cannot discern objects, then each object must be taken individually.

The study of particular cases of information bireducts is concluded remarking that there are no information bireducts whose subset of objects is the empty set. This occurs because we can always consider at least one object, as it is shown in Proposition 18.

This section finishes highlighting a feature of Definition 11 which could be strange in some practical examples. Notice that, if there exist $x, y \in U$ such that $[x]_{I(\mathcal{A})} = [y]_{I(\mathcal{A})}$ then x, y are B -indiscernible, for all $B \subseteq \mathcal{A}$, as a consequence x, y cannot belong to the same information bireduct (X, B) . However, the choice of B has not caused x, y to be B -indiscernible due to they are $\mathcal{A}$ -indiscernible. Since this fact is independent of B , it has no sense considering these objects to compute the information bireducts. This fact would lead us to obtain two information bireducts whose only difference is that one of them contains the object x and the other one the object y , being in essence exactly the same information provided by both bireducts. In order to avoid this redundancy, another type of bireducts defined in information tables could be considered taking into account this fact. For that, next definition proposes to include all $\mathcal{A}$ -indiscernible objects in equivalence classes so they can belong to a same information bireduct.

Definition 19. Let $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ be an information table. An information table reduced by classes is a tuple $(U^*, \mathcal{A}_s, \mathcal{V}_{\mathcal{A}_s}, \overline{\mathcal{A}}_s)$ such that $U^* = \{[x]_{I(\mathcal{A})} | x \in U\}$, $\mathcal{A}_s = \mathcal{A} \cup \{s\}$ with $s \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_s} = \{V_a | a \in \mathcal{A}_s\}$ being V_a the set of values associated with the attribute a over U^* , and $\overline{\mathcal{A}}_s$ is the set which collects the mappings associated with each $a \in \mathcal{A}$ defined as

$$\begin{aligned} \overline{a}^* : U^* &\rightarrow V_a \\ [x]_{I(\mathcal{A})} &\mapsto \tilde{a}(x) \end{aligned}$$

and the mapping associated with the attribute s defined as

$$\begin{aligned} \bar{s} : U^* &\rightarrow V_s \\ [x]_{I(\mathcal{A})} &\mapsto \text{card}([x]_{I(\mathcal{A})}) \end{aligned}$$

Notice that the attribute s provides the number of objects $\mathcal{A}$ -indiscernible with x and $V_s \subset \mathbb{N}$.

Following the same philosophy that the last two results, we present the following properties of information tables reduced by classes.

Proposition 20. *Let $(U^*, \mathcal{A}_s, \mathcal{V}_{\mathcal{A}_s}, \overline{\mathcal{A}}_s)$ be an information table reduced by classes and $B \subseteq \mathcal{A}$. The pair (U^*, B) is an information bireduct if and only if B is a reduct of $\mathcal{A}$.*

Proof. Suppose that B is a reduct of $\mathcal{A}$. Given $x^*, y^* \in U^*$, by the definition of U^* , we have that x^*, y^* are $\mathcal{A}$ -discernible. Since $I(\mathcal{A}) = I(B)$ because B is a reduct, then we deduce that x^*, y^* are B -discernible. Therefore, the pair (U^*, B) is an information bireduct.

The counterpart is obtained from Proposition 14 taking into account that the elements of U^* are the classes $[x]_{I(\mathcal{A})}$, for all $x \in U$. $\square$

The last property shows that the information tables considered in Benítez-Caballero et al²⁶ are actually information tables reduced by classes.

Proposition 21. *Let $(U^*, \mathcal{A}_s, \mathcal{V}_{\mathcal{A}_s}, \overline{\mathcal{A}}_s)$ be an information table reduced by classes and (X, B) be an information bireduct, with $X \subseteq U^*$, $B \subseteq \mathcal{A}$. Then $B = \emptyset$ if and only if $X = [x]_{I(\mathcal{A})}$ for all $x \in U$.*

Proof. The proof is obtained directly by Proposition 18, taking into account that the elements of U^* are the classes $[x]_{I(\mathcal{A})}$, for all $x \in U$. $\square$

The previous properties and examples highlight the need of studying information bireducts for a useful management of information in data analysis. Now, we carry out the study of decision bireducts following a similar structure to this section.

3.2 | Decision bireducts

Decision bireducts are useful to analyze if a group of different subsets of attributes classifies suitably a subset of objects in an easier way.^{19,24,25} Now, a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ will be fixed. The formal notion of decision bireduct together with an illustrative example are given below.

Definition 22. Given $X \subseteq U$ and $B \subseteq \mathcal{A}$, we say that the pair (X, B) is a decision bireduct if every pair $x, y \in X$ is B -discernible when they are d -discernible, B is X -irreducible and X is B -inextensible with respect to this property.

It is important to emphasize that decision bireducts generate decision subtables where there are no contradictions. Therefore, this notion is very useful for decision-making. Next, we illustrate this notion in the following example.

Example 23. Consider the decision table given in Example 10 and the subset of attributes $B = \{O, T, H\} \subseteq \mathcal{A}$ in order to compute all decision bireducts with subset of attributes B . Taking into account the information contained in Table 3 and applying Definition 2, we obtain

$$U/I(B) = \{[1]_{I(B)}, [3]_{I(B)}, [4]_{I(B)}, [5]_{I(B)}, [7]_{I(B)}, [8]_{I(B)}, [9]_{I(B)}, [10]_{I(B)}, [11]_{I(B)}, [12]_{I(B)}, [13]_{I(B)}\}$$

where $[1]_{I(B)} = \{1, 2\}$, $[4]_{I(B)} = \{4, 14\}$, $[5]_{I(B)} = \{5, 6\}$ and the other classes are the trivial classes. Recall that

$$U/I(\{d\}) = \{[1]_{I(\{d\})}, [3]_{I(\{d\})}\}$$

where $[1]_{I(\{d\})} = \{1, 2, 6, 8, 14\}$, $[3]_{I(\{d\})} = \{3, 4, 5, 7, 9, 10, 11, 12, 13\}$. Therefore, it is easy to see that the pairs of objects 4,14 and 5,6 are the only d -discernible and B -indiscernible objects simultaneously. Hence, these pairs of objects cannot appear in the same decision bireduct.

If we withdraw an object of each pair 4,14 and 5,6 from U , we obtain the sets $U \setminus \{6, 14\}$, $U \setminus \{5, 14\}$, $U \setminus \{4, 6\}$ and $U \setminus \{4, 5\}$ which are B -inextensible by construction. Moreover, all objects included in these sets are B -discernible when they are d -discernible. As a result, by Definition 22, the candidates to decision bireducts are $(U \setminus \{6, 14\}, B)$, $(U \setminus \{5, 14\}, B)$, $(U \setminus \{4, 6\}, B)$ and $(U \setminus \{4, 5\}, B)$. In order to guarantee that they are actually decision bireducts, we need to see that B is X -irreducible, being $X \in \{U \setminus \{6, 14\}, U \setminus \{5, 14\}, U \setminus \{4, 6\}, U \setminus \{4, 5\}\}$.

From Table 3 and the classes of the quotient sets $U/I(\{d\})$ and $U/I(B)$, we can ensure that

- The pairs of objects 8,11 and 10,14 are B -discernible, d -discernible and $\{O, T\}$ -indiscernible.
- The pairs of objects 1,3; 2,3; 4,8; 8,12; 6,7; 6,9 and 12,14 are B -discernible, d -discernible and $\{T, H\}$ -indiscernible.
- The pair of objects 6,10 are B -discernible, d -discernible and $\{O, H\}$ -indiscernible.

With respect to the two first items, we have that at least one pair appears in all candidates to decision bireduct. Therefore, attributes H and O cannot be removed from any candidate. In addition, the pair of objects 6,10 only appear in candidates $(U \setminus \{5, 14\}, B)$ and $(U \setminus \{4, 5\}, B)$, so that it is not possible to eliminate the attribute T in $(U \setminus \{5, 14\}, B)$ and $(U \setminus \{4, 5\}, B)$.

Consequently, B is X -irreducible only for the pairs $(U \setminus \{5, 14\}, B)$ and $(U \setminus \{4, 5\}, B)$ and then, they are the only decision bireducts whose subset of attributes is B . We also have that $(U \setminus \{6, 14\}, \{O, H\})$, and $(U \setminus \{4, 6\}, \{O, H\})$ are decision bireducts.

It is important to mention that Proposition 12 is not satisfied by decision bireducts. In fact, the previous example is a counterexample. Specifically, we have that object 5 does not belong to any decision bireduct with B as subset of attributes.

On the other hand, next result makes easier the computation of decision bireducts thanks to the notion of positive region.

Proposition 24. Given $X \subseteq U$ and $B \subseteq \mathcal{A}$, if the pair (X, B) is a decision bireduct then $POS_B(\{d\}) \subseteq X$.

Proof. Let (X, B) be a decision bireduct. We will prove that $POS_B(\{d\}) \subseteq X$ by reductio ad absurdum. Let $x \in POS_B(\{d\})$ such that $x \notin X$. As (X, B) is a decision bireduct, the subset X is B -inextensible. Then, there exists $y \in X$ such that x, y are d -discernible and B -indiscernible. As a consequence, $[x]_{I(B)} \not\subseteq [x]_{I(\{d\})}$. Therefore, $x \notin POS_B(\{d\})$, obtaining a contradiction. $\square$

This result allows to simplify the procedure to obtain decision bireducts, since the positive region is very useful to the construction of the subset of objects X . On the other hand, it is important to emphasize that the inclusion is usually strict, that is $POS_B(\{d\}) \neq X$. For instance, in Example 23, $(U \setminus \{5, 14\}, B)$ and $(U \setminus \{4, 5\}, B)$ are all decision bireducts with $B = \{O, T, H\}$ as subset of attributes. In addition, from the computation of $U/I(\{d\})$ and $U/I(B)$ it is easy to see that $POS_B(\{d\}) = U \setminus \{4, 5, 6, 14\}$. Therefore, in both cases, $POS_B(\{d\}) \neq X$.

Following the same scheme of the previous section, we introduce a result which relates the definitions of decision bireduct and d -reduct.

Proposition 25. *Let $B \subseteq \mathcal{A}$ be a subset of attributes. If the pair (U, B) is a decision bireduct then $POS_B(\{d\}) = U$. Moreover, B is a d -reduct of $\mathcal{A}$.*

Proof. First of all, we will prove that $POS_B(\{d\}) = U$. Clearly, $POS_B(\{d\}) \subseteq U$. Now, given $x \in U$ and $y \in [x]_{I(B)} \subseteq U$, we have that x, y are B -indiscernible. Since (U, B) is a decision bireduct, then x, y are also d -indiscernible. Therefore, $y \in [x]_{I(\{d\})}$. Consequently, $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$ which implies that $x \in POS_B(\{d\})$. Hence, the equality $POS_B(\{d\}) = U$ holds.

Furthermore, we must prove that B is a d -reduct of $\mathcal{A}$. First of all, notice that $POS_{\mathcal{A}}(\{d\}) = U$ since $POS_B(\{d\}) \subseteq POS_{\mathcal{A}}(\{d\})$. Now, we will prove the irreducibility condition by reductio ad absurdum. We suppose that there exists $B' \subset B$ such that $POS_{B'}(\{d\}) = POS_B(\{d\}) = U$. Since (U, B) is a decision bireduct, we obtain that for each $B' \subset B$ there exist $x_{B'}, y_{B'} \in U$ such that they are d -discernible, B -discernible and B' -indiscernible. As a consequence, $x_{B'}, y_{B'} \in POS_B(\{d\})$ but $x_{B'}, y_{B'} \notin POS_{B'}(\{d\})$. Hence, $POS_B(\{d\}) \neq POS_{B'}(\{d\})$, obtaining a contradiction. As a result, B is d -independent and therefore B is a d -reduct of $\mathcal{A}$. $\square$

The following example shows that B be a d -reduct of $\mathcal{A}$ is not enough to ensure that (U, B) is a decision bireduct.

Example 26. We consider the information table of Example 15, and we add a decision attribute $\{d\} = \{\text{Test}(t)\}$ which represents a medical test. This decision table is represented in Table 6.

Consider the subset of attributes $B = \{h, g\}$. It is easy to see that

$$U/I(\{d\}) = \{[1]_{I(\{d\})}, [2]_{I(\{d\})}\} \\ U/I(B) = \{[1]_{I(B)}, [3]_{I(B)}, [4]_{I(B)}\}$$

where $[1]_{I(\{d\})} = \{1, 3\}$, $[2]_{I(\{d\})} = \{2, 4\}$, $[1]_{I(B)} = \{1, 2\}$ and the other classes are the trivial classes. In addition, in Example 15, we showed that $U/I(\mathcal{A}) = U/I(B)$. Hence, it is obtained that $POS_{\mathcal{A}}(\{d\}) = POS_B(\{d\}) = \{3, 4\}$. On the other hand, it is easy to see that $POS_{\{h\}}(\{d\}) = \{4\}$ and $POS_{\{g\}}(\{d\}) = \{3\}$, so both attributes are d -indispensable in B , and therefore, B is a d -reduct of $\mathcal{A}$. However, since $POS_B(\{d\}) = \{3, 4\} \neq U$, by Proposition 25, we obtain that (U, B) is not a decision bireduct. Indeed, in this case, there is no subset of objects $X \subseteq U$ such that (X, B) is a decision bireduct, as we show next.

We have that the pair of objects 1,2 are B -indiscernible and d -discernible, whereas the other pairs of objects are B -discernible if they are d -discernible. As a consequence, the only candidates to decision bireduct with B as subset of attributes are $(\{1, 3, 4\}, B)$ and $(\{2, 3, 4\}, B)$. However, it is easy to see that B is not X -irreducible with $X \in \{\{1, 3, 4\}, \{2, 3, 4\}\}$ because $(\{1, 3, 4\}, \{h\})$ and $(\{2, 3, 4\}, \{g\})$ are decision bireducts. Therefore, since B is a d -reduct of $\mathcal{A}$ and does not exist $X \subseteq U$ such that (X, B) is a decision bireduct, we have that Proposition 16 is not true for decision tables.

	Height	Age	Gender	Test
1	Medium	Young	Woman	+
2	Medium	Young	Woman	−
3	Medium	Young	Man	+
4	Short	Old	Woman	−

TABLE 6 Table associated with the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 26

TABLE 7 Table associated with the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 26 modified

	Height	Age	Gender	Test
1	Medium	Young	Woman	+
2	Medium	Young	Woman	−
3	Medium	Young	Man	+
4	Short	Old	Woman	−
5	Medium	Old	Woman	+
6	Medium	Old	Woman	−

In addition, we will show that Corollary 17 is not true in decision tables. With this purpose, we add two objects to the original table. This decision table is represented in Table 7.

We will show that $(\{2, 3, 4, 5\}, \mathcal{A})$ is a decision bireduct. From Table 7, it is easy to see that all elements in $\{2, 3, 4, 5\}$ are $\mathcal{A}$ -discernible if they are d -discernible. In addition, $\{2, 3, 4, 5\}$ is $\mathcal{A}$ -inextensible because the pairs of objects 1, 2 and 5, 6 are d -discernible and $\mathcal{A}$ -indiscernible.

Now, we will prove that $\mathcal{A}$ is $\{2, 3, 4, 5\}$ -irreducible. On the one hand, the pair of objects 2, 3 are $\{h, a\}$ -indiscernible and d -discernible, so that the attribute g is necessary to distinguish these objects. On the other hand, objects 4, 5 are $\{a, g\}$ -indiscernible and d -discernible, so that the attribute h is also necessary. Finally, the objects 2, 5 are $\{h, g\}$ -indiscernible and d -discernible, so the attribute a is also necessary. As a consequence, $\mathcal{A}$ is $\{2, 3, 4, 5\}$ -irreducible and therefore $(\{2, 3, 4, 5\}, \mathcal{A})$ is a decision bireduct.

To conclude, we will show that $\mathcal{A}$ is not a d -reduct of $\mathcal{A}$. As we have mentioned, the pairs of objects 1, 2 and 5, 6 are the only pairs of objects d -discernible and $\mathcal{A}$ -indiscernible. Therefore, $POS_{\mathcal{A}}(\{d\}) = \{3, 4\}$. On the other hand, as $[3]_{I(\{h,g\})} = \{3\}$ and $[4]_{I(\{h,g\})} = \{4\}$ and $POS_{\{h,g\}}(\{d\}) \subseteq POS_{\mathcal{A}}(\{d\})$ it is obtained that $POS_{\{h,g\}}(\{d\}) = POS_{\mathcal{A}}(\{d\}) = \{3, 4\}$. Finally, we obtain that $3 \notin POS_{\{h\}}(\{d\})$ due to $[3]_{I(\{h\})} = U \setminus \{4\}$. As a consequence, the attribute g is indispensable in $\{h, g\}$. For the same reason, $4 \notin POS_{\{g\}}(\{d\})$ and the attribute h is indispensable in $\{h, g\}$. Therefore, $\{h, g\}$ is a d -reduct of $\mathcal{A}$. Therefore, there exists $X \subseteq U$ such that $(X, \mathcal{A})$ is a decision bireduct and $\mathcal{A}$ is not d -reduct of $\mathcal{A}$, and so Corollary 17 is not true in decision tables.

The next result recalls that, if a decision table satisfies that $POS_{\mathcal{A}}(\{d\}) = U$ then the reciprocal of Proposition 25 is true, as it is showed in Stawicki et al.²⁵

Proposition 27 (Stawicki et al.²⁵). *Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table such that $POS_{\mathcal{A}}(\{d\}) = U$ and $B \subseteq \mathcal{A}$. Then, the pair (U, B) is a decision bireduct if and only if B is a d -reduct of $\mathcal{A}$.*

Notice that, the decision table given in Example 10 satisfies the equality $POS_{\mathcal{A}}(\{d\}) = U$. Taking into account the set of attributes $B = \{O, H, W\}$, we obtained that $\{O, H, W\}$ is a d -reduct of $\mathcal{A}$ in the aforementioned example. Then, applying Proposition 27, we can ensure that $(U, \{O, H, W\})$ is a decision bireduct.

The following property studies the connection between the indiscernibility classes of the objects in a decision table and the bireducts with no attribute.

Proposition 28. *Let (X, B) be a decision bireduct with $X \subseteq U$ and $B \subseteq \mathcal{A}$. Then, $B = \emptyset$ if and only if $X = [x]_{I(\{d\})}$ for all $x \in U$.*

Proof. Suppose that $B = \emptyset$. Hence, each pair of objects $x, y \in X$ are B -indiscernible. Therefore, since (X, B) is a decision bireduct, each pair of objects $x, y \in X$ must be d -indiscernible. As a consequence, $X = [x]_{I(\{d\})}$ for all $x \in U$.

We suppose now $X = [x]_{I(\{d\})}$ for all $x \in U$. By hypothesis, every $x, y \in X$ are d -indiscernible. Therefore, since B does not need to discern any pair of objects, by the irreducibility condition of B , this set must be the empty set. $\square$

Notice that, unlike to information bireducts, a table reduced by classes is not needed.

The correspondence between the set of objects X of a decision bireduct $(X, \emptyset)$ and the indiscernibility classes $[x]_{I(\{d\})}$, for all $x \in U$, is illustrated below.

Example 29. Considering the subsets of objects $X_1 = \{1, 2, 6, 8, 14\}$ and $X_2 = \{3, 4, 5, 7, 9, 10, 11, 12, 13\}$ of the decision table given in Example 10, from the information displayed in Table 3, it is easy to see that $(X_1, \emptyset)$ and $(X_2, \emptyset)$ are all decision bireducts with the empty set as subset of attributes. In addition, taking into account the computations carried out in Example 23, we have that $[1]_{I(\{d\})} = \{1, 2, 6, 8, 14\}$ and $[3]_{I(\{d\})} = \{3, 4, 5, 7, 9, 10, 11, 12, 13\}$. As a consequence, comparing these subsets of objects with the decision bireducts $(X_1, \emptyset)$ and $(X_2, \emptyset)$, we obtain clearly a correspondence between the sets of objects X_i with $i \in \{1, 2\}$ and the indiscernibility classes $[x]_{I(\{d\})}$ for all $x \in U$.

3.3 | U -decision bireducts

This section continues our study of decision bireducts taking into account an alternative definition, which also has been studied in previous works.^{19,24,25} In addition, we will show some interesting results about this definition.

Definition 30. Given $X \subseteq U$ and $B \subseteq \mathcal{A}$, we say that the pair (X, B) is a U -decision bireduct if every pair $x \in X, y \in U$ are B -discernible when they are d -discernible, B is X -irreducible and X is B -inextensible with respect to this property.

Notice that, this definition is more restrictive than Definition 22 since each object $x \in X$ is compared with each object $y \in U$. As a consequence, in the decision subtables generated by U -decision bireducts there is no element triggering a contradiction in the original decision table. Now, we will clarify this definition in the following example.

Example 31. Consider the decision table given in Example 10 and the subset of attributes $B = \{O, T, H\} \subseteq \mathcal{A}$. As we mentioned in Example 23, the pair of objects 4,14 and 5,6 are the only objects d -discernible and B -indiscernible. As a consequence, none of these pairs of elements should be considered, because in the definition of U -decision bireduct every object in X is compared with any object in the universe U . Hence, none of them can appear in a U -decision bireduct. Therefore, according to Definition 30, the unique candidate to U -decision bireduct with B as subset of attributes is $(U \setminus \{4, 5, 6, 14\}, B)$.

Clearly, all objects are B -discernible if they are d -discernible and the subset of objects is B -inextensible by construction. Following the same reasoning that the one given in Example 23, we can ensure that B is $U \setminus \{4, 5, 6, 14\}$ -irreducible. Therefore, according to Definition 30, $(U \setminus \{4, 5, 6, 14\}, B)$ is a U -decision bireduct.

Continuing in the same line that the previous sections, we will show some technical properties related to U -decision bireducts.

Proposition 32. Given (X, B) a U -decision bireduct, where $X \subseteq U$ and $B \subseteq \mathcal{A}$, the following equivalence holds:

$$x \in X \text{ if and only if } [x]_{I(B)} \subseteq [x]_{I(\{d\})}$$

Proof. Supposing that $x \in X$, we prove that $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$. Since $x \in X$, for every $y \in U$, if x, y are d -discernible then they are B -discernible. Let $z \in [x]_{I(B)}$, then x, z are B -indiscernible and, as a consequence, they are d -indiscernible. Therefore, $z \in [x]_{I(\{d\})}$.

Now, suppose that $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$ and $y \in U$ such that x, y are d -discernible. Then $y \notin [x]_{I(\{d\})}$. Therefore, $y \notin [x]_{I(B)}$ and x, y are B -discernible. Taking into account that X is B -inextensible, we can ensure that $x \in X$. $\square$

It is convenient to emphasize the importance of this result. Specifically, we deduce that $x \in X$ if and only if there is no other object $y \in U$ such that x, y are B -indiscernible and d -discernible. Consequently, Proposition 32 proves that if there exists a U -decision bireduct with $B \subseteq \mathcal{A}$ as subset of attributes that is the unique U -decision bireduct with B as subset of attributes. In addition, from Proposition 32, we can deduce that if (X, B) is a U -decision bireduct, then $X = POS_B(\{d\})$ because $POS_B(\{d\}) = \{x \in U \mid [x]_{I(B)} \subseteq [x]_{I(\{d\})}\}$. In this case, we can compute U -decision bireducts in an easier way than decision bireducts, since Proposition 32 is more powerful than Proposition 24.

On the other hand, it is possible to obtain a stronger relationship between the positive region and U -decision bireducts, allowing the calculus of these bireducts through the positive region, as it is showed in Stawicki et al.²⁵

Proposition 33 (Stawicki et al.²⁵). Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table such that $X \subseteq U$ and $B \subseteq \mathcal{A}$. Then (X, B) is a U -decision bireduct if and only if $X = POS_B(\{d\})$ and there is no subset $B' \subset B$ such that $POS_{B'}(\{d\}) = POS_B(\{d\})$.

Analogously to decision bireducts, we introduce a result which relates the definitions of U -decision bireduct to d -reduct.

Proposition 34. Given $B \subseteq \mathcal{A}$, if the pair (U, B) is a U -decision bireduct then $POS_B(\{d\}) = U$. Moreover, B is a d -reduct of $\mathcal{A}$.

Proof. The proof is straightforwardly obtained from Proposition 33. $\square$

Just like it occurs in Proposition 25, the reciprocal of the last property is not true in general. Coming back to Example 26, from Table 6, we deduce that $B = \{h, g\}$ is a d -reduct of $\mathcal{A}$. Therefore, there is no subset $B' \subset B$ such that $POS_{B'}(\{d\}) = POS_B(\{d\})$. Taking into account that $POS_B(\{d\}) = \{3, 4\}$ and applying Proposition 33, we obtain that $(\{3, 4\}, B)$ is the U -decision bireduct with B as subset of attributes.

Equivalent results to Proposition 16 and Corollary 17 for decision tables are deduced from Proposition 33.

Proposition 35. *Given $B \subseteq \mathcal{A}$, if B is a d -reduct of $\mathcal{A}$ then there exists $X \subseteq U$ such that (X, B) is a U -decision bireduct.*

Proof. The proof is straightforwardly obtained from Proposition 33. $\square$

Corollary 36. *There exists $X \subseteq U$ such that $(X, \mathcal{A})$ is a U -decision bireduct if and only if $\mathcal{A}$ is a d -reduct of $\mathcal{A}$.*

Proof. The proof is straightforwardly obtained from Proposition 33. $\square$

The reciprocal of Proposition 34 is true when a decision table satisfies that $POS_{\mathcal{A}}(\{d\}) = U$, as it is showed in Stawicki et al.²⁵

Proposition 37 (Stawicki et al.²⁵). *Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table such that $POS_{\mathcal{A}}(\{d\}) = U$ and $B \subseteq \mathcal{A}$. Then, the pair (U, B) is a U -decision bireduct if and only if B is a d -reduct of $\mathcal{A}$.*

Finally, we introduce the result which studies U -decision bireducts whose subset of attributes is $B = \emptyset$.

Proposition 38. *Let $X \subseteq U$, $B \subseteq \mathcal{A}$ and (X, B) be a U -decision bireduct. Suppose that there exist $x, y \in U$ such that $y \notin [x]_{I(\{d\})}$. Then, $B = \emptyset$ if and only if $X = \emptyset$.*

Proof. First of all, we suppose $B = \emptyset$. By Proposition 32, $X \neq \emptyset$ if and only if there exists $x \in U$ such that $[x]_{I(\emptyset)} \subseteq [x]_{I(\{d\})}$. However, $[x]_{I(\emptyset)} = U$ and, by hypothesis $y \notin [x]_{I(\{d\})}$, then we deduce that $[x]_{I(\{d\})} \neq U$. Then $X = \emptyset$.

Now, we suppose $X = \emptyset$. Then, by Proposition 32, $[x]_{I(B)} \not\subseteq [x]_{I(\{d\})}$ for all $x \in U$. On the other hand, if $B' \subseteq B$ then $[x]_{I(B)} \subseteq [x]_{I(B')} \not\subseteq [x]_{I(\{d\})}$. Therefore, since $[x]_{I(\emptyset)} \not\subseteq [x]_{I(\{d\})}$, by the irreducibility condition of B , this set must be the empty set. $\square$

Hence, several properties of the three different classes of bireducts have been studied, which have shed more light on the notion of bireduct. Next section compares bireducts with another useful notion: value reduct.

4 | RELATION BETWEEN VALUE REDUCTS AND BIREDUCTS

This section is devoted to study the relationships among the different notions of value reducts and bireducts. The importance of this study lies on the comparison of two notions originally independent, so that the study is divided into three different parts depending on the different notions of bireduct in information tables and in decision tables.

4.1 | Relation between information value reducts and information bireducts

This first part focuses on information tables in order to connect information value reducts and information bireducts, that is, Definitions 5 and 11. In this section, an information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ will be fixed. Next result studies the belonging of objects to information bireducts by using the notion of value reduct.

Proposition 39. *Let $x \in U$ and $B \subseteq \mathcal{A}$ be a value reduct of $\mathcal{A}$ for x . Then, for each $y_0 \in [x]_{I(\mathcal{A})}$ there exists (X_{y_0}, B) information bireduct such that $y_0 \in X_{y_0}$. Indeed, all information bireducts (X, B) satisfy that $[x]_{I(\mathcal{A})} \cap X \neq \emptyset$.*

Proof. Given $B \subseteq \mathcal{A}$ being B a value reduct of $\mathcal{A}$ for $x \in U$, we will prove that there exists (X_0, B) information bireduct with $y_0 \in X_0$ being $y_0 \in [x]_{I(\mathcal{A})}$. Now, we consider the set $X'_0 = \{y_0\}$ and we add to X'_0 all objects satisfying that they are B -discernible among them, so that the new obtained set X_0 be B -inextensible. In this way, we have that X_0 is maximal. Now, we will prove that B is X_0 -irreducible by reductio ad absurdum. We suppose that there exists $B' \subset B$ such that, for all $y, z \in X_0$, we have that y, z are B' -discernible. Since B is a value reduct of $\mathcal{A}$ for x , we deduce that $[x]_{I(B')} \neq [x]_{I(B)}$. Specifically, $[x]_{I(B)} \subseteq [x]_{I(B')}$ but $[x]_{I(B')} \not\subseteq [x]_{I(B)}$. Therefore, there exists $w \in [x]_{I(B')}$ such that $w \notin [x]_{I(B)}$. In addition, as $y_0 \in [x]_{I(\mathcal{A})}$, we obtain that $[x]_{I(B')} = [y_0]_{I(B')}$ and $[x]_{I(B)} = [y_0]_{I(B)}$. We distinguish two cases:

- If $w \in X_0$, since y_0, w are B' -indiscernible, we obtain a contradiction.
- If $w \notin X_0$, since X_0 is B -inextensible by construction, there exists $z \in [w]_{I(B)}$ such that $z \in X_0$. Then the following chain holds:

$$z \in [w]_{I(B)} \subseteq [w]_{I(B')} = [x]_{I(B')} = [y_0]_{I(B')}$$

which lead us to conclude that y_0, z are B' -indiscernible, obtaining again a contradiction.

In any case, B is X_0 -irreducible and, as a consequence, (X_0, B) is an information bireduct with $y_0 \in X_0$ being $y_0 \in [x]_{I(A)}$.

Now, we will prove that for each information bireduct (X, B) there exists $y \in [x]_{I(A)}$ such that $y \in X$. We will prove it by reductio ad absurdum. Given (X', B) an information bireduct, suppose that for each $y \in [x]_{I(A)}$, we obtain that $y \notin X'$. Then, given $y \in [x]_{I(A)}$, there exists $z \in X'$ such that y, z are B -indiscernible, that is, $z \in [y]_{I(B)}$. Since B is a value reduct of A for x , we deduce the equality $[x]_{I(B)} = [x]_{I(A)}$. As $y \in [x]_{I(A)}$, we have $[y]_{I(A)} = [x]_{I(A)}$ and $[y]_{I(B)} = [x]_{I(B)}$ too. Hence, the following chain of equalities holds:

$$[y]_{I(A)} = [x]_{I(A)} = [x]_{I(B)} = [y]_{I(B)} = [z]_{I(B)}$$

which implies that $z \in [x]_{I(A)}$, and as $z \in X'$, this fact contradicts the supposition considered on (X', B) . $\square$

In the particular case of $[x]_{I(A)} = \{x\}$, the last result is simplified as follows.

Corollary 40. *Given $B \subseteq A$ and $x \in U$ such that $[x]_{I(A)} = \{x\}$ with B value reduct of A for x , then every information bireduct (X, B) satisfies that $x \in X$.*

Proof. The proof follows from Proposition 39 considering that $[x]_{I(A)} = \{x\}$. $\square$

This corollary allows to ensure that, given a value reduct for a certain object x , then x belongs to all information bireducts whose subset of attributes is the aforementioned value reduct. This fact is illustrated by means of the following example.

Example 41. Coming back to Examples 10 and 13, we can ensure that $[x]_{I(A)} = \{x\}$ for all $x \in U$, so that we are in the environment of Corollary 40, as well as $B = \{O, T, W\}$ is a value reduct of A for the object 5 and $(U \setminus \{10, 13\}, B)$, $(U \setminus \{3, 10\}, B)$, $(U \setminus \{4, 13\}, B)$ and $(U \setminus \{3, 4\}, B)$ are all the information bireducts, with B as subset of attributes. Clearly, we obtain that

$$U \setminus \{3, 4, 10, 13\} = U \setminus \{10, 13\} \cap U \setminus \{3, 10\} \cap U \setminus \{4, 13\} \cap U \setminus \{3, 4\}$$

and object 5 belongs to all information bireducts.

Next result shows the close relationship between the notions of value reducts and information bireducts. Specifically, we provide the required conditions to the subsets of objects of all information bireducts with B as subset of attributes, in order to guarantee the existence of a subset $B' \subseteq B$ being a value reduct of A , and vice versa.

Proposition 42. *Given $B \subseteq A$, $X_1, \dots, X_n \subseteq U$ and $x \in U$ such that $(X_1, B), \dots, (X_n, B)$ are all the information bireducts with B as subset of attributes. Then for each $i \in \{1, \dots, n\}$ there exists $y_i \in [x]_{I(A)}$ such that $y_i \in X_i$ if and only if there exists $B' \subseteq B$ such that B' is a value reduct of A for x .*

Proof. Supposing that for each X_i there exists $y_i \in [x]_{I(A)}$ such that $y_i \in X_i$, we will prove that there exists $B' \subseteq B$ such that B' is a value reduct of A for x . The proof will be done by reductio ad absurdum. We suppose that $B' \subseteq B$ is not a value reduct of A for x , for all $B' \subseteq A$. Then $[x]_{I(B')} \neq [x]_{I(A)}$ and since $[x]_{I(A)} \subseteq [x]_{I(B')}$, there exists $z \in [x]_{I(B')}$ satisfying $z \notin [x]_{I(A)}$. In particular, considering $B' = B$, we obtain x, z are B -indiscernible, that is, y_i, z are B -indiscernible for all $y_i \in [x]_{I(A)}$. Therefore, necessarily $z \notin X_i$ for all $i \in \{1, \dots, n\}$. Hence, $z \notin \bigcup_{i=1}^n X_i$ and by Proposition 12, we have that $z \notin U$, which leads us to a contradiction.

Now, assuming that there exists $B' \subseteq B$ such that B' is a value reduct of A for x , we will prove that there exists $y_i \in [x]_{I(A)}$ such that $y_i \in X_i$, for all $i \in \{1, \dots, n\}$. By reductio ad absurdum we will suppose that there exists $i \in \{1, \dots, n\}$, we have that $y \notin X_i$ for all $y \in [x]_{I(A)}$. Then, given $y_i \in [x]_{I(A)}$, there exists $z \in X_i$ such that y_i, z are B -indiscernible, so $z \in [y_i]_{I(B)}$. Since $y_i \in [x]_{I(A)}$, it is obtained that $[x]_{I(B)} = [y_i]_{I(B)}$ and $[x]_{I(B')} = [y_i]_{I(B')}$. Hence, we obtain

$$[x]_{I(A)} \subseteq [x]_{I(B)} = [y_i]_{I(B)} \subseteq [y_i]_{I(B')} = [x]_{I(B')} = [x]_{I(A)}$$

That is due to monotony of equivalence classes and because B' is a value reduct of A for x . As a consequence, $[y_i]_{I(B)} = [x]_{I(A)}$ and then $z \in [x]_{I(A)}$, obtaining a contradiction with the hypothesis considered about X_i . $\square$

Now, we include the last result in the particular case of $[x]_{I(A)} = \{x\}$. Specifically, we show that if we know all the information bireducts, associated with a fixed subset of attributes, we can delimit the study of value reducts for the objects belonging to the intersection of these information bireducts.

Corollary 43. Let $B \subseteq \mathcal{A}$, $X_1, \dots, X_n \subseteq U$, $x \in U$ such that $[x]_{I(\mathcal{A})} = \{x\}$ and $(X_1, B), \dots, (X_n, B)$ all the information bireducts with B as subset of attributes. Then $x \in \bigcap_{i=1}^n X_i$ if and only if there exists $B' \subseteq B$ such that B' is a value reduct of $\mathcal{A}$ for x .

Proof. The proof follows from Proposition 42 due to $[x]_{I(\mathcal{A})} = \{x\}$. $\square$

On the other hand, next example will show that the condition $B' \subseteq B$ is indispensable in Corollary 43, that is, B could not be a value reduct of $\mathcal{A}$ for x , in general. As a consequence, the mentioned condition will be also essential in Proposition 42.

Example 44. Consider the information table given in Example 10 and the subset of attributes $B = \{O, T, W\} \subseteq \mathcal{A}$. In Example 13, we calculated all information bireducts with B as subset of attributes, obtaining that they are $(U \setminus \{10, 13\}, B)$, $(U \setminus \{3, 10\}, B)$, $(U \setminus \{4, 13\}, B)$ and $(U \setminus \{3, 4\}, B)$. Hence, it is easy to see that object 9 belongs to all of them.

On the other hand, from Table 2, we conclude that $\{O, T\}$ is a value reduct of $\mathcal{A}$ for the object 9. Therefore, $\{O, T, W\}$ is not a value reduct of $\mathcal{A}$ for the object 9 and so, we must include the condition $B' \subseteq B$ in Proposition 42 and in Corollary 43.

4.2 | Relation between decision value reducts and decision bireducts

This second part focuses on decision tables in order to relate d -value reducts and decision bireducts, that is, Definitions 9 and 22. From now on, a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ will be fixed. First of all, we introduce a result which will be useful later.

Lemma 45. Let $B \subseteq \mathcal{A}$ be a d -value reduct of $\mathcal{A}$ for $x \in U$ such that $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$ and (X, B) be a decision bireduct. If $x \in X$ then $[x]_{I(B)} \subseteq X$.

Proof. Given $y \in [x]_{I(B)}$, we will prove that $y \in X$ by reductio ad absurdum. Suppose that $y \notin X$. Then, there exists $z \in X$ such that y, z are d -discernible and B -indiscernible. Since B is a d -value reduct of $\mathcal{A}$ for x and $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$, we obtain that $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$. Hence, $y \in [x]_{I(\{d\})}$. As a consequence, $x, z \in X$, they are d -discernible and B -indiscernible, which leads us to a contradiction. Therefore, $[x]_{I(B)} \subseteq X$. $\square$

Following the same scheme of the previous section, the following result studies the belonging of objects to decision bireducts by using the notion of d -value reduct.

Proposition 46. Let $B \subseteq \mathcal{A}$ be a d -value reduct of $\mathcal{A}$ for $x \in U$ such that $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Then, there exists a decision bireduct (X_0, B) such that $x \in X_0 \subseteq U$. Indeed, all decision bireducts (X, B) satisfy that $[x]_{I(B)} \subseteq X$.

Proof. First of all, we will prove the existence of a decision bireduct (X_0, B) with $x \in X_0$. The strategy followed in this proof consists of defining a set of objects $X_0 \subseteq U$ such that (X_0, B) is a decision bireduct with $x \in X_0$. Consider $X'_0 = \{x\}$. Since B is a d -value reduct of $\mathcal{A}$ for x with $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$, we have for each $a \in B$ that there exists $z_a \in [x]_{I(B \setminus \{a\})}$ such that $z_a \notin [x]_{I(\{d\})}$. Now, consider $X''_0 = \{x, z_{a_1}, \dots, z_{a_{|B|}}\}$. As B is a d -value reduct of $\mathcal{A}$ for x and $z_{a_i} \notin [x]_{I(\{d\})}$ for all $i \in \{1, \dots, |B|\}$, we deduce that $z_{a_i} \notin [x]_{I(B)}$. Then x, z_{a_i} are d -discernible, B -discernible and $B \setminus \{a_i\}$ -indiscernible, for all $i \in \{1, \dots, |B|\}$.

We suppose now that there exist $j, k \in \{1, \dots, |B|\}$ such that z_{a_j}, z_{a_k} are d -discernible and B -indiscernible. Hence,

$$z_{a_k} \in [z_{a_j}]_{I(B)} \subseteq [z_{a_j}]_{I(B \setminus \{a_j\})} = [x]_{I(B \setminus \{a_j\})}$$

Therefore, x, z_{a_k} are $B \setminus \{a_j\}$ -indiscernible. By construction, x, z_{a_k} are also $B \setminus \{a_k\}$ -indiscernible. As a consequence, we can consider $X'''_0 \setminus \{z_{a_j}\}$ so that, for each $a \in B$, there exists $z_a \in X'''_0 \setminus \{z_{a_j}\}$ such that x, z_a are d -discernible, B -discernible and $B \setminus \{a\}$ -indiscernible, being BX'''_0 -irreducible. Let X''''_0 be the set obtained from X'''_0 , applying the previous procedure to every pair of d -discernible and B -indiscernible objects z_{a_j}, z_{a_k} . Therefore, the new obtained set X''''_0 satisfies that every pair of objects d -discernible are B -discernible and B is X''''_0 -irreducible with respect to this property. Now, we can increase X''''_0 keeping this property until the set is B -inextensible, being this last set X_0 . As a consequence, (X_0, B) is a decision bireduct with $x \in X_0$.

Now, we will prove that for each decision bireduct (X, B) it obtains that $[x]_{I(B)} \subseteq X$. First of all, we will prove that $x \in X$ by reductio ad absurdum. Let (X', B) be a decision bireduct such that $x \notin X'$. As X' is B -inextensible, there exists $y \in X'$ such that x, y are d -discernible and B -indiscernible, and therefore $y \in [x]_{I(B)}$. Since B is a d -value reduct of $\mathcal{A}$ for x , we obtain that $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$. Therefore $y \in [x]_{I(\{d\})}$. As a consequence, x, y are d -indiscernible, which contradicts that x, y be d -discernible. Therefore, $x \in X$ and applying Lemma 45, $[x]_{I(B)} \subseteq X$. $\square$

Notice that the assumption $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$ is not restrictive since otherwise the objects in the class $[x]_{I(\mathcal{A})}$ do not classify correctly (they have different decisions) and so, reducing attributes is meaningless. Indeed, in this case it is necessary to increase the number of attributes to obtain a better classification.

Hence, this proposition guarantees that given a d -value reduct for a certain object x , then x belongs to all decision bireducts whose subset of attributes is the aforementioned d -value reduct. This fact is exemplified below.

Example 47. Returning to Example 10, we have that $B = \{O, H\}$ is a d -value reduct of $\{O, T, H\}$ and a d -value reduct of $\mathcal{A}$ for the object 1. In addition, following a similar procedure to the one given in Example 23, we have that $(U \setminus \{4, 6\}, B)$, $(U \setminus \{6, 14\}, B)$, $(U \setminus \{4, 5, 10\}, B)$ and $(U \setminus \{5, 10, 14\}, B)$ are all decision bireducts with $B = \{O, H\}$ as subset of attributes. Clearly, object 1 belongs to all of them, as Proposition 46 states.

To finish this section, it is important to emphasize that a similar result to Proposition 42 is not verified to decision bireducts. From the computations of Example 23, we have that $(U \setminus \{5, 14\}, \{O, T, H\})$ and $(U \setminus \{4, 5\}, \{O, T, H\})$ are all decision bireducts with $\{O, T, H\}$ as subset of attributes. Hence, it is easy to see that the object 6 belongs to all of them.

On the other hand, from Table 4, we have that $\{O, W\}$ is a d -value reduct of $\mathcal{A}$ for the object 6 and the pair of objects 5, 6 are d -discernible and $\{O, T, H\}$ -indiscernible. Therefore, $\{O, T, H\}$ is not a d -value reduct of $\mathcal{A}$ for the object 6. As a consequence, the attribute W must belong to all d -value reducts of $\mathcal{A}$ for the object 6. Following an analogous reasoning with the pair of objects 6, 7 and the subset of attributes $\{T, H, W\}$ we can conclude that the attribute O must belong to all d -value reducts of $\mathcal{A}$ for the object 6. In short, since both attributes, O and W , must belong to all d -value reducts of $\mathcal{A}$ for the object 6 and $\{O, W\}$ is a d -value reduct, we can conclude that $\{O, W\}$ is the unique d -value reduct of $\mathcal{A}$ for the object 6. Since the object 6 belongs to all decision bireducts with $\{O, T, H\}$ as subset of attributes, $\{O, W\}$ is the unique d -value reduct of $\mathcal{A}$ for the object 6 and $\{O, W\} \not\subseteq \{O, T, H\}$, we conclude that if an object x belongs to all decision bireducts with a fixed subset of attributes B , then it may not exist $B' \subseteq B$ with B' a d -value reduct of $\mathcal{A}$ for the object x . As a consequence, an analogous result to Proposition 42 for decision tables is not verified to decision bireducts.

4.3 | Relation between decision value reducts and U -decision bireducts

This section finishes our study relating decision value reducts and U -decision bireducts. First of all, we introduce a result which will be useful later.

Lemma 48. Let (X, B) be a U -decision bireduct where $X \subseteq U$ and $B \subseteq \mathcal{A}$. Given $x \in U$, if $x \in X$ then $[x]_{I(B)} \subseteq X$.

Proof. Given $y \in [x]_{I(B)}$, we will prove that $y \in X$. Since $x \in X$, by using Proposition 32, we obtain that $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$. As a consequence, we obtain the following chain of equalities:

$$[y]_{I(B)} = [x]_{I(B)} \subseteq [x]_{I(\{d\})} = [y]_{I(\{d\})}$$

As a consequence, by using Proposition 32, we obtain that $y \in X$. Therefore, $[x]_{I(B)} \subseteq X$. $\square$

Next result studies the belonging of objects to U -decision bireducts by using the notion of d -value reduct, obtaining a result similar to Proposition 46.

Proposition 49. Let $B \subseteq \mathcal{A}$ be a d -value reduct of $\mathcal{A}$ for $x \in U$ such that $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Then, there exists $X \subseteq U$, such that the pair (X, B) is a U -decision bireduct satisfying that $[x]_{I(B)} \subseteq X$.

Proof. We define the set $X = \{y \in U \mid [y]_{I(B)} \subseteq [y]_{I(\{d\})}\}$. Notice that, $y \in X$ if and only if for each object $z \in U$ such that y, z are d -discernible, they also are B -discernible. Therefore, X is B -inextensible. Furthermore, $x \in X$ because B is a d -value reduct of $\mathcal{A}$ for x and $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Finally, since B is a d -value reduct of $\mathcal{A}$ for x , for each $B' \subset B$ we obtain that $[x]_{I(B')} \not\subseteq [x]_{I(\{d\})}$. As a consequence, for each $B' \subset B$ there exists $y_{B'} \in U$ such that $x, y_{B'}$ are d -discernible and B' -indiscernible. Hence, B is X -irreducible. Since $x \in X$, (X, B) is a U -decision bireduct and applying Lemma 48, we obtain that $[x]_{I(B)} \subseteq X$. $\square$

The following result also relates the notions of U -decision bireduct to d -value reduct, in particular, it studies the existence of d -value reducts for those objects which belong to U -decision bireducts. Notice that, this result is equivalent to Proposition 42 for decision tables, by using the notion of U -decision bireducts.

Proposition 50. *Let (X, B) be a U -decision bireduct with $X \subseteq U$, $B \subseteq \mathcal{A}$ and $x \in U$ such that $[x]_{I(\mathcal{A})} \subseteq [x]_{I(\{d\})}$. Then, $[x]_{I(B)} \subseteq X$ if and only if there exists $B' \subseteq B$ such that B' is a d -value reduct of $\mathcal{A}$ for x .*

Proof. Supposing that $[x]_{I(B)} \subseteq X$, we will prove that there exists $B' \subseteq B$ such that B' is a d -value reduct of $\mathcal{A}$ for x . We will prove it by reductio ad absurdum. Consider $x \in X$ and suppose that any subset of attributes $B' \subseteq B$ is not a d -value reduct of $\mathcal{A}$ for x . Then $[x]_{I(B')} \not\subseteq [x]_{I(\{d\})}$, for all $B' \subseteq B$. However, by Proposition 32, $[x]_{I(B)} \subseteq [x]_{I(\{d\})}$ which leads us to a contradiction when $B' = B$ is considered in the expression above.

Now, supposing that there exists $B' \subseteq B$ such that B' is a d -value reduct of $\mathcal{A}$ for x , we will prove that $[x]_{I(B)} \subseteq X$. Given $y \in [x]_{I(B)}$ we obtain

$$[y]_{I(B)} = [x]_{I(B)} \subseteq [x]_{I(B')} \subseteq [x]_{I(\{d\})} = [y]_{I(\{d\})}$$

where $[x]_{I(B)} \subseteq [x]_{I(B')}$ because $B' \subseteq B$ and $[x]_{I(B')} \subseteq [x]_{I(\{d\})}$, since B' is a d -value reduct of $\mathcal{A}$ for x . Therefore, by using Proposition 32, we obtain that $y \in X$. As a consequence, $[x]_{I(B)} \subseteq X$. $\square$

It is important to observe that B could not be a d -value reduct of $\mathcal{A}$ for x , that is, the condition $B' \subseteq B$ is indispensable in Proposition 50 as we will show next.

Example 51. Consider the decision table given in Example 10 and the subset of attributes $B = \{O, T, H\} \subseteq \mathcal{A}$. In Example 31, we obtained that $(U \setminus \{4, 5, 6, 14\}, \{O, T, H\})$ is the U -decision bireduct with $\{O, T, H\}$ as subset of attributes. From now on, we will focus on the object 1, which belongs to that U -decision bireduct. From Table 3, we can conclude that the pair of objects 1,3 are d -discernible and $\{T, H, W\}$ -indiscernible. Therefore, attribute O must belong to all d -value reducts of $\mathcal{A}$ for the object 1. Now, we compute all d -value reducts of $\mathcal{A}$ for the object 1.

On the one hand, the pair of objects 1,9 are d -discernible and $\{O\}$ -indiscernible. Therefore, $\{O\}$ is not a d -value reduct of $\mathcal{A}$ for the object 1. From Table 4, we have that $B'_1 = \{O, H\}$ is a d -value reduct of $\mathcal{A}$ for the object 1.

On the other hand, from Table 3, it is easy to see that

$$[1]_{I(\{O,T\})} = \{1, 2\} \subseteq [1]_{I(\{d\})} = \{1, 2, 6, 8, 14\}$$

and, as a result, $B'_2 = \{O, T\}$ is a d -value reduct of $\mathcal{A}$ for the object 1.

Notice that, $\{O, W\}$ is not a d -value reduct of $\mathcal{A}$ for the object 1, since we also obtain that the pair of objects 1,9 are d -discernible and $\{O, W\}$ -indiscernible from Table 3.

Therefore, B'_1 and B'_2 are all the d -value reducts of $\mathcal{A}$ for the object 1, being $B'_1, B'_2 \neq B$. Hence, B is not a d -value reduct of $\mathcal{A}$ for the object 1 and so, we must include the condition $B' \subseteq B$ in Proposition 50.

5 | CONCLUSIONS AND FURTHER WORK

This paper has delved into the study of bireducts, analyzing new properties about bireducts whose subsets of objects or attributes are extreme cases, that is, they are the whole set of objects or attributes contained in the information/decision table or the empty set. Furthermore, we have proven that the attribute set of an information bireduct provides a covering of the universe of objects into maximal consistent subsystems.

The second part of the paper is focused on the relationship between bireducts and value reducts in both information and decision tables, and contains interesting results of these two originally independent notions which really prove both are very well related. Furthermore, in order to clarify the content of the paper, we have accompanied all the notions and results with examples. This relationship will also be useful for obtaining more optimal decision rules extraction mechanisms, which will be analyzed in detail in the future.

Moreover, we will study the generalization of the results introduced in this paper to the fuzzy environment. Furthermore, we will analyze more relationships between bireducts and value reducts, and their relation and application to other frameworks, such as logic, formal concept analysis and non-linear relation equations.^{16,27–29} In particular, we are also interested in the study of the relationship between decision rules and attribute implication, studied in formal concept analysis.

ACKNOWLEDGEMENTS

Partially supported by the 2014-2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

CONFLICT OF INTERESTS

This work does not have any conflict of interests.

ORCID

Fernando Chacón-Gómez  <https://orcid.org/0000-0001-9710-7813>

M. Eugenia Cornejo  <https://orcid.org/0000-0002-8230-0044>

Jesús Medina  <https://orcid.org/0000-0002-3931-5873>

Eloísa Ramírez-Poussa  <https://orcid.org/0000-0002-2391-5217>

REFERENCES

1. Pawlak Z. Rough sets. *Int J Comput Inform Sci.* 1982;11:341-356.
2. Brentan BM, Carpitella S, Izquierdo J, Luvizotto Jr E, Meirelles G. District metered area design through multicriteria and multiobjective optimization. *Math Methods Appl Sci.* 45(6):3254-3271.
3. Cornejo ME, Díaz-Moreno JC, Medina J. Multi-adjoint relation equations: A decision support system for fuzzy logic. *Int J Intell Syst.* 2017;32(8):778-800.
4. Dursun M, Gumus G. Intuitionistic fuzzy cognitive map approach for the evaluation of supply chain configuration criteria. *Math Methods Appl Sci.* 2020;43(13):7788-7801.
5. El Barbary OG, Salama AS, Atlam ES. Granular information retrieval using neighborhood systems. *Math Methods Appl Sci.* 2018;41(15):5737-5753.
6. Rodríguez-Jiménez JM, Cordero P, Enciso M, Mora A. Data mining algorithms to compute mixed concepts with negative attributes: an application to breast cancer data analysis. *Math Methods Appl Sci.* 2016;39(16):4829-4845.
7. Antoni L, Cornejo ME, Medina J, Ramírez E. Attribute classification and reduct computation in multi-adjoint concept lattices. *IEEE Trans Fuzzy Syst.* 2020;29(5):1121-1132.
8. Aragón RG, Medina J, Ramírez-Poussa E. Identifying non-sublattice equivalence classes induced by an attribute reduction in FCA. *Mathematics.* 2021;9(5):565.
9. Aragón RG, Medina J, Ramírez-poussa E. Impact of local congruences in variable selection from datasets. *J Comput Appl Math.* 2022;404. doi:10.1016/j.cam.2021.113416
10. Benítez-caballero M.J, Medina J, Ramírez-Poussa E, Ślęczak D. Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis, 391; 2020.
11. Chen J, Mi J, Xie B, Lin Y. A fast attribute reduction method for large formal decision contexts. *Int J Approx Reason.* 2019;106:1-17.
12. Pawlak Z. *Rough Sets: Theoretical Aspects of Reasoning About Data.* Norwell, MA USA: Kluwer Academic Publishers; 1992.
13. Yao Y, Fu R. *The Concept of Reducts in Pawlak Three-Step Rough Set Analysis.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2013.
14. Cornelis C, Jensen R, Hurtado G, Ślęczak D. Attribute selection with fuzzy decision reducts. *Inform Sci.* 2010;180:209-224.
15. Janusz A, Ślęczak D. Rough set methods for attribute clustering and selection. *Appl Artif Intell.* 2014;28(3):220-242.
16. Medina J. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Comput Math Appl.* 2012;64(6):1992-2002.
17. Yao Y, Zhao Y. Attribute reduction in decision-theoretic rough set models. *Inform Sci.* 2008;178(17):3356-3373.
18. Zhao Y, Yao Y, Luo F. Data analysis based on discernibility and indiscernibility. *Inform Sci.* 2007;177(22):4959-4976.
19. Benítez-Caballero MJ, Medina J, Ramírez-Poussa E, Ślęczak D. Bireducts with tolerance relations. *Inform Sci.* 2018;435:26-39.
20. Benítez-Caballero MJ, Medina J, Ramírez-Poussa E, Ślęczak D. A computational procedure for variable selection preserving different initial conditions. *Int J Comput Math.* 2020;97(1-2):387-404.
21. Benítez-Caballero MJ, Medina-Moreno J, Ramírez-poussa E. *Bireducts in Formal Concept Analysis.* Cham: Springer International Publishing; 2020.
22. Janusz A, Ślęczak D, Nguyen HS. Unsupervised similarity learning from textual data. *Fundamenta Informaticae.* 2012;119:319-336.
23. Mac Parthalain N, Jensen R. Simultaneous feature and instance selection using fuzzy-rough bireducts. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE, 2013); 2013:1-8.
24. Stawicki S, Ślęczak D. Recent advances in decision bireducts: Complexity, heuristics and streams. In: Lingras P, Wolski M, Cornelis C, Mitra S, Wasilewski P, eds. *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, vol. 8171: Springer Berlin Heidelberg; 2013:200-212.

25. Stawicki S, Śleczak D, Janusz A, Widz S. Decision bireducts and decision reducts - a comparison. *Int J Approx Reason*. 2017;84:75-109.
26. Benítez-Caballero MJ, Medina J, Ramírez-Poussa E. Towards a classification of rough set bireducts. In: Lesot M-J, Vieira S, Reformat MZ, et al., eds. *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Cham: Springer International Publishing; 2020:759-770.
27. Cornejo ME, Lobo D, Medina J. Bipolar fuzzy relation equations systems based on the product t-norm. *Math Methods Appl Sci*. 2019;42(17):5779-5793.
28. Madrid N, Ojeda-Aciego M. A measure of consistency for fuzzy logic theories. *Math Methods Appl Sci*. 2021;1-14. doi:10.1002/mma.7470
29. Quirós P, Alonso P, Díaz I, Jurío A, Montes S. A hybrid construction method based on weight functions to obtain interval-valued fuzzy relations. *Math Methods Appl Sci*. 2016;39(16):4723-4735.

How to cite this article: Chacón-Gómez F, Cornejo ME, Medina J, Ramírez-Poussa E. Value reducts and bireducts: a comparative study. *Math Meth Appl Sci*. 2022;1-20. doi:10.1002/mma.8597

Capítulo 5

Algoritmos de decisión en conjuntos rugosos para modelar con incertidumbre

Título:	Rough set decision algorithms for modeling with uncertainty
Autores:	Fernando Chacón-Gómez, M. Eugenia Cornejo, Jesús Medina, Eloísa Ramírez-Poussa
Revista:	Journal of Computational and Applied Mathematics
Factor de impacto JCR:	2,4. Posición 45 de 267 (Q1)
Año:	2023
Categoría:	Mathematics, Applied
Publicación:	Febrero 2024
DOI:	10.1016/j.cam.2023.115413



Rough set decision algorithms for modeling with uncertainty[☆]

Fernando Chacón-Gómez, M. Eugenia Cornejo, Jesús Medina^{*},
Eloísa Ramírez-Poussa

Department of Mathematics, University of Cádiz, Spain



ARTICLE INFO

Article history:

Received 16 September 2022

Received in revised form 3 May 2023

Keywords:

Rough set theory

Fuzzy rough set theory

Decision rules

Decision algorithms

ABSTRACT

The use of decision rules allows to extract information and to infer conclusions from relational databases in a reliable way, thanks to some indicators like support and certainty. Moreover, decision algorithms collect a group of decision rules that satisfies desirable properties to describe the relational system. However, when a decision table is considered within a fuzzy environment, it is necessary to extend all notions related to decision algorithms to this framework. This paper presents a generalization of these notions, highlighting the new definitions of indicators of relevance to describe decision rules and decision algorithms.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Pawlak [1,2] introduced a formal tool to deal with imprecise or incomplete information contained in databases, which was called Rough Set Theory (RST). This theory considers information systems, which are relational databases composed of a set of objects and a set of attributes which characterizes the objects, and it has attracted the attention of other (related) theories. For example, attribute reduction in the extensions of RST from the philosophy of Formal Concept Analysis (FCA) were studied in [3–6]. The paper [7] combined the K-nearest neighbors algorithm with RST, the algebraic notion of congruence was adapted in order to introduce an optimal attribute reduction based on RST in the FCA framework in [8], fuzzy soft sets were used to propose a combined forecasting approach for complementing a methodology based on rough sets [9], and fuzzy rough sets and probability statistics were studied and combined in [10]. A particular type of information system widely studied within RST is the one provided by decision tables. Decision rules [11–16] can be used to describe decision tables, allowing the extraction of information and the inference of significant conclusions. Decision rules are usually accompanied by some relevance indicators which describe them. In addition, decision algorithm is another important notion related to decision rules. This notion was introduced with the purpose of collecting some desirable properties that a set of decision rules should satisfy in order to describe the decision table in a suitable way.

On the other hand, the management of real databases can be a really difficult task. Fuzzy Rough Set Theory (FRST) [17–19] arises with the goal of increasing the flexibility of RST, addressing a wider range of real applications. FRST is a generalization of RST in the fuzzy environment, essential to deal with the information contained in databases, especially when they are composed by continuous quantities.

[☆] Partially supported by the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI), Spain in project PID2019-108991GB-I00, with the Ecological and Digital Transition Projects 2021 of the Ministry of Science and Innovation, Spain in project TED2021-129748B-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

^{*} Corresponding author.

E-mail addresses: fernando.chacon@uca.es (F. Chacón-Gómez), mariaeugenia.cornejo@uca.es (M.E. Cornejo), jesus.medina@uca.es (J. Medina), eloisa.ramirez@uca.es (E. Ramírez-Poussa).

The aim of this paper is to provide a generalization of the notion of decision algorithm from the classical environment to the fuzzy framework given by the multi-adjoint paradigm. With this goal, all the necessary notions related to decision rules and their corresponding relevance indicators are also extended to the fuzzy environment. We will prove that the new notions introduced in this work are indeed generalizations of the classical ones. This fact shows the possibility of applying this approach to the mechanisms focused on obtaining decision rules, which are mainly based on a crisp point of view. Notice that, for example, the last contributions in the area, such as [20], which considers an equivalent algorithm to the one provided by Pawlak [14] in the crisp case. The authors present a mechanism on a real decision table in which they need to consider a discretization, however, it is more natural to study this table from a fuzzy point of view, as we will do at the end of the paper. Hence, the contribution of this paper will also allow to extend the theory given in [20] to the fuzzy case. Granular computing is also applied to rule acquisition [21] in which the fuzzy point of view will also allow to be more precise and flexible, and give more information about the decision table. Moreover, in [22], the authors consider a variant of the notion of algorithm [14] in crisp incomplete decision systems, in which the approach introduced in this paper also extends the results given in the aforementioned paper to the fuzzy case, being also possible its adaptation to incomplete decision systems. Thus, it is possible to apply the framework introduced in this paper to classification problems with a greater flexibility than the current mechanisms. We will also include some examples to illustrate the content of the paper.

The paper is organized as follows: Section 2 recalls some preliminary classical notions of RST. In Section 3, a study about decision rules in FRST is introduced, including the generalizations of the relevance indicators to the fuzzy framework. Section 4 introduces a new notion of decision algorithm and shows the existing connection between the classical and the fuzzy version. Section 6 finishes with some conclusions and prospects for future work.

2. Basic notions in rough set theory

This section presents some important definitions of RST [13,14]. First of all, it is convenient to recall that databases can be represented as decision tables in this framework.

Definition 1. Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A *decision table* is a tuple $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with the attribute a over U , and $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a}: U \rightarrow V_a\}$. In this case, the attributes of $\mathcal{A}$ are called *condition attributes* and d is called *decision attribute*.

This paper will focus on decision tables with one decision attribute, as it is shown in this definition. However, a similar study could be carried out in decision tables with more than one decision attribute.

Now, an equivalence relation is defined on the set of objects of a decision table, in order to compare them by using a given subset of attributes.

Definition 2. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table. The *indiscernibility mapping* $I: \mathcal{P}(\mathcal{A}_d) \rightarrow \mathcal{P}(U \times U)$ is defined, for each $B \subseteq \mathcal{A}_d$, as the equivalence relation

$$I(B) = \{(x, y) \in U \times U \mid \bar{a}(x) = \bar{a}(y), \text{ for all } a \in B\}$$

which is called *B-indiscernibility relation*. Each class of $I(B)$ can be written as $[x]_{I(B)} = \{y \in U \mid (x, y) \in I(B)\}$. The partition determined by $I(B)$ on the set of objects U is denoted as $U/I(B) = \{[x]_{I(B)} \mid x \in U\}$.

The positive region is introduced now. This notion is deeply studied in order to analyze and extract information from decision tables.

Definition 3. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. The *positive region* of the partition $U/I(\{d\})$ with respect to B is defined as

$$\text{POS}_B(\{d\}) = \bigcup_{X \in U/I(\{d\})} B_*(X)$$

where $B_*(X) = \{x \in U \mid [x]_{I(B)} \subseteq X\}$.

2.1. Decision rules

The previous notions are useful for the management of the information contained in decision tables. Nevertheless, decision rules [14] arise in order to analyze these tables by using logic equivalences. This fact provides easier interpretation of decision tables, describing decisions in terms of conditions that must be satisfied. In order to present the notion of decision rule, we must introduce a formal language to describe approximations in logical terms.

Definition 4. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}_d$. The set of formulas associated with B , denoted as $For(B)$, is built from attribute–value pairs (a, v) , where $a \in B$ and $v \in V_a$, by means of the conjunction and disjunction logical connectives, $\wedge$ and $\vee$, respectively.

For each $\Phi \in For(B)$, with $\Phi = (a, v)$, the set of objects $X \subseteq U$ that satisfies Φ in S is defined as:

$$\|\Phi\|_S = \|(a, v)\|_S = \{x \in U \mid \tilde{a}(x) = v\}$$

Inductively, given $\Phi, \Psi \in For(B)$, the set of objects that satisfies $\Phi \wedge \Psi$ in S is defined as $\|\Phi \wedge \Psi\|_S = \|\Phi\|_S \cap \|\Psi\|_S$ and the set of objects that satisfies $\Phi \vee \Psi$ in S is defined as $\|\Phi \vee \Psi\|_S = \|\Phi\|_S \cup \|\Psi\|_S$.

Now, we introduce the notion of decision rule. As we mentioned above, it is an important pillar of the synthesis of information from relational databases.

Definition 5. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. A *decision rule* in S is an expression $\Phi \rightarrow \Psi$, with $\Phi \in For(B)$, $\Psi \in For(\{d\})$ where Φ and Ψ are the antecedent and the consequent of the decision rule, respectively. In addition, we will say that an object $x \in U$ satisfies the decision rule $\Phi \rightarrow \Psi$ if $x \in \|\Phi \wedge \Psi\|_S$.

As we commented above, decision rules provide a logic representation of decision tables. However, it is necessary to interpret these rules for the draw of conclusions from the decision table. For this purpose, we recall the indicators which describe decision rules from different points of view. Specifically, the *support* considers the number of objects satisfying the decision rule, the *strength* represents the proportion of objects satisfying the decision rule, that is, the representativeness of that rule in the table, the *certainty* provides us with the proportion of objects which satisfy the consequent satisfying the antecedent of the given decision rule and the *coverage* determines the proportion of objects which satisfy the antecedent satisfying the consequent of the given decision rule.

Definition 6. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $B \subseteq \mathcal{A}$ and $\Phi \rightarrow \Psi$ be a decision rule in S , with $\Phi \in For(B)$ and $\Psi \in For(\{d\})$. We call

- *support* of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$supp_S(\Phi, \Psi) = \|\Phi \wedge \Psi\|_S$$

- *strength* of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$\sigma_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{|U|}$$

- *certainty* of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$cer_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{\|\Phi\|_S}$$

when $\|\Phi\|_S \neq 0$.

- *coverage* of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$cov_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{\|\Psi\|_S}$$

when $\|\Psi\|_S \neq 0$.

From the notion of certainty, we will say that $\Phi \rightarrow \Psi$ is a *true decision rule* if $cer_S(\Phi, \Psi) = 1$. If $cer_S(\Phi, \Psi) = 0$, we say that the decision rule is *false*. Otherwise, it will be called a *not entirely true decision rule*.

These notions are of great importance in the study of decision algorithms and their efficiency, as it is shown in the next section.

2.2. Decision algorithm

Decision algorithms are a collection of decision rules of a decision table verifying certain conditions, focused on providing a global representation of the decision table. The definition of a decision algorithm is given below [14].

Definition 7. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$. We say that:

1. $Dec(S)$ is a set of *pairwise mutually exclusive (independent) decision rules*, if each pair of decision rules $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisfies that $\Phi = \Phi'$ or $\|\Phi \wedge \Phi'\|_S = \emptyset$, and $\Psi = \Psi'$ or $\|\Psi \wedge \Psi'\|_S = \emptyset$.
2. $Dec(S)$ *covers* U , if $\|\bigvee_{i=1}^m \Phi_i\|_S = \|\bigvee_{i=1}^m \Psi_i\|_S = U$.
3. A decision rule $\Phi \rightarrow \Psi \in Dec(S)$ is *admissible* in S if $supp_S(\Phi, \Psi) \neq 0$.

4. $Dec(S)$ preserves the consistency of S if

$$POS_{\mathcal{A}}(\{d\}) = \bigvee_{\Phi \rightarrow \Psi \in Dec^+(S)} \|\Phi\|_S$$

where $Dec^+(S)$ is the set of true decision rules of $Dec(S)$.

A set of decision rules $Dec(S)$ satisfying properties (1)–(4) is called *decision algorithm* in S and it is denoted as $DA(S)$.

Notice that, the notion of consistency is not defined from the positive region in other works, such as [22]. However, the approach proposed in [22] considers the notions of admissible and redundancy, being this last one closely related to Property (1) in Definition 7. Hence, our contribution also extends the results given in [22], being possible its adaptation to incomplete decision systems.

On the other hand, a pair of conclusions can be extracted from the definition of decision algorithm. By Properties (1) and (2), if $DA(S)$ is a decision algorithm, the antecedents and consequents of the decision rules of $DA(S)$ define a partition of U . Furthermore, by Property (4), the set of true decision rules define a partition of the positive region with respect to $\mathcal{A}$. On the other hand, it is convenient to emphasize that there may exist objects which do not satisfy any decision rule of the decision algorithm $DA(S)$, as it is shown in the following example.

Example 8. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table where the set of objects is $U = \{x_1, x_2, x_3\}$ and the set of attributes is $\mathcal{A} = \{a\}$. This decision table is represented in Table 1.

Table 1
Table associated with the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 8.

	a	d
x_1	0	0
x_2	0	1
x_3	1	1

Consider the set of decision rules $Dec(S) = \{\Phi_1 \rightarrow \Psi_1, \Phi_2 \rightarrow \Psi_2\}$, denoted as r_1 and r_2 respectively, given as:

$$r_1 : (a, 0) \rightarrow (d, 0)$$

$$r_2 : (a, 1) \rightarrow (d, 1)$$

Hence, it is immediate to check that $\Phi_1 \neq \Phi_2$, $\Psi_1 \neq \Psi_2$ and

$$\|\Phi_1 \wedge \Phi_2\|_S = \|\Phi_1\|_S \cap \|\Phi_2\|_S = \{x_1, x_2\} \cap \{x_3\} = \emptyset$$

$$\|\Psi_1 \wedge \Psi_2\|_S = \|\Psi_1\|_S \cap \|\Psi_2\|_S = \{x_1\} \cap \{x_2, x_3\} = \emptyset$$

In addition, $\|\bigvee_{i=1}^2 \Phi_i\|_S = \|\bigvee_{i=1}^2 \Psi_i\|_S = U$ and $supp_S(\Phi_1, \Psi_1) = supp_S(\Phi_2, \Psi_2) = 1 \neq 0$. Finally, $cer_S(\Phi_1, \Psi_1) = 0.5$ and $cer_S(\Phi_2, \Psi_2) = 1$. Therefore,

$$POS_{\mathcal{A}}(\{d\}) = \{x_3\} = \|\Phi_2\|_S = \bigvee_{\Phi \rightarrow \Psi \in Dec^+(S)} \|\Phi\|_S$$

As a consequence, $Dec(S)$ is a decision algorithm and the object x_2 does not satisfy any decision rule of $Dec(S)$. $\square$

Decision algorithms where all the objects of the decision table satisfy a decision rule of that algorithm were also defined by Pawlak in [13] and they are recalled now.

Definition 9. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $DA(S)$ be a decision algorithm. If for each $x \in U$ there exists $\Phi \rightarrow \Psi \in DA(S)$ such that $x \in \|\Phi \wedge \Psi\|_S$ then $DA(S)$ is called a *complete decision algorithm*.

An important consequence of Definition 9, taking into account the first item of Definition 7, is that the addition of the supports of all decision rules is the cardinal of the set of objects, since for each object there exists only one decision rule which is satisfied by that object. Therefore, the addition of the strengths of all decision rules is 1.

3. Decision rules in fuzzy rough set theory

This section will focus on the study of decision tables in the fuzzy environment by using decision rules. With this purpose, it is needed to recall some important notions of FRST [17]. Most of them are the generalization of the previous definitions to the fuzzy framework. The first definition is essential to compare a pair of objects of a decision table according to an attribute. It presents a map that, evaluated on two objects, indicates the relationship between these objects, taking into account the considered attribute.

Definition 10. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $a \in \mathcal{A}_d$. An a -indiscernibility relation is a $[0, 1]$ -fuzzy tolerance relation $R_a: U \times U \rightarrow [0, 1]$, that is, a reflexive and symmetrical fuzzy relationship.

Depending on the value of $R_a(x, y)$ we can conclude how similar the objects x, y are, according to the attribute a . Next definition is required to compare pairs of objects by using a subset of attributes $B \subseteq \mathcal{A}$ instead of a single attribute. Aggregation operators [17] are indispensable for that.

Definition 11. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $R_a: U \times U \rightarrow [0, 1]$ be an a -indiscernibility relation for all $a \in \mathcal{A}$. The B -indiscernibility relation $R_B: U \times U \rightarrow [0, 1]$ is defined for each pair of objects $x, y \in U$ as

$$R_B(x, y) = @(\mathcal{R}_B^{x,y}(a_1), \dots, \mathcal{R}_B^{x,y}(a_m))$$

where $\mathcal{R}_B^{x,y}: \mathcal{A} \rightarrow [0, 1]$ is defined for each $a \in \mathcal{A}$ as

$$\mathcal{R}_B^{x,y}(a) = \begin{cases} R_a(x, y) & \text{if } a \in B \\ 1 & \text{otherwise} \end{cases}$$

and $@: [0, 1]^m \rightarrow [0, 1]$ is an aggregation operator, that is, an increasing operator on each argument satisfying $@(1, \dots, 1) = 1$ and $@(0, \dots, 0) = 0$.

The interpretation of the value given by this mapping is the same as Definition 10. In addition, notice that R_B is a $[0, 1]$ -fuzzy tolerance relation for all $B \subseteq \mathcal{A}$.

In the following, some previous notions given in [17] are recalled, which are necessary to define the positive region in the fuzzy environment. First of all, we introduce the notion of adjoint triple.

Definition 12. Let $(P_1, \leq_1)$, $(P_2, \leq_2)$ and $(P_3, \leq_3)$ be partially ordered sets (posets) and $\&: P_1 \times P_2 \rightarrow P_3$, $\swarrow: P_3 \times P_2 \rightarrow P_1$ and $\searrow: P_3 \times P_1 \rightarrow P_2$ be mappings. Then $(\&, \swarrow, \searrow)$ is an *adjoint triple* with respect to P_1, P_2 and P_3 if these mappings satisfy the adjoint property, that is, $x \leq_1 z \swarrow y$ iff $x \& y \leq_3 z$ iff $y \leq_2 z \searrow x$ for all $x \in P_1, y \in P_2, z \in P_3$.

Notice that, the mappings $\swarrow, \searrow$ can be univocally obtained from the operator $\&$. Therefore, we can refer to the adjoint triple $(\&, \swarrow, \searrow)$ simply as $\&$. On the other hand, in the particular case that the mapping $\&$ is commutative and $P_1 = P_2$, we obtain that $\swarrow = \searrow$. More properties can be seen in [23].

Left-continuous t-norms together with their corresponding residuated implications clearly are particular cases of adjoint triples [24], such as the well-known Gödel, product and Łukasiewicz t-norms. In order to illustrate the impact of the use of adjoint triples in the management of information, a pair of adjoint triples are presented in the following example.

Example 13. We include an adjoint triple $(\&, \swarrow, \searrow)$ with respect to the unit interval in which the mapping $\&$ is not commutative. This fact will lead to obtain two different mappings $\swarrow$ and $\searrow$.

$$\begin{aligned} x \& y &= x^2 y \\ z \swarrow y &= \begin{cases} 1 & \text{if } y \leq z \\ \sqrt{z/y} & \text{otherwise} \end{cases} \\ z \searrow x &= \begin{cases} 1 & \text{if } x^2 \leq z \\ z/x^2 & \text{otherwise} \end{cases} \end{aligned}$$

From this adjoint triple we can deduce that the values of the variable x are penalized by the square (the square of values less or equal to 1 provides a less value than the original). Hence, we can say that with the use of this conjunctive the variable y has a major impact (for obtaining a large aggregated value) than the variable x . This fact allows us to give different importance to each variable of the corresponding study.

Now, we present the family of Łukasiewicz adjoint triples $(\&_{\mathbf{t}}^{\beta}, \swarrow_{\mathbf{t}}^{\beta}, \searrow_{\mathbf{t}}^{\beta})$, which depends on a parameter $\beta \in \mathbb{N}$:

$$\begin{aligned} x \&_{\mathbf{t}}^{\beta} y &= \max\{0, \sqrt[\beta]{x^{\beta} + y^{\beta} - 1}\} \\ z \swarrow_{\mathbf{t}}^{\beta} x &= \sqrt[\beta]{\min\{1, 1 + z^{\beta} - x^{\beta}\}} \end{aligned} \quad (1)$$

Since the mapping $\&_{\mathbf{t}}^{\beta}$ is commutative we obtain that $\swarrow_{\mathbf{t}}^{\beta} = \searrow_{\mathbf{t}}^{\beta}$. In this family of adjoint triples, the parameter β affects to the monotonicity. Specifically, it can be checked that $\&_{\mathbf{t}}^{\beta}$ is decreasing in β and $\swarrow_{\mathbf{t}}^{\beta}$ is increasing in β , as it is shown in Fig. 1. Hence, the user can consider different values of β depending on the objects under consideration, choosing higher or lower values to give them different importance [17]. $\square$

Now, we recall the notion of multi-adjoint property-oriented frame.

Definition 14. Given a poset $(P_1, \leq_1)$, two complete lattices $(L_2, \leq_2)$ and $(L_3, \leq_3)$ and adjoint triples $(\&_i, \swarrow^i, \searrow_i)$, with respect to P_1, L_2, L_3 , for all $i \in \{1, \dots, n\}$, a *multi-adjoint property-oriented frame* is the tuple

$$(P_1, L_2, L_3, \&_1, \dots, \&_n)$$

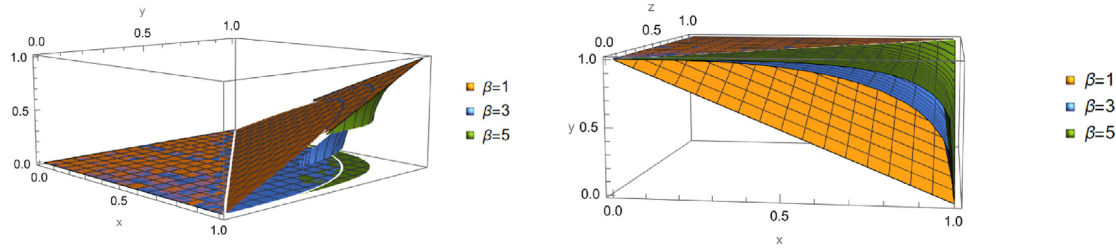


Fig. 1. Operators $\&L^\beta$ (left) and $<L^\beta$ (right), with $\beta \in \{1, 3, 5\}$.

Next, the notion of context is introduced.

Definition 15. Let A and B be non-empty sets and $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ be a multi-adjoint property-oriented frame. A *context* is a tuple (A, B, R, τ) , where R is a P_1 -fuzzy relation $R: A \times B \rightarrow P_1$ and $\tau: A \times B \rightarrow \{1, \dots, n\}$ is a mapping which associates any pair of elements in $A \times B$ with some particular adjoint triple in the frame.

In the following definition, we recall the generalization of the notion of upper and lower approximation of fuzzy RST given in [17].

Definition 16. Let $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ be a multi-adjoint property oriented frame and (A, B, R, τ) be a context. Given $g \in L_2^B$ and $f \in L_3^A$ we define the *possibility and necessity operators*, $\uparrow^\pi: L_2^B \rightarrow L_3^A$ and $\downarrow^N: L_3^A \rightarrow L_2^B$, respectively, as:

$$g^{\uparrow^\pi}(a) = \sup\{R(a, b) \&_{\tau(a, b)} g(b) \mid b \in B\}$$

$$f^{\downarrow^N}(b) = \inf\{f(a) \<_{\tau(a, b)} R(a, b) \mid a \in A\}$$

where $g^{\uparrow^\pi}$ is interpreted as the upper approximation of g and $f^{\downarrow^N}$ as the lower approximation of f .

Notice that, the use of the mapping τ allows the consideration of different degrees of preference on the sets of objects and attributes [25]. Furthermore, the possibility of taking into account non-commutative operators allows to consider different levels of relevance in the computation of the aggregated value. For example, if the first adjoint triple in Example 13 is considered in the possibility operator, then we are assuming that it is more relevant the values $g(b)$ (the object b belong to the fuzzy subset g) than the values given by the relation.

In this paper, we will focus on a multi-adjoint property-oriented frame $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$ and a context (U, U, R_B, τ) with $B \subseteq A$. Finally, we recall the notion of positive region (generalization of Definition 3) in the fuzzy setting when R_d only takes boolean values, which was given in [17].

Definition 17. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. Given a multi-adjoint property oriented frame $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$, a context (U, U, R_B, τ) and a boolean relation R_d , the *multi-adjoint fuzzy B-positive region* is defined, for each $y \in U$, as:

$$\text{POS}_B^f(y) = (R_d y)^{\downarrow^N}(y) = \inf\{(R_d y)(x) \<_{\tau(x, y)} R_B(x, y) \mid x \in U\}$$

where $R_d y: U \rightarrow \{0, 1\}$ is defined as $(R_d y)(x) = R_d(y, x)$ and $\<_{\tau(x, y)}$ is the left residuated fuzzy implication of $\&_{\tau(x, y)}$ associated with the pair of objects x, y .

The authors in [17] asserted that Definition 17 is a generalization of Definition 3. The following result shows the details of this statement.

Proposition 18. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $x \in U$ and R_A and R_a be boolean relations for all $a \in \mathcal{A}_d$. Then $x \in \text{POS}_{\mathcal{A}}(\{d\})$ if and only if $\text{POS}_{\mathcal{A}}^f(x) = 1$.

Proof. First of all, notice that if R_A is a boolean relation then, given two objects $x, y \in U$, $R_A(x, y) = 1$ if and only if $R_a(x, y) = 1$ for all $a \in \mathcal{A}$, that is, $\bar{a}(x) = \bar{a}(y)$ for all $a \in \mathcal{A}$, and $R_A(x, y) = 0$ otherwise. Let $x \in \text{POS}_{\mathcal{A}}(\{d\})$ and $y \in U$. We distinguish two cases.

- We suppose that $y \in [x]_{I(\mathcal{A})}$. Since $x \in \text{POS}_{\mathcal{A}}(\{d\})$ we obtain that $[x]_{I(\mathcal{A})} \subseteq [x]_{I(d)}$. Therefore, $y \in [x]_{I(d)}$. Hence, $R_a(x, y) = 1$ for all $a \in \mathcal{A}_d$. As a consequence,

$$R_d(y, x) \<_{\tau(x, y)} R_A(x, y) = 1 \<_{\tau(x, y)} 1 = 1$$

- Now, we suppose that $y \notin [x]_{I(\mathcal{A})}$. Then there exists $a \in \mathcal{A}$ such that $\bar{a}(x) \neq \bar{a}(y)$. Hence, $R_A(x, y) = 0$. As a consequence,

$$R_d(y, x) \<_{\tau(x, y)} R_A(x, y) = R_d(y, x) \<_{\tau(x, y)} 0 = 0$$

In conclusion, $R_d(y, x) \frown_{\tau(x,y)} R_A(x, y) = 1$ for all $y \in U$. Therefore,

$$\text{POS}_A^f(x) = (R_d y)^{\downarrow^N}(y) = \inf\{R_d(y, x) \frown_{\tau(x,y)} R_A(x, y) \mid x \in U\} = 1$$

Now, we suppose that $\text{POS}_A^f(x) = 1$. As a consequence, we obtain that $R_d(y, x) \frown_{\tau(x,y)} R_A(x, y) = 1$ for all $y \in U$. If $x \notin \text{POS}_A(\{d\})$ then there exists $y \in [x]_{I(A)} \setminus [x]_{I(d)}$. Hence, $R_A(x, y) = 1$ and $R_d(x, y) = 0$. Therefore,

$$R_d(y, x) \frown_{\tau(x,y)} R_A(x, y) = 0 \frown_{\tau(x,y)} 1 = 0,$$

obtaining a contradiction. In conclusion, $x \in \text{POS}_A(\{d\})$. $\square$

Since in the computation of $\text{POS}_A^f(y)$ all the objects of the decision table are taken into account, by using indiscernibility relations and a fuzzy implication, this notion provides a degree of consistency of the object y in the decision table. As a consequence, the higher the value of $\text{POS}_A^f(y)$, the more dependence there will be between the set of attributes A and the decision attribute d for the object y .

Real datasets usually present numerous attributes, making difficult the management of the information contained in them. The notion of fuzzy decision reduct arises to reduce the number of attributes without losing information, which enables us to deal with them in an easier way. Before introducing this notion, we must recall the following definition.

Definition 19. Let $S = (U, A_d, \mathcal{V}_{A_d}, \overline{A_d})$ be a decision table, $B \subseteq A$ and $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$ be a multi-adjoint property oriented frame. A monotonic mapping $m: \mathcal{P}(A) \rightarrow [0, 1]$ is a $[0, 1]$ -valued measure associated with S if the condition $R_d(x, y) \frown_{\tau(x,y)} R_B(x, y) = R_d(x, y) \frown_{\tau(x,y)} R_A(x, y)$, for all $x, y \in U$ implies $m(B) = 1$.

Now, we can present the notion of fuzzy decision reduct, which is supported by $[0, 1]$ -valued measures.

Definition 20 ([17]). Let $S = (U, A_d, \mathcal{V}_{A_d}, \overline{A_d})$ be a decision table, $B \subseteq A$, m be a $[0, 1]$ -valued measure associated with S and $\alpha \in (0, 1]$. The set B is called a *fuzzy m -decision superreduct to degree α* if $\alpha \preccurlyeq m(B)$. Moreover, if $\alpha \not\preccurlyeq m(B')$ for each $B' \subset B$ then B is called *fuzzy m -decision reduct to degree α* .

Now, we recall the notion of cardinal of a fuzzy set, which will be used later. This definition is introduced focusing on the multi-adjoint fuzzy positive region.

Definition 21. Given a universe U and $f: U \rightarrow [0, 1]$, the *cardinal* of f is given as

$$\text{card}_f(f) = \sum_{x \in U} f(x)$$

Furthermore, it is necessary to generalize the notions defined in Sections 2.1 and 2.2 to this framework.

First of all, we introduce a particular kind of tolerance relation that will be used in the rest of the paper.

Definition 22 ([26]). Let $S = (U, A_d, \mathcal{V}_{A_d}, \overline{A_d})$ be a decision table, $a \in A$ and $T_a: V_a \times V_a \rightarrow [0, 1]$ be a tolerance relation. If $T_a(v, w) = 1$ implies $v = w$ then T_a is called a *separable tolerance relation*.

Notice that, the counterpart is verified for all tolerance relations since they are reflexive. As a consequence, a separable tolerance relation T_a satisfies that $T_a(v, w) = 1$ if and only if $v = w$.

Regarding to the notion of formula in the fuzzy setting, we must emphasize that this notion is the same as Definition 4. The reason is that the value v can be any element of an arbitrary set associated with the attribute a , providing a great level of flexibility, which is in consonance with the fuzzy environment. However, $\|\Phi\|_S$ has a different meaning.

Definition 23. Let $S = (U, A_d, \mathcal{V}_{A_d}, \overline{A_d})$ be a decision table, $B \subseteq A_d$ and $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in A_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. The mapping $\|\cdot\|_S^T: \text{For}(B) \rightarrow [0, 1]^U$ is inductively defined as:

$$\|\Phi\|_S^T(x) = T_a(\bar{a}(x), v)$$

for all $x \in U$ and $\Phi = (a, v)$, where $a \in B$ and $v \in V_a$. For every $\Phi, \Psi \in \text{For}(B)$, the conjunction and disjunction of formulas are defined, for all $x \in U$, as follows:

$$\|\Phi \wedge \Psi\|_S^T(x) = \inf\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$

$$\|\Phi \vee \Psi\|_S^T(x) = \sup\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$

Therefore, $\|\Phi\|_S^T(x)$ represents how much the object x satisfies the formula Φ , through the relationships between the value of the attribute a in the object x and the value of the attribute a in the formula Φ .

Notice that, the infimum and the supremum are the extensions of the intersection and the union of sets to the fuzzy environment, respectively. In this way, $\|\Phi \wedge \Psi\|_S^T(x)$ and $\|\Phi \vee \Psi\|_S^T(x)$ generalize the conjunction and disjunction of formulas in the classical environment given in Definition 4. In addition, Definition 23 generalizes Definition 4 when boolean separable tolerance relations are considered, as we show below.

Table 2Table associated with the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 25.

	a_1	a_2	a_3	a_4	d
x_1	0.34	0.31	0.75	0.2	0
x_2	0.21	0.71	0.5	0.2	1
x_3	0.52	0.92	1	0.7	0
x_4	0.85	0.65	1	0.7	1
x_5	0.43	0.89	0.5	0.2	0
x_6	0.21	0.47	0.25	0.5	1
x_7	0.09	0.93	0.25	0.5	0

Proposition 24. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $B \subseteq \mathcal{A}_d$, $\Phi \in \text{For}(B)$ and $x \in U$. If $T = \{T_a : V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ is a family of boolean separable tolerance relations then the following property holds:

$$x \in \|\Phi\|_S \text{ if and only if } \|\Phi\|_S^T(x) = 1$$

Proof. We will consider an attribute-value pair $\Phi = (a, v)$ and T_a a boolean separable tolerance relation. The proof can be extended inductively to the conjunction or disjunction of attribute-value pairs. Suppose that $x \in \|\Phi\|_S = \|(a, v)\|_S$. Then, x satisfies the formula Φ , that is, $\bar{a}(x) = v$. As a consequence, since T_a is reflexive, we obtain that $T_a(\bar{a}(x), v) = 1$. Hence, $\|\Phi\|_S^T(x) = 1$.

The proof of the counterpart is completely analogous taking into account that T_a is a separable relation. $\square$

On the other hand, the definition of decision rules in FRST is the same as Definition 5. An example is presented in order to show the notions of formula and decision rule in FRST.

Example 25. Consider the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ where the set of objects is $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, the set of attributes is $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ and $\mathcal{V}_a = [0, 1]$ for all $a \in \mathcal{A}_d$. This decision table is represented in Table 2. In this example, we will show the decision rules that can be obtained from these data considering a reduct instead of all the attributes $\mathcal{A}$. In addition, we will compute how much each object satisfies each antecedent and consequent of the decision rules, respectively.

First of all, we are interested in finding a fuzzy m -decision reduct to degree 1 in order to deal with the decision table S without considering all the attributes while all the information contained in S is preserved. With this purpose, we will use the $[0, 1]$ -valued measure $m : \mathcal{P}(\mathcal{A}) \rightarrow [0, 1]$ defined for each $B \subseteq \mathcal{A}$ as

$$m(B) = \text{card}_F(\text{POS}_B^f) \swarrow^L \text{card}_F(\text{POS}_{\mathcal{A}}^f) \quad (2)$$

$$= \min \left\{ 1, 1 - \frac{\text{card}_F(\text{POS}_{\mathcal{A}}^f)}{|U|} + \frac{\text{card}_F(\text{POS}_B^f)}{|U|} \right\} \quad (3)$$

Now, we fix the set $B = \{a_1, a_2, a_3\}$ in order to check if it is possible to discard the attribute a_4 without losing information. Next step is to compute $\text{POS}_{\mathcal{A}}^f(x)$ and $\text{POS}_B^f(x)$, for all $x \in U$. For this purpose, we consider the a -indiscernibility relation $R_a : U \times U \rightarrow [0, 1]$ given as:

$$R_a(x, y) = 1 - |\bar{a}(x) - \bar{a}(y)| \quad (4)$$

for all $a \in \mathcal{A}_d$ and $x, y \in U$. Furthermore, we will consider the following indiscernibility relations:

$$R_{\mathcal{A}}(x, y) = \frac{R_{a_1}(x, y) + 2(R_{a_2}(x, y) + R_{a_3}(x, y)) + R_{a_4}(x, y)}{6}$$

$$R_B(x, y) = \frac{R_{a_1}(x, y) + 2(R_{a_2}(x, y) + R_{a_3}(x, y)) + 1}{6}$$

for all $x, y \in U$. Moreover, we define a fuzzy implication in order to calculate the fuzzy positive region in the context $(U, U, R_{\mathcal{A}}, \tau)$. It will depend on the objects under consideration and on a value β given from the mapping $\tau : U \times U \rightarrow \{1, 3, 5\}$ defined for each $x_i, x_j \in U$ as:

$$\tau(x_i, x_j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are even} \\ 3 & \text{if } i \text{ and } j \text{ are odd} \\ 5 & \text{otherwise} \end{cases}$$

We will consider the fuzzy implication in Eq. (1), that is, $\lhd_{\tau}^{\beta} : [0, 1] \times [0, 1] \rightarrow [0, 1]$ defined, for each $x_i, x_j \in U$, as:

$$x_i \lhd_{\tau}^{\beta} x_j = \sqrt[\beta]{\min\{1, 1 + x_i^{\beta} - x_j^{\beta}\}} \quad (5)$$

Table 3
Relation R_{a_1} of Example 25.

R_{a_1}	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.87	0.82	0.49	0.91	0.87	0.75
x_2	0.87	1	0.69	0.36	0.78	1	0.88
x_3	0.82	0.69	1	0.67	0.91	0.69	0.57
x_4	0.49	0.36	0.67	1	0.58	0.36	0.24
x_5	0.91	0.78	0.91	0.58	1	0.78	0.66
x_6	0.87	1	0.69	0.36	0.78	1	0.88
x_7	0.75	0.88	0.57	0.24	0.66	0.88	1

Table 4
Relation R_{a_2} of Example 25.

R_{a_2}	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.6	0.39	0.66	0.42	0.84	0.38
x_2	0.6	1	0.79	0.94	0.82	0.76	0.78
x_3	0.39	0.79	1	0.73	0.97	0.55	0.99
x_4	0.66	0.94	0.73	1	0.76	0.82	0.72
x_5	0.42	0.82	0.97	0.76	1	0.58	0.96
x_6	0.84	0.76	0.55	0.82	0.58	1	0.54
x_7	0.38	0.78	0.99	0.72	0.96	0.54	1

Table 5
Relation R_{a_3} of Example 25.

R_{a_3}	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.75	0.75	0.75	0.75	0.5	0.5
x_2	0.75	1	0.5	0.5	1	0.75	0.75
x_3	0.75	0.5	1	1	0.5	0.25	0.25
x_4	0.75	0.5	1	1	0.5	0.25	0.25
x_5	0.75	1	0.5	0.5	1	0.75	0.75
x_6	0.5	0.75	0.25	0.25	0.75	1	1
x_7	0.5	0.75	0.25	0.25	0.75	1	1

Table 6
Relation R_{a_4} of Example 25.

R_{a_4}	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	1	0.5	0.5	1	0.7	0.7
x_2	1	1	0.5	0.5	1	0.7	0.7
x_3	0.5	0.5	1	1	0.5	0.8	0.8
x_4	0.5	0.5	1	1	0.5	0.8	0.8
x_5	1	1	0.5	0.5	1	0.7	0.7
x_6	0.7	0.7	0.8	0.8	0.7	1	1
x_7	0.7	0.7	0.8	0.8	0.7	1	1

with $\tau(x_i, x_j) = \beta$. Notice that, the case $\beta = 1$ corresponds to the Łukasiewicz implication. The relations $R_{a_1}, R_{a_2}, R_{a_3}, R_{a_4}, R_B$ and R_A given in Tables 3–8 will be needed to compute $\text{POS}_A^f(x)$ and $\text{POS}_B^f(x)$, for all $x \in U$. In order to illustrate the computations, we include the following examples:

$$\begin{aligned}
 R_{a_1}(x_1, x_2) &= 1 - |\bar{a}_1(x_1) - \bar{a}_1(x_2)| = 1 - |0.34 - 0.21| = 0.87 \\
 R_A(x_1, x_2) &= \frac{R_{a_1}(x_1, x_2) + 2(R_{a_2}(x_1, x_2) + R_{a_3}(x_1, x_2)) + R_{a_4}(x_1, x_2)}{6} \\
 &= \frac{0.87 + 2(0.6 + 0.75) + 1}{6} \\
 &= \frac{4.57}{6} \approx 0.76
 \end{aligned}$$

On the other hand, since d is a boolean attribute, the tolerance relation $R_d(x, y) = 1 - |\bar{d}(x) - \bar{d}(y)|$ is given as follows:

$$R_d(x, y) = \begin{cases} 1 & \text{if } \bar{d}(x) = \bar{d}(y) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Table 7
 R_B of Example 25.

R_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.76	0.68	0.72	0.71	0.76	0.59
x_2	0.76	1	0.71	0.71	0.9	0.84	0.82
x_3	0.68	0.71	1	0.86	0.81	0.55	0.68
x_4	0.72	0.71	0.86	1	0.68	0.58	0.53
x_5	0.71	0.9	0.81	0.68	1	0.74	0.85
x_6	0.76	0.84	0.55	0.58	0.74	1	0.83
x_7	0.59	0.82	0.68	0.53	0.85	0.83	1

Table 8
 R_A of Example 25.

R_A	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.76	0.6	0.64	0.71	0.71	0.54
x_2	0.76	1	0.63	0.62	0.9	0.79	0.77
x_3	0.6	0.63	1	0.86	0.73	0.52	0.64
x_4	0.64	0.62	0.86	1	0.6	0.55	0.5
x_5	0.71	0.9	0.73	0.6	1	0.69	0.8
x_6	0.71	0.79	0.52	0.55	0.69	1	0.83
x_7	0.54	0.77	0.64	0.5	0.8	0.83	1

Table 9
 $\text{POS}_{\mathcal{A}}^f$ of Example 25.

	$\text{POS}_{\mathcal{A}}^f$
x_1	0.94
x_2	0.83
x_3	0.89
x_4	0.89
x_5	0.83
x_6	0.91
x_7	0.91

Finally, Tables 7 and 8 show the relations R_B and R_A . From Tables 7 and 8 the multi-adjoint fuzzy B -positive region and A -positive region are computed, obtaining that $\text{POS}_{\mathcal{A}}^f(x) = \text{POS}_B^f(x)$, for all $x \in U$. These values are shown in Table 9. As a consequence, it is immediately obtained that $m(B) = 1$.

In a similar way it can be computed that $m(\{a_1, a_2\}) = 0.97$, $m(\{a_1, a_3\}) = 0.86$ and $m(\{a_2, a_3\}) = 0.98$. Therefore, by Definition 20, B is a fuzzy m -decision reduct to degree 1. In this way, we can discard the attribute a_4 without losing information. Thus, we will only consider the attributes a_1 , a_2 and a_3 , and by using the conjunction of formulas, to compute the set of decision rules which will describe the decision table of Table 2. In this way, one decision rule is obtained for each object. For instance, for the object x_1 we obtain the formulas $(a_1, 0.34)$, $(a_2, 0.31)$ and $(a_3, 0.75)$ considering each attribute of $\mathcal{A}$. By using the conjunction of these pairs we obtain the formula $\Phi_1 = (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75)$. On the other hand, considering the attribute d we obtain the formula $\Psi_1 = (d, 0)$. As a consequence, the decision rule $r_1 = \Phi_1 \rightarrow \Psi_1$ associated with the object x_1 is $r_1 : (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75) \rightarrow (d, 0)$. Following an analogous reasoning next decision rules are obtained.

$$\begin{aligned}
r_1 &: (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75) \rightarrow (d, 0) \\
r_2 &: (a_1, 0.21) \wedge (a_2, 0.71) \wedge (a_3, 0.5) \rightarrow (d, 1) \\
r_3 &: (a_1, 0.52) \wedge (a_2, 0.92) \wedge (a_3, 1) \rightarrow (d, 0) \\
r_4 &: (a_1, 0.85) \wedge (a_2, 0.65) \wedge (a_3, 1) \rightarrow (d, 1) \\
r_5 &: (a_1, 0.43) \wedge (a_2, 0.89) \wedge (a_3, 0.5) \rightarrow (d, 0) \\
r_6 &: (a_1, 0.21) \wedge (a_2, 0.47) \wedge (a_3, 0.25) \rightarrow (d, 1) \\
r_7 &: (a_1, 0.09) \wedge (a_2, 0.93) \wedge (a_3, 0.25) \rightarrow (d, 0)
\end{aligned}$$

Now, we compute the degree of satisfaction to each antecedent for each object, that is, $\|\Phi_i\|_S^T(x_j)$ with $i, j \in \{1, 2, \dots, 7\}$. For this purpose, we use the family of separable $[0, 1]$ -fuzzy tolerance relation $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}\}$ defined as $T_a(\bar{a}(x), v) = 1 - |\bar{a}(x) - v|$, for all $a \in \mathcal{A}$, $x \in U$ and $v \in [0, 1]$. Notice that, $\|\Phi_i\|_S^T(x_j) = \|\Psi_i\|_S^T(x_j) = 1$ if $i = j$, since each object generates the decision rule with the same subscript. The computation of $\|\Phi_2\|_S^T(x_1)$ is shown for a non-trivial example.

$$\begin{aligned}
\|\Phi_2\|_S^T(x_1) &= \|(a_1, 0.21) \wedge (a_2, 0.71) \wedge (a_3, 0.5)\|_S^T(x_1) \\
&= \|(a_1, 0.21)\|_S^T(x_1) \wedge \|(a_2, 0.71)\|_S^T(x_1) \wedge \|(a_3, 0.5)\|_S^T(x_1)
\end{aligned}$$

Table 10

Degree of satisfaction to each antecedent of decision rules for each object.

	$\ \Phi_1\ _S^T$	$\ \Phi_2\ _S^T$	$\ \Phi_3\ _S^T$	$\ \Phi_4\ _S^T$	$\ \Phi_5\ _S^T$	$\ \Phi_6\ _S^T$	$\ \Phi_7\ _S^T$
x_1	1	0.6	0.39	0.49	0.42	0.5	0.38
x_2	0.6	1	0.5	0.36	0.78	0.75	0.75
x_3	0.39	0.5	1	0.67	0.5	0.25	0.25
x_4	0.49	0.36	0.67	1	0.5	0.25	0.24
x_5	0.42	0.78	0.5	0.5	1	0.58	0.66
x_6	0.5	0.75	0.25	0.25	0.58	1	0.54
x_7	0.38	0.75	0.25	0.24	0.66	0.54	1

$$\begin{aligned}
&= (1 - |\bar{a}_1(x_1) - 0.21|) \wedge (1 - |\bar{a}_2(x_1) - 0.71|) \wedge (1 - |\bar{a}_3(x_1) - 0.5|) \\
&= (1 - |0.34 - 0.21|) \wedge (1 - |0.31 - 0.71|) \wedge (1 - |0.75 - 0.5|) \\
&= 0.87 \wedge 0.6 \wedge 0.75 = 0.6
\end{aligned}$$

Following an analogous procedure, the rest of calculations are obtained. They are shown in Table 10.

On the other hand, we compute $\|\Psi_i\|_S^T(x_j) = \|(d, v_i)\|_S^T(x_j)$ with $i, j \in \{1, 2, \dots, 7\}$. In this case, since d is a boolean attribute the calculations are easier, obtaining

$$\|\Psi_i\|_S^T(x_j) = \begin{cases} 1 & \text{if } \bar{d}(x_j) = v_i \\ 0 & \text{otherwise} \end{cases} \quad \square \quad (7)$$

Now, we present the notions of support, strength, certainty and coverage in the fuzzy environment, in order to describe decision rules in this framework. These notions generalize the classical ones, so they have an analogous interpretation.

Definition 26. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\Phi \rightarrow \Psi$ be a decision rule in S and $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. We call:

- T -support of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$\text{supp}_S^T(\Phi, \Psi) = \text{card}_F(\|\Phi \wedge \Psi\|_S^T)$$

- T -strength of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$\sigma_S^T(\Phi, \Psi) = \frac{\text{supp}_S^T(\Phi, \Psi)}{|U|}$$

- T -certainty of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$\text{cer}_S^T(\Phi, \Psi) = \frac{\text{supp}_S^T(\Phi, \Psi)}{\text{card}_F(\|\Phi\|_S^T)}$$

when $\text{card}_F(\|\Phi\|_S^T) \neq 0$.

- T -coverage of the decision rule $\Phi \rightarrow \Psi$ to the value:

$$\text{cov}_S^T(\Phi, \Psi) = \frac{\text{supp}_S^T(\Phi, \Psi)}{\text{card}_F(\|\Psi\|_S^T)}$$

when $\text{card}_F(\|\Psi\|_S^T) \neq 0$.

Notice that, $\text{cer}_S^T(\Phi, \Psi) \leq 1$ for each $\Phi \rightarrow \Psi$ in S since $\|\Phi \wedge \Psi\|_S^T(x) = \|\Phi\|_S^T(x) \wedge \|\Psi\|_S^T(x) \leq \|\Phi\|_S^T(x)$ for all $x \in U$. In the same way, it can be concluded that $\sigma_S^T(\Phi, \Psi), \text{cov}_S^T(\Phi, \Psi) \leq 1$. Analogously to the classical environment, we characterize decision rules depending on the value of its T -certainty.

Definition 27. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\Phi \rightarrow \Psi$ be a decision rule in S and $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. If $\text{cer}_S^T(\Phi, \Psi) = 1$ we will say that $\Phi \rightarrow \Psi$ is a T -true decision rule. If $\text{cer}_S^T(\Phi, \Psi) = 0$, we say that the decision rule is T -false. Otherwise, it will be called a *not entirely T -true decision rule*.

Now, we prove that these notions generalize to the classical ones.

Proposition 28. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $\Phi \rightarrow \Psi$ be a decision rule in S . If $T = \{T_a : V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ is a family of boolean separable tolerance relations then

$$\text{supp}_S^T(\Phi, \Psi) = \text{supp}_S(\Phi, \Psi)$$

$$\begin{aligned}\sigma_S^T(\Phi, \Psi) &= \sigma_S(\Phi, \Psi) \\ \text{cer}_S^T(\Phi, \Psi) &= \text{cer}_S(\Phi, \Psi) \\ \text{cov}_S^T(\Phi, \Psi) &= \text{cov}_S(\Phi, \Psi)\end{aligned}$$

Proof. Let $B \subseteq \mathcal{A}_d$. Since it is well-known that card_F generalizes the classical definition of cardinality, we have that $\|\Phi\|_S = \text{card}_F(\|\Phi\|_S^T)$ with $\Phi \in \text{For}(B)$. Thus, because of the notions of T -support, T -strength, T -certainty and T -coverage are based on the cardinal of fuzzy sets we obtain that these notions generalize to the classical ones. $\square$

These notions are illustrated in the following example.

Example 29. Coming back to the environment of [Example 25](#), we compute the T -support, T -strength, T -certainty and T -coverage of each decision rule by using [Table 10](#), [Eq. \(7\)](#) and [Definition 26](#). We will only show some calculus since the rest are analogous.

First of all, we will calculate $\text{supp}_S^T(\Phi_1, \Psi_1)$.

$$\begin{aligned}\text{supp}_S^T(\Phi_1, \Psi_1) &= \text{card}_F(\|\Phi_1 \wedge \Psi_1\|_S^T) \\ &= \sum_{x \in U} \|\Phi_1 \wedge \Psi_1\|_S^T(x) \\ &= \sum_{x \in U} (\|\Phi_1\|_S^T(x) \wedge \|\Psi_1\|_S^T(x)) \\ &= \|\Phi_1\|_S^T(x_1) \wedge \|\Psi_1\|_S^T(x_1) + \|\Phi_1\|_S^T(x_2) \wedge \|\Psi_1\|_S^T(x_2) \\ &\quad + \|\Phi_1\|_S^T(x_3) \wedge \|\Psi_1\|_S^T(x_3) + \|\Phi_1\|_S^T(x_4) \wedge \|\Psi_1\|_S^T(x_4) \\ &\quad + \|\Phi_1\|_S^T(x_5) \wedge \|\Psi_1\|_S^T(x_5) + \|\Phi_1\|_S^T(x_6) \wedge \|\Psi_1\|_S^T(x_6) \\ &\quad + \|\Phi_1\|_S^T(x_7) \wedge \|\Psi_1\|_S^T(x_7) \\ &= 1 \wedge 1 + 0.6 \wedge 0 + 0.39 \wedge 1 + 0.49 \wedge 0 + 0.42 \wedge 1 + 0.5 \wedge 0 + 0.38 \wedge 1 \\ &= 1 + 0 + 0.39 + 0 + 0.42 + 0 + 0.38 \\ &= 2.19\end{aligned}$$

Hence, the T -strength of the decision rule $\Phi_1 \rightarrow \Psi_1$ is

$$\sigma_S^T(\Phi_1, \Psi_1) = \frac{\text{supp}_S^T(\Phi_1, \Psi_1)}{|U|} = \frac{2.19}{7} \approx 0.31$$

In order to compute the T -certainty of this decision rule we calculate $\text{card}_F(\|\Phi_1\|_S^T)$.

$$\begin{aligned}\text{card}_F(\|\Phi_1\|_S^T) &= \sum_{x \in U} \|\Phi_1\|_S^T(x) = \|\Phi_1\|_S^T(x_1) + \|\Phi_1\|_S^T(x_2) + \|\Phi_1\|_S^T(x_3) \\ &\quad + \|\Phi_1\|_S^T(x_4) + \|\Phi_1\|_S^T(x_5) + \|\Phi_1\|_S^T(x_6) + \|\Phi_1\|_S^T(x_7) \\ &= 1 + 0.6 + 0.39 + 0.49 + 0.42 + 0.5 + 0.38 \\ &= 3.78\end{aligned}$$

As a consequence,

$$\text{cer}_S^T(\Phi_1, \Psi_1) = \frac{\text{supp}_S^T(\Phi_1, \Psi_1)}{\text{card}_F(\|\Phi_1\|_S^T)} = \frac{2.19}{3.78} \approx 0.58$$

Now, we calculate the T -coverage of the decision rule $\Phi_1 \rightarrow \Psi_1$. With this purpose, we compute $\text{card}_F(\|\Psi_1\|_S^T)$.

$$\begin{aligned}\text{card}_F(\|\Psi_1\|_S^T) &= \sum_{x \in U} \|\Psi_1\|_S^T(x) = \|\Psi_1\|_S^T(x_1) + \|\Psi_1\|_S^T(x_2) + \|\Psi_1\|_S^T(x_3) \\ &\quad + \|\Psi_1\|_S^T(x_4) + \|\Psi_1\|_S^T(x_5) + \|\Psi_1\|_S^T(x_6) + \|\Psi_1\|_S^T(x_7) \\ &= 1 + 0 + 1 + 0 + 1 + 0 + 1 \\ &= 4\end{aligned}$$

Hence,

$$\text{cov}_S^T(\Phi_1, \Psi_1) = \frac{\text{supp}_S^T(\Phi_1, \Psi_1)}{\text{card}_F(\|\Psi_1\|_S^T)} = \frac{2.19}{4} \approx 0.55$$

The other indicators are calculated analogously and they are presented in [Table 11](#). $\square$

Table 11*T*-support, *T*-strength, *T*-certainty and *T*-coverage of the decision rules of Example 25.

Rule	$supp_S^T$	σ_S^T	cer_S^T	cov_S^T
r_1	2.19	0.31	0.58	0.55
r_2	2.11	0.3	0.45	0.7
r_3	2.14	0.31	0.6	0.54
r_4	1.61	0.23	0.58	0.54
r_5	2.58	0.37	0.58	0.65
r_6	2	0.29	0.52	0.67
r_7	2.29	0.33	0.6	0.57

Table 12

Relation between each pair of antecedents of decision rules of Example 25.

R_{Fd}	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
Φ_1	1	0.6	0.39	0.49	0.42	0.5	0.38
Φ_2	0.6	1	0.5	0.36	0.78	0.75	0.75
Φ_3	0.39	0.5	1	0.67	0.5	0.25	0.25
Φ_4	0.49	0.36	0.67	1	0.5	0.25	0.24
Φ_5	0.42	0.78	0.5	0.5	1	0.58	0.66
Φ_6	0.5	0.75	0.25	0.25	0.58	1	0.54
Φ_7	0.38	0.75	0.25	0.24	0.66	0.54	1

4. Decision algorithms in fuzzy rough set theory

This section is devoted to the study of decision algorithms in the fuzzy environment. First of all, it is necessary to compare formulas. These will be conjunction of attribute-value pairs because it is the usual form of the antecedent of decision rules. With this purpose, a fuzzy tolerance relation is introduced below.

Definition 30. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations and $\Phi, \Phi' \in For(\mathcal{A}_d)$, with $\Phi = (a_1, v_1) \wedge \dots \wedge (a_n, v_n)$ and $\Phi' = (a'_1, w_1) \wedge \dots \wedge (a'_m, w_m)$. The *F-indiscernibility relation* is a $[0, 1]$ -fuzzy tolerance relation $R_{Fd}: For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$ given as

$$R_{Fd}(\Phi, \Phi') = \begin{cases} \bigwedge_{i \in \{1, \dots, n\}} T_{a_i}(v_i, w_i) & \text{if } n = m \text{ and } a_i = a'_i \text{ for all } i \in \{1, \dots, n\} \\ 0 & \text{otherwise} \end{cases}$$

Given $\alpha \in [0, 1]$, the $R_{Fd} - \alpha$ -block is defined for each $\Phi \in For(\mathcal{A}_d)$ as:

$$[\Phi]_\alpha = \{\Phi' \in For(\mathcal{A}_d) \mid \alpha \leq R_{Fd}(\Phi, \Phi')\}$$

If $\Phi' \in [\Phi]_\alpha$ then we will say that Φ and Φ' are $R_{Fd} - \alpha$ -related.

It is important to emphasize that R_{Fd} is not a transitive fuzzy tolerance relation, in general, this is why we have tolerance blocks instead of equivalence classes. In addition, notice that given a pair of formulas Φ, Φ' and a fuzzy tolerance relation R_{Fd} , the value α determines if each formula belongs to the $R_{Fd} - \alpha$ -block of the other formula and vice versa.

In the following example, we will compare each pair of antecedents of decision rules of Example 25 by using a *F-indiscernibility relation* R_{Fd} .

Example 31. We will compute $R_{Fd}(\Phi, \Phi')$ for each Φ, Φ' antecedents of decision rules of Example 25. With this purpose, we fix the $[0, 1]$ -fuzzy tolerance relation $T_a: V_a \times V_a \rightarrow [0, 1]$ defined as $T_a(v, w) = 1 - |v - w|$ for all $a \in \mathcal{A}$. In particular, we illustrate the computation of $R_{Fd}(\Phi_1, \Phi_2)$. This is made by using Table 2 and Definition 30.

$$\begin{aligned} R_{Fd}(\Phi_1, \Phi_2) &= \bigwedge_{i \in \{1, 2, 3\}} T_{a_i}(v_i, w_i) \\ &= T_{a_1}(0.34, 0.21) \wedge T_{a_2}(0.31, 0.71) \wedge T_{a_3}(0.75, 0.5) \\ &= 0.87 \wedge 0.6 \wedge 0.75 = 0.6 \end{aligned}$$

The rest of results are shown in Table 12. $\square$

Before defining the notion of decision algorithm in the fuzzy framework it is necessary to introduce the following definition.

Definition 32. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\alpha \in (0, 1]$, $R_{Fd}: For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$ be a *F-indiscernibility relation* and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$

and $m \geq 2$. The set of α -consistent decision rules is defined as follows:

$$Dec_{\alpha}^{+}(S) = \{\Phi \rightarrow \Psi \in Dec(S) \mid \text{if for each } \Phi' \rightarrow \Psi' \in Dec(S) \text{ such} \\ \text{that } \Phi' \in [\Phi]_{\alpha} \text{ then } \Psi' \in [\Psi]_{\alpha}\}$$

Notice that, this set does not contain contradictory decision rules, that is, this set is composed of decision rules that, if their antecedents are $R_{Fd} - \alpha$ -related then their consequents are also $R_{Fd} - \alpha$ -related. Therefore, it is not possible to find decision rules in $Dec_{\alpha}^{+}(S)$ whose antecedents are $R_{Fd} - \alpha$ -related but their consequents not.

Finally, it is important to emphasize that we have not defined $Dec_{\alpha}^{+}(S)$ as the set of T -true decision rules because, by [Definition 26](#), given $\Phi \rightarrow \Psi \in Dec(S)$, we have that $cer_S^T(\Phi, \Psi) = 1$ if and only if $\|\Phi\|_S(x) \leq \|\Psi\|_S(x)$ for all $x \in U$. Hence, in order to avoid a restrictive condition we have adapted the notion of $Dec^{+}(S)$ to obtain interesting results.

Now, we present the notion of decision algorithm in the fuzzy setting.

Definition 33. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\alpha \in (0, 1]$, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1]$,¹ $\alpha_3 > 0$, $T = \{T_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable tolerance relations and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$. We say that:

1. $Dec(S)$ is a set of $\alpha_1\alpha_2$ -pairwise mutually exclusive (independent) decision rules, if each pair of decision rules $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisfies that $\Phi = \Phi'$ or $\|\Phi \wedge \Phi'\|_S^T(x) \leq \alpha_1$ and $\Psi = \Psi'$ or $\|\Psi \wedge \Psi'\|_S^T(x) \leq \alpha_2$, for all $x \in U$.
2. $Dec(S)$ covers U , if $card_F(\|\bigvee_{i=1}^m \Phi_i\|_S^T) = card_F(\|\bigvee_{i=1}^m \Psi_i\|_S^T) = |U|$.
3. The decision rule $\Phi \rightarrow \Psi \in Dec(S)$ is α_3 -admissible in S if $\alpha_3 < sup_S^T(\Phi, \Psi)$.
4. $Dec(S)$ preserves the α -consistency of S with a degree α_4 if the next inequality holds for all $x \in U$:

$$|\text{POS}_{\mathcal{A}}^f(x) - \bigvee_{\Phi \rightarrow \Psi \in Dec_{\alpha}^{+}(S)} \|\Phi\|_S^T(x)| \leq \alpha_4$$

The set of decision rules $Dec(S)$ satisfying the previous properties for the values $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ is called $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm in S and it is denoted as $DA_T(S)$.

In the previous definition only α affects α_4 , however, since the value α is a 'threshold' that the user must fix from the beginning, it should be known before defining the decision algorithm, and so it appears as subindex of the tuple of thresholds $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$.

It is important to emphasize that the set $Dec(S)$ in which each decision rule is extracted from each object of the decision table S is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm, taking appropriate values of each α_i , with $i \in \{1, 2, 3, 4\}$ as we will show in [Example 38](#). Each value has a different role. The value α_1 sets a bound to discern each pair of antecedents of the decision rules of $Dec(S)$, except the cases in which they are equals. The interpretation of α_2 is the same considering the consequents instead of the antecedents. Moreover, α_3 ensures a minimum value of satisfaction for all decision rules in $Dec(S)$ taking into account all the objects, which provides a meaningful description of the set of decision rules. Finally, α_4 is established to compare two important notions involved in the study of dependence between attributes, the fuzzy positive region and the decision rules of Dec_{α}^{+} .

In order to obtain a relationship between the sets $Dec_{\alpha}^{+}(S)$ and $Dec^{+}(S)$, we introduce the following notion.

Definition 34. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $DA_T(S)$ be a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm. If for each $x \in U$ there exists $\Phi \rightarrow \Psi \in DA_T(S)$ such that $\|\Phi \wedge \Psi\|_S^T(x) = 1$ then $DA_T(S)$ is called a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm.

Considering complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithms, it is possible to relate $Dec_{\alpha}^{+}(S)$ of [Definition 33](#) and the set $Dec^{+}(S)$ of [Definition 7](#), as the following result reveals.

Proposition 35. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $T = \{T_a : V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ be a family of boolean separable tolerance relations. If $\alpha > 0$ and $DA_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm then

$$Dec_{\alpha}^{+}(S) = Dec^{+}(S)$$

Proof. To begin with, notice that if each T_a tolerance relation of the family T is boolean, then R_{Fd} is also a boolean tolerance relation. In addition, since $\alpha > 0$ we obtain that

$$\begin{aligned} [\Phi]_{\alpha} &= \{\Phi' \mid \alpha \leq R_{Fd}(\Phi, \Phi'), \Phi' \rightarrow \Psi' \in DA_T(S)\} \\ &= \{\Phi' \mid R_{Fd}(\Phi, \Phi') = 1, \Phi' \rightarrow \Psi' \in DA_T(S)\} \\ &= \{\Phi' \mid \Phi' = \Phi, \Phi' \rightarrow \Psi' \in DA_T(S)\} \end{aligned}$$

¹ The value 1 is not taken into account since these cases are always satisfied and they are valueless.

Hence, $\Phi' \in [\Phi]_\alpha$ if and only if $\Phi = \Phi'$. Therefore, under these hypotheses we obtain that

$$\text{Dec}_\alpha^+(S) = \{\Phi \rightarrow \Psi \in DA_T(S) \mid \text{if for each } \Phi' \rightarrow \Psi' \in DA_T(S) \text{ such} \\ \text{that } \Phi' = \Phi \text{ then } \Psi' = \Psi\}$$

First of all, we prove that $\text{Dec}_\alpha^+(S) \subseteq \text{Dec}^+(S)$. Let $\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)$. By reductio ad absurdum, we suppose that $\Phi \rightarrow \Psi \notin \text{Dec}^+(S)$. Then $\text{cer}_S^T(\Phi, \Psi) < 1$. Since T is a family of boolean separable tolerance relations, by Proposition 28, we obtain $\text{cer}_S^T(\Phi, \Psi) = \text{cer}_S(\Phi, \Psi)$. Therefore,

$$\frac{\sum_{x \in U} \|\Phi \wedge \Psi\|_S^T(x)}{\sum_{x \in U} \|\Phi\|_S^T(x)} = \frac{\text{supp}_S^T(\Phi, \Psi)}{\text{card}_F(\|\Phi\|_S^T)} = \text{cer}_S^T(\Phi, \Psi) = \text{cer}_S(\Phi, \Psi) < 1$$

As a consequence, there exists $x \in U$ such that $\|\Phi \wedge \Psi\|_S^T(x) < \|\Phi\|_S^T(x)$. Taking into account that T is a family of boolean separable tolerance relations, $\|\Phi\|_S^T(x) = 1$ and $\|\Psi\|_S^T(x) = 0$.

On the other hand, since $DA_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm, there exists $\Phi' \rightarrow \Psi' \in DA_T(S)$ such that $\|\Phi' \wedge \Psi'\|_S^T(x) = 1$. Hence, $\|\Phi\|_S^T(x) = \|\Phi'\|_S^T(x) = \|\Phi \wedge \Phi'\|_S^T(x) = 1$. As a result, taking into account that $\alpha_1 < 1$ it is deduced that $\alpha_1 < \|\Phi \wedge \Phi'\|_S^T(x)$, so $\Phi = \Phi'$. In addition, $\|\Psi'\|_S^T(x) = 1$ and $\|\Psi \wedge \Psi'\|_S^T(x) = 0$. Therefore, it must be $\Psi \neq \Psi'$. This fact leads to a contradiction, since $\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)$, so $\Phi \rightarrow \Psi \in \text{Dec}^+(S)$. Therefore, we can conclude that $\text{Dec}_\alpha^+(S) \subseteq \text{Dec}^+(S)$.

Now, we prove that $\text{Dec}^+(S) \subseteq \text{Dec}_\alpha^+(S)$. Let $\Phi \rightarrow \Psi \in \text{Dec}^+(S)$. Then $\text{cer}_S(\Phi, \Psi) = 1$ and by Proposition 28, $\text{cer}_S^T(\Phi, \Psi) = 1$. By reductio ad absurdum, we suppose that $\Phi \rightarrow \Psi \notin \text{Dec}_\alpha^+(S)$. Then, there exists $\Phi \rightarrow \Psi' \in DA_T(S)$ such that $\Psi \neq \Psi'$. Since $DA_T(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm, it is obtained that $\text{supp}_S^T(\Phi, \Psi), \text{supp}_S^T(\Phi, \Psi') > \alpha_3 \geq 0$. In particular, since each mapping of T is boolean and separable, $\text{supp}_S^T(\Phi, \Psi), \text{supp}_S^T(\Phi, \Psi') \geq 1$. Therefore, there exists $x \in U$ such that $\|\Phi \wedge \Psi'\|_S^T(x) = 1$. Taking into account that $\Psi \neq \Psi'$, it is concluded that $\|\Phi \wedge \Psi\|_S^T(x) = 0$. Hence,

$$\text{cer}_S^T(\Phi, \Psi) = \frac{\text{supp}_S^T(\Phi, \Psi)}{\text{card}_F(\|\Phi\|_S^T)} = \frac{\sum_{x \in U} \|\Phi \wedge \Psi\|_S^T(x)}{\sum_{x \in U} \|\Phi\|_S^T(x)} < 1$$

leading to a contradiction. As a consequence, $\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)$, and $\text{Dec}_\alpha^+(S) = \text{Dec}^+(S)$. $\square$

From this result, we can relate the set $\text{Dec}_\alpha^+(S)$ to the set of T -true decision rules as it is shown below.

Corollary 36. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $T = \{T_a: V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ be a family of boolean separable tolerance relations. If $\alpha > 0$ and $DA_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm then

$$\text{Dec}_\alpha^+(S) = \{\Phi \rightarrow \Psi \in DA_T(S) \mid \text{cer}_S^T(\Phi, \Psi) = 1\}$$

Proof. It is immediately obtained, by using Propositions 28 and 35

$$\begin{aligned} \text{Dec}_\alpha^+(S) &= \text{Dec}^+(S) \\ &= \{\Phi \rightarrow \Psi \in DA_T(S) \mid \text{cer}_S(\Phi, \Psi) = 1\} \\ &= \{\Phi \rightarrow \Psi \in DA_T(S) \mid \text{cer}_S^T(\Phi, \Psi) = 1\} \quad \square \end{aligned}$$

Hence, when boolean and separable relations are considered, the set $\text{Dec}_\alpha^+(S)$ coincides with the set of T -true decision rules.

Finally, we show that complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithms generalize complete decision algorithms.

Proposition 37. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $T = \{T_a: V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ be a family of boolean separable tolerance relations and R_A and R_a be boolean separable tolerance relations for all $a \in \mathcal{A}_d$. Given $DA_T(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$, where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$, $\alpha \in (0, 1]$, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1)$ and $0 \leq \alpha_3 < \min\{\text{supp}_S^T(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in DA_T(S)\}$, then $DA_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm if and only if it is a complete decision algorithm.

Proof. Suppose that $DA_T(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm. We will prove the equivalence between each item of Definitions 7 and 33.

1. Let $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in DA_T(S)$ with $\Phi \neq \Phi'$. If $\|\Phi \wedge \Phi'\|_S^T(x) \leq \alpha_1 < 1$ for all $x \in U$, taking into account that T is a family of boolean separable tolerance relations, it will be $\|\Phi \wedge \Phi'\|_S^T(x) = 0$ for all $x \in U$. Then, by Proposition 24, $x \notin \|\Phi \wedge \Phi'\|_S$ for all $x \in U$. As a consequence, $\|\Phi \wedge \Phi'\|_S = \emptyset$. Analogously, if $\|\Psi \wedge \Psi'\|_S^T(x) \leq \alpha_2 < 1$ for all $x \in U$ then $\|\Psi \wedge \Psi'\|_S = \emptyset$.

2. We suppose that $\text{card}_F(\|\bigvee_{i=1}^m \Phi_i\|_S^T) = |U|$. By the definition of cardinal of a fuzzy set and the disjunction of formulas we obtain

$$\text{card}_F(\|\bigvee_{i=1}^m \Phi_i\|_S^T) = \sum_{x \in U} \|\bigvee_{i=1}^m \Phi_i\|_S^T(x) = \sum_{x \in U} \sum_{i=1}^m \|\Phi_i\|_S^T(x) = |U|$$

Since $0 \leq \|\bigvee_{i=1}^m \Phi_i\|_S^T(x) \leq 1$ and that addition has $|U|$ addends, it is obtained that $\|\bigvee_{i=1}^m \Phi_i\|_S^T(x) = 1$, for all $x \in U$. Therefore, for each $x \in U$, there exists $k \in \{1, 2, \dots, m\}$ such that $\|\Phi_k\|_S^T(x) = 1$. Hence, by [Proposition 24](#), $x \in \|\Phi_k\|_S \subseteq \|\bigvee_{i=1}^m \Phi_i\|_S$. Consequently, $U \subseteq \|\bigvee_{i=1}^m \Phi_i\|_S$. The other inclusion is immediately obtained because of

Definition 4. It can be proved analogously that $\text{card}_F(\|\bigvee_{i=1}^m \Psi_i\|_S^T) = |U|$ implies that $\|\bigvee_{i=1}^m \Psi_i\|_S = U$.

3. The third item is obtained immediately since, by [Proposition 28](#), $\text{supp}_S^T(\Phi, \Psi) = \text{supp}_S(\Phi, \Psi)$ for each decision rule $\Phi \rightarrow \Psi$ when T is a family of boolean separable tolerance relations.
4. For the fourth item notice that $\text{Dec}_\alpha^+(S) = \text{Dec}^+(S)$ by [Proposition 35](#). Since $\alpha_4 < 1$ and $\text{POS}_\mathcal{A}^f(x)$, $\bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)} \|\Phi\|_S^T(x) \in \{0, 1\}$ for all $x \in U$, we obtain that

$$\text{POS}_\mathcal{A}^f(x) = \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)} \|\Phi\|_S^T(x)$$

for all $x \in U$. Now, we will prove that $\text{POS}_\mathcal{A}(\{d\}) = \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}^+(S)} \|\Phi\|_S$. Let $x \in \text{POS}_\mathcal{A}(\{d\})$. By [Proposition 18](#),

$$\text{POS}_\mathcal{A}^f(x) = 1. \text{ Hence, } \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)} \|\Phi\|_S^T(x) = 1. \text{ As a consequence, } x \in \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}^+(S)} \|\Phi\|_S \text{ and } \text{POS}_\mathcal{A}(\{d\}) \subseteq \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}^+(S)} \|\Phi\|_S.$$

Suppose now that $x \in \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}^+(S)} \|\Phi\|_S$. Then, $\bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_\alpha^+(S)} \|\Phi\|_S^T(x) = 1$. As a consequence, $\text{POS}_\mathcal{A}^f(x) = 1$. Therefore, by [Proposition 18](#) $x \in \text{POS}_\mathcal{A}(\{d\})$ and $\bigvee_{\Phi \rightarrow \Psi \in \text{Dec}^+(S)} \|\Phi\|_S \subseteq \text{POS}_\mathcal{A}(\{d\})$.

Finally, we suppose that $DA_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm. Then, for each $x \in U$ there exists $\Phi \rightarrow \Psi \in DA_T(S)$ such that $\|\Phi \wedge \Psi\|_S^T(x) = 1$. Taking into account that T is a family of boolean separable tolerance relations and applying [Proposition 24](#), we obtain that $x \in \|\Phi \wedge \Psi\|_S$. As a consequence, $DA_T(S)$ is a complete decision algorithm.

The counterpart can be analogously proved, taking into account the requirement of the value α_3 given in the hypothesis. $\square$

This connection between complete decision algorithms in the classical and fuzzy frameworks gives more importance to [Definition 33](#), since it provides more flexibility than [Definition 7](#), thanks to the family of tolerance relations T and the values $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

The following example is introduced to illustrate the notion of $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm.

Example 38. Now, we will calculate maximum/minimum values of α_k with $k \in \{1, 2, 3, 4\}$ to obtain a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm composed of the decision rules $\text{Dec}(S) = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$ of [Example 25](#). Firstly, we compute α_1 by using [Table 10](#) and [Definition 23](#). From the first item of [Definition 33](#), we compute the value

$$\begin{aligned} \max\{\|\Phi \wedge \Phi'\|_S^T(x) \mid \Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in \text{Dec}(S), \Phi \neq \Phi', x \in U\} &= \|\Phi_2 \wedge \Phi_5\|_S^T(x_2) \\ &= 1 \wedge 0.78 \\ &= 0.78 \end{aligned}$$

Consequently, any value $\alpha_1 \geq 0.78$ verifies that $\|\Phi \wedge \Phi'\|_S^T(x) \leq \alpha_1$ for each Φ, Φ' antecedents of decision rules of $\text{Dec}(S)$. The condition $\alpha_2 \geq 0$ clearly arises because d is a boolean attribute. Hence, $\text{Dec}(S)$ is a set of $\alpha_1\alpha_2$ -pairwise mutually exclusive decision rules.

On the other hand, notice that each decision rule of $\text{Dec}(S)$ satisfies that $\|\Phi_i \wedge \Psi_i\|_S^T(x_i) = \|\Phi_i\|_S^T(x_i) \wedge \|\Psi_i\|_S^T(x_i) = 1$, that is, $\|\Phi_i\|_S^T(x_i) = \|\Psi_i\|_S^T(x_i) = 1$ for all $i \in \{1, 2, \dots, 7\}$. As a consequence,

$$\text{card}_F(\|\bigvee_{i=1}^7 \Phi_i\|_S^T) = \sum_{x \in U} \|\bigvee_{i=1}^7 \Phi_i\|_S^T(x) = \sum_{x \in U} \sum_{i=1}^7 \|\Phi_i\|_S^T(x) = \sum_{j=1}^{|U|} \|\Phi_j\|_S^T(x_j) = \sum_{j=1}^{|U|} 1 = |U|$$

Analogously, $\text{card}_F(\|\bigvee_{i=1}^m \Psi_i\|_S^T) = |U|$. Therefore, according to Definition 33, $\text{Dec}(S)$ covers U .

Now, for the computation of α_3 we take into account the third item of Definition 33, obtaining:

$$\min\{\text{supp}_S^T(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in \text{Dec}(S)\} = \text{supp}_S^T(\Phi_4, \Psi_4) = 1.61$$

Therefore, any value $\alpha_3 < 1.61$ satisfies that $\alpha_3 < \text{supp}_S^T(\Phi, \Psi)$ for all $\Phi \rightarrow \Psi \in \text{Dec}(S)$. Consequently, each decision rule of $\text{Dec}(S)$ is α_3 -admissible.

Finally, the value α_4 is computed. In Table 9, we showed the values of $\text{POS}_{\mathcal{A}}^f(x)$, for all $x \in U$. Now, we must obtain the set $\text{Dec}_{\alpha}^+(S)$ to compute $\bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_{\alpha}^+(S)} \|\Phi\|_S^T(x)$ for all $x \in U$. In order to consider a good level of flexibility and a value close to 1 at the same time, we choose the value $\alpha = 0.75$. Hence, according to Definition 33, it is necessary to calculate the $R_{Fd} - 0.75$ -block of each antecedent. By using Table 12, these blocks are

$$\begin{aligned} [\Phi_1]_{0.75} &= \{\Phi_1\} \\ [\Phi_2]_{0.75} &= \{\Phi_2, \Phi_5, \Phi_6, \Phi_7\} & [\Phi_4]_{0.75} &= \{\Phi_4\} & [\Phi_6]_{0.75} &= \{\Phi_2, \Phi_6\} \\ [\Phi_3]_{0.75} &= \{\Phi_3\} & [\Phi_5]_{0.75} &= \{\Phi_2, \Phi_5\} & [\Phi_7]_{0.75} &= \{\Phi_2, \Phi_7\} \end{aligned}$$

Analogously, the $R_{Fd} - 0.75$ -block of each consequent is given as

$$[\Psi_i]_{0.75} = \begin{cases} \{\Psi_1, \Psi_3, \Psi_5, \Psi_7\} & \text{if } i \text{ is odd} \\ \{\Psi_2, \Psi_4, \Psi_6\} & \text{if } i \text{ is even} \end{cases}$$

Consequently, from these blocks we obtain that $\text{Dec}_{0.75}^+(S) = \{r_1, r_3, r_4, r_6\}$ because $\Phi_5, \Phi_7 \in [\Phi_2]_{0.75}$ but $\Psi_5, \Psi_7 \notin [\Psi_2]_{0.75}$. For the fourth item of Definition 33 consider the value

$$\begin{aligned} &\max\{|\text{POS}_{\mathcal{A}}^f(x) - \bigvee_{\Phi \rightarrow \Psi \in \text{Dec}_{0.75}^+(S)} \|\Phi\|_S^T(x)| \mid x \in U\} \\ &= \max\{|0.96 - 1|, |0.86 - 0.75|, |0.91 - 1|, |0.91 - 1|, |0.86 - 0.58|, \\ &\quad |0.93 - 1|, |0.93 - 0.54|\} \\ &= 0.39 \end{aligned}$$

Hence, any value $\alpha_4 \geq 0.39$ makes $\text{Dec}(S)$ preserves the 0.75-consistency of S to degree α_4 .

Therefore, $\text{Dec}(S) = \text{DA}_T(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm. In fact, it is easy to see that $\text{DA}_T(S)$ is a complete $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{\alpha}$ -decision algorithm since $\|\Phi_i \wedge \Psi_i\|_S^T(x_i) = 1$ for all $i \in \{1, 2, \dots, 7\}$. $\square$

In conclusion, in this section we have presented the fuzzy notion of decision rule and all the rest notions related to it such as the relevance indicators which describe them and decision algorithms. It is important to emphasize that all these notions generalize the classical ones and they present interesting properties thanks to the flexibility the fuzzy framework offers. Next, we present an example with real information, to illustrate all of these definitions, showing in detail some calculus as a sample.

5. A toy example

This section applies the developed theory to analyze a real dataset. We include an example based on the knowledge system displayed in Table 13, whose dataset has been extracted from Statistical Yearbook of Zhejiang Province (2016) [20], in order to optimize of water conservancy project investment decision-making. It is important to mention that the knowledge system, obtained from the previously mentioned data, was already studied by using classical RST techniques in [20]. Now, we will carry out its study by using the notions and results obtained in the introduced fuzzy framework.

Specifically, the knowledge system collects eight project evaluations, based on the attributes construction expense (CE), financial income (FI), strategy benefit (SB) and external influence (EI), together with the corresponding decisions made by an investor, that is, to make an investment (I), a delayed investment (DI) or not to invest (NI). Every unit in Table 13 correspond to 10.000 yuan. With the goal of applying the proposed fuzzy RST technique, we transform the data given in Table 13 into the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ represented in Table 14, where the set of objects is $U = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\}$, the set of attributes is $\mathcal{A} = \{CE, FI, SB, EI\}$, $\mathcal{V}_a = [0, 1]$, for all $a \in \mathcal{A}$ and $\mathcal{V}_d = \{I, NI, DI\}$. Notice that, we have applied a normalization factor by columns so that the domain is the interval $[0, 1]$. As the values of the first three columns are expected significantly higher than the values of the last one, we have divided by 1000 in the first case and by 500 in the second one. We could have normalized from the maximum and minimum values in each column of the decision table. However, we have considered a bigger range in order to have the possibility of taking into consideration projects with amounts greater and lower than these maximums and minimums.

To begin with, we compute a fuzzy m -decision reduct to degree 1 in order to consider less attributes in the decision rules, without losing information. The corresponding $[0, 1]$ -valued measure $m: \mathcal{P}(\mathcal{A}) \rightarrow [0, 1]$ is the one given in Eq. (2)

Table 13

Knowledge system on water conservancy project investment decision-making.

	CE	FI	SB	EI	Decision (d)
Project 1 (p_1)	500	200	690	150	<i>DI</i>
Project 2 (p_2)	700	470	650	440	<i>NI</i>
Project 3 (p_3)	300	410	500	280	<i>NI</i>
Project 4 (p_4)	200	455	550	290	<i>I</i>
Project 5 (p_5)	250	260	490	105	<i>DI</i>
Project 6 (p_6)	510	380	430	130	<i>DI</i>
Project 7 (p_7)	350	550	255	145	<i>I</i>
Project 8 (p_8)	650	600	570	120	<i>NI</i>

Table 14Decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ extracted from Table 13.

	CE	FI	SB	EI	d
p_1	0.5	0.2	0.69	0.3	<i>DI</i>
p_2	0.7	0.47	0.65	0.88	<i>NI</i>
p_3	0.3	0.41	0.5	0.56	<i>NI</i>
p_4	0.2	0.455	0.55	0.58	<i>I</i>
p_5	0.25	0.26	0.49	0.21	<i>DI</i>
p_6	0.51	0.38	0.43	0.26	<i>DI</i>
p_7	0.35	0.55	0.255	0.29	<i>I</i>
p_8	0.65	0.6	0.57	0.24	<i>NI</i>

and the a -indiscernibility relation $R_a: U \times U \rightarrow [0, 1]$ is the one given in Eq. (4), for all $a \in \mathcal{A}$. For the d -indiscernibility relation R_d , we define the following mapping, for all $x, y \in U$:

$$R_d(x, y) = \begin{cases} 1 & \text{if } \bar{d}(x) = \bar{d}(y) \\ 0.5 & \text{if } \bar{d}(x) = DI \text{ and } \bar{d}(y) \neq DI \\ 0 & \text{otherwise} \end{cases}$$

It is important to recall that R_d is a symmetrical relation. On the other hand, since we are interested in all attributes have the same impact in the evaluation of construction projects, we will use the B -indiscernibility relation $R_B: U \times U \rightarrow [0, 1]$ defined, for each $B \subseteq \mathcal{A}$, $x, y \in U$, as:

$$R_B(x, y) = \frac{\mathcal{R}_B^{x,y}(CE) + \mathcal{R}_B^{x,y}(FI) + \mathcal{R}_B^{x,y}(SB) + \mathcal{R}_B^{x,y}(EI)}{4}$$

where $\mathcal{R}_B^{x,y}(a)$ was given in Definition 11, for all $a \in \mathcal{A}$. Finally, the multi-adjoint fuzzy $\mathcal{A}$ -positive region is computed from the fuzzy implication given in Eq. (5), considering the following symmetrical mapping $\tau: U \times U \rightarrow \{1, 3, 5\}$ defined, for each $x, y \in U$, as:

$$\tau(x, y) = \begin{cases} 1 & \text{if } \bar{d}(x) = \bar{d}(y) \\ 3 & \text{if } \bar{d}(x) = DI \text{ and } \bar{d}(y) \neq DI \\ 5 & \text{if } \bar{d}(x) = I \text{ and } \bar{d}(y) = NI \end{cases}$$

We would like to cushion the relatively big differences between the decision attribute values. For example, if we consider p_2 and p_7 , we obtain $R_d(p_2, p_7) = 0$, which has a great impact in the computation of the implication $R_d(p_2, p_7) \curvearrowright_{\tau(p_2, p_7)} R_A(p_2, p_7)$, and so in the positive region. Hence, the selection of the implication $\tau(p_2, p_7) = 5$ reduces the fact they have some similarity on $\mathcal{A}$ and gives a greater value to the implication.

Fig. 2 shows the computed multi-adjoint fuzzy $\mathcal{A}$ -positive region. The values of the multi-adjoint fuzzy $\mathcal{A}$ -positive region allow us to compute $m(B)$, for all $B \subseteq \mathcal{A}$. The most remarkable results are also presented in Fig. 2.

We have shown the $[0, 1]$ -valued measure m for all the subsets of three attributes and the three subsets of two attributes with the greatest $[0, 1]$ -valued measure. It is important to emphasize the relevance of the attribute CE in the decision table since, as we can see in Fig. 2, it is the only attribute which belongs to the two subsets with two attributes of greater $[0, 1]$ -valued measure. Therefore, when a project is evaluated to invest or not, it is fundamental to pay attention to the construction expense, which is also very natural. However, from these values we can conclude that $\mathcal{A}$ is the only fuzzy m -decision reduct to degree 1, so all its attributes are indispensable to process all the information contained in

	$\text{POS}_{\mathcal{A}}^f$	
p_1	0.83	$m(\{CE, FI, EI\}) = 0.95$
p_2	0.9	$m(\{FI, SB, EI\}) = 0.94$
p_3	0.75	$m(\{CE, SB, EI\}) = 0.94$
p_4	0.75	$m(\{CE, FI, SB\}) = 0.93$
p_5	0.79	$m(\{CE, FI\}) = 0.88$
p_6	0.78	$m(\{CE, EI\}) = 0.87$
p_7	0.76	$m(\{FI, FI\}) = 0.85$
p_8	0.78	

Fig. 2. Multi-adjoint fuzzy $\mathcal{A}$ -positive region.

Table 15

Degree of satisfaction for each antecedent in $\text{Dec}(S)$ for each object.

	$\ \Phi_1\ _S^T$	$\ \Phi_2\ _S^T$	$\ \Phi_3\ _S^T$	$\ \Phi_4\ _S^T$	$\ \Phi_5\ _S^T$	$\ \Phi_6\ _S^T$	$\ \Phi_7\ _S^T$	$\ \Phi_8\ _S^T$
p_1	1	0.42	0.74	0.7	0.75	0.74	0.565	0.6
p_2	0.42	1	0.6	0.5	0.33	0.38	0.41	0.36
p_3	0.74	0.6	1	0.9	0.65	0.7	0.73	0.65
p_4	0.7	0.5	0.9	1	0.63	0.68	0.705	0.55
p_5	0.75	0.33	0.65	0.63	1	0.74	0.71	0.6
p_6	0.74	0.38	0.7	0.68	0.74	1	0.825	0.78
p_7	0.565	0.41	0.73	0.705	0.71	0.825	1	0.685
p_8	0.6	0.36	0.65	0.55	0.6	0.78	0.685	1

Table 16

Degree of satisfaction for each consequent in $\text{Dec}(S)$ for each object.

	$\ \Psi_1\ _S^T$	$\ \Psi_2\ _S^T$	$\ \Psi_3\ _S^T$	$\ \Psi_4\ _S^T$	$\ \Psi_5\ _S^T$	$\ \Psi_6\ _S^T$	$\ \Psi_7\ _S^T$	$\ \Psi_8\ _S^T$
p_1	1	0.5	0.5	0.5	1	1	0.5	0.5
p_2	0.5	1	1	0	0.5	0.5	0	1
p_3	0.5	1	1	0	0.5	0.5	0	1
p_4	0.5	0	0	1	0.5	0.5	1	0
p_5	1	0.5	0.5	0.5	1	1	0.5	0.5
p_6	1	0.5	0.5	0.5	1	1	0.5	0.5
p_7	0.5	0	0	1	0.5	0.5	1	0
p_8	0.5	1	1	0	0.5	0.5	0	1

Table 14. As a consequence, to evaluate projects in this framework it is necessary to take into account the whole set of attributes to make a decision. We obtain the following decision rules:

- $$\begin{aligned}
r_1 &: (CE, 0.5) \wedge (FI, 0.2) \wedge (SB, 0.69) \wedge (EI, 0.3) \rightarrow (d, DI) \\
r_2 &: (CE, 0.7) \wedge (FI, 0.47) \wedge (SB, 0.65) \wedge (EI, 0.88) \rightarrow (d, NI) \\
r_3 &: (CE, 0.3) \wedge (FI, 0.41) \wedge (SB, 0.5) \wedge (EI, 0.56) \rightarrow (d, NI) \\
r_4 &: (CE, 0.2) \wedge (FI, 0.455) \wedge (SB, 0.55) \wedge (EI, 0.58) \rightarrow (d, I) \\
r_5 &: (CE, 0.25) \wedge (FI, 0.26) \wedge (SB, 0.49) \wedge (EI, 0.21) \rightarrow (d, DI) \\
r_6 &: (CE, 0.51) \wedge (FI, 0.38) \wedge (SB, 0.43) \wedge (EI, 0.26) \rightarrow (d, DI) \\
r_7 &: (CE, 0.35) \wedge (FI, 0.55) \wedge (SB, 0.255) \wedge (EI, 0.29) \rightarrow (d, I) \\
r_8 &: (CE, 0.65) \wedge (FI, 0.6) \wedge (SB, 0.57) \wedge (EI, 0.24) \rightarrow (d, NI)
\end{aligned}$$

We denote this set of decision rules as $\text{Dec}(S)$. The degree of satisfaction for each antecedent and consequent, for each object, is computed by using the family of separable $[0, 1]$ -fuzzy tolerance relations $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ defined as $T_a(\bar{a}(x), v) = 1 - |\bar{a}(x) - v|$, for all $a \in \mathcal{A}$, $T_d = R_d$, $x \in U$ and $v \in [0, 1]$. All the results are shown in Tables 15 and 16.

Now, we compute the T -support, T -strength, T -certainty and T -coverage of each decision rule, which are values obtained from new measures presented in this paper, that could not be taken into account in the framework of [20], and that offer more information about the set of decision rules. The results are presented in Table 17.

From Table 17 we can extract some conclusions. First of all, we can detect that the rules with largest T -support are r_1 , r_5 and r_6 , whose decision attribute takes the value DI . This fact arises because of the three rules have DI as decision attribute value and DI has a level of similarity with the other decision attribute values. Moreover, these rules have also the largest T -certainty, showing that the decision DI is the best one to recommend in case of doubt about the investment.

Table 17*T*-support, *T*-strength, *T*-certainty and *T*-coverage of each rule in *Dec*(*S*).

Rule	$supp_S^T$	σ_S^T	cer_S^T	cov_S^T
r_1	4.91	0.61	0.89	0.89
r_2	3.09	0.39	0.77	0.69
r_3	3.75	0.47	0.63	0.83
r_4	3.21	0.4	0.57	0.92
r_5	4.82	0.6	0.89	0.88
r_6	4.86	0.61	0.83	0.88
r_7	3.21	0.4	0.57	0.92
r_8	3.51	0.44	0.67	0.78

Table 18Relation between each pair of antecedents of *Dec*(*S*).

R_{Fd}	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7	Φ_8
Φ_1	1	0.42	0.74	0.7	0.75	0.74	0.565	0.6
Φ_2	0.42	1	0.6	0.5	0.33	0.38	0.41	0.36
Φ_3	0.74	0.6	1	0.9	0.65	0.7	0.73	0.65
Φ_4	0.7	0.5	0.9	1	0.63	0.68	0.705	0.55
Φ_5	0.75	0.33	0.65	0.63	1	0.74	0.71	0.6
Φ_6	0.74	0.38	0.7	0.68	0.74	1	0.825	0.78
Φ_7	0.565	0.41	0.73	0.705	0.71	0.825	1	0.685
Φ_8	0.6	0.36	0.65	0.55	0.6	0.78	0.685	1

Table 19Relation between each pair of consequents of *Dec*(*S*).

R_{Fd}	Ψ_1	Ψ_2	Ψ_3	Ψ_4	Ψ_5	Ψ_6	Ψ_7	Ψ_8
Ψ_1	1	0.5	0.5	0.5	1	1	0.5	0.5
Ψ_2	0.5	1	1	0	0.5	0.5	0	1
Ψ_3	0.5	1	1	0	0.5	0.5	0	1
Ψ_4	0.5	0	0	1	0.5	0.5	1	0
Ψ_5	1	0.5	0.5	0.5	1	1	0.5	0.5
Ψ_6	1	0.5	0.5	0.5	1	1	0.5	0.5
Ψ_7	0.5	0	0	1	0.5	0.5	1	0
Ψ_8	0.5	1	1	0	0.5	0.5	0	1

On the other hand, the rules r_4 and r_7 , which correspond to the decision *I*, have the lowest values in the *T*-certainty. This is due to the fact that no project presents properties that make it a safe investment, since the objects associated with these rules are similar to others with a different decision. However, these rules have the greatest *T*-coverage because, in case that having to invest in some projects, they are the best candidates according to their condition attributes. Finally, the decision rules with lowest *T*-coverage are r_2 , r_3 and r_8 , in which their decision attribute is *NI*. This fact shows that there are reasons not to invest in the projects related to the rules r_2 , r_3 and r_8 .

Now, we are going to prove whether the considered set of decision rules forms a decision algorithm (Definition 33) and, in affirmative case, the associated thresholds $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ will be analyzed, if the tolerance value is fixed at 0.75. The relationship between each pair of antecedents and consequents is computed based on Definition 30. The obtained results are shown in Tables 18 and 19.

As in Example 38, we have set the tolerance value $\alpha = 0.75$ to compute the $R_{Fd} - \alpha$ -blocks of the antecedents and consequents. By using Tables 15 and 16, it can be checked that these blocks are the following:

$$\begin{aligned} [\Phi_1]_{0.75} &= \{\Phi_1, \Phi_5\} & [\Phi_4]_{0.75} &= \{\Phi_3, \Phi_4\} & [\Phi_7]_{0.75} &= \{\Phi_6, \Phi_7\} \\ [\Phi_2]_{0.75} &= \{\Phi_2\} & [\Phi_5]_{0.75} &= \{\Phi_1, \Phi_5\} & [\Phi_8]_{0.75} &= \{\Phi_6, \Phi_8\} \\ [\Phi_3]_{0.75} &= \{\Phi_3, \Phi_4\} & [\Phi_6]_{0.75} &= \{\Phi_6, \Phi_7, \Phi_8\} \end{aligned}$$

$$\begin{aligned} [\Psi_1]_{0.75} &= [\Psi_5]_{0.75} = [\Psi_6]_{0.75} = \{\Psi_1, \Psi_5, \Psi_6\} \\ [\Psi_2]_{0.75} &= [\Psi_3]_{0.75} = [\Psi_8]_{0.75} = \{\Psi_2, \Psi_3, \Psi_8\} \\ [\Psi_4]_{0.75} &= [\Psi_7]_{0.75} = \{\Psi_4, \Psi_7\} \end{aligned}$$

Consequently, by Definition 32, we deduce that $Dec_{0.75}^+(S) = \{r_1, r_2, r_5\}$. Following an analogous procedure to the one given in Example 38, we obtain that *Dec*(*S*) is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{0.75}$ -decision algorithm, with $\alpha_1 \geq 0.9$, $\alpha_2 \geq 0.5$, $\alpha_3 < 3.09$ and $\alpha_4 \geq 0.21$. From the bound α_1 we deduce that it is difficult to discern objects of the table, which shows the complexity in the decision-making process for these evaluation projects. In fact, by Table 18 we know this value has been obtained from $\|\Phi_3 \wedge \Phi_4\|_S^T(p_3) = 0.9$, where $\Psi_3 = (d, NI)$ and $\Psi_4 = (d, I)$, which are totally contrary decisions. Therefore, this dataset presents inconsistencies, which were not detected in [20].

Table 20Degree of satisfaction for each antecedent in $Dec(S')$ for each object.

	$\ \Phi'_1\ _S^T$	$\ \Phi'_2\ _S^T$	$\ \Phi'_3\ _S^T$	$\ \Phi'_4\ _S^T$	$\ \Phi'_5\ _S^T$	$\ \Phi'_6\ _S^T$	$\ \Phi'_7\ _S^T$	$\ \Phi'_8\ _S^T$
p_1	1	0.42	0.74	0.7	0.75	0.82	0.65	0.6
p_2	0.42	1	0.6	0.5	0.33	0.38	0.41	0.36
p_3	0.74	0.6	1	0.9	0.65	0.7	0.73	0.65
p_4	0.7	0.5	0.9	1	0.63	0.68	0.71	0.55
p_5	0.75	0.33	0.65	0.63	1	0.74	0.71	0.6
p_6	0.82	0.38	0.7	0.68	0.74	1	0.83	0.78
p_7	0.65	0.41	0.73	0.71	0.71	0.83	1	0.7
p_8	0.6	0.36	0.65	0.55	0.6	0.78	0.7	1

Table 21 T -support, T -strength, T -certainty and T -coverage of each rule in $Dec(S')$.

Rule	$supp_S^T$	σ_S^T	cer_S^T	cov_S^T
r'_1	4.99	0.62	0.88	0.91
r'_2	3.09	0.39	0.77	0.69
r'_3	3.75	0.47	0.63	0.83
r'_4	3.21	0.4	0.57	0.92
r'_5	4.82	0.6	0.89	0.88
r'_6	4.94	0.62	0.83	0.9
r'_7	3.21	0.4	0.56	0.92
r'_8	3.51	0.44	0.67	0.78

On the other hand, due to the decision attribute is not fuzzy, it is easy to discern objects with different decisions, as the bound α_2 shows. With respect to the bound α_3 , it illustrates that all the rules have a high level of T -support, taking into account the reduced number of objects. This aspect is in line with the similarity in the values of the attributes. Notice that the T -support will be reduced if the normalization is done through the maximum and the minimum values in the decision table. However, as we previously commented, we have considered the range from 0 to 1.000 in order to have the possibility of taking into consideration projects with amounts lower than 2 millions of yuan and greater than 7 millions of yuan. Finally, α_4 exposes that the consistency of the dataset, provided by the fuzzy positive region, and the consistency of the considered set of decision rules in the algorithm is similar, taking into account that the tolerance value has been set in $\alpha = 0.75$. As a consequence, the chosen decision algorithm is suitable to study the corresponding dataset.

Notice that, if we avoid the inclusion of the decision rule associated with the object p_4 (a similar study can be done with respect to object p_3), then the inconsistency decreases. Specifically, the new obtained thresholds are: $\alpha_1 \geq 0.83$, $\alpha_2 \geq 0.5$, $\alpha_3 < 3.09$ and $\alpha_4 \geq 0.29$. However, the new set of rules does not form an algorithm, because of no formula Φ exists such that $\|\Phi\|_S^T(p_4) = 1$, and so the new set of rules does not cover U .

Although we remarked previously that only the set $\mathcal{A}$ is a fuzzy m -decision reduct to degree 1, we also showed the importance of the set of attributes $B = \{CE, FI, EI\}$. In fact, it can be checked that it is a fuzzy m -decision reduct to degree α with $0.94 < \alpha \leq 0.98$. Therefore, we will also study the set of decision rules in which the antecedents are the conjunction of the attributes of B , comparing the obtained results with those of the previous rules. The new set of rules is as follows:

$$\begin{aligned}
r'_1 &: (CE, 0.5) \wedge (FI, 0.2) \wedge (EI, 0.3) \rightarrow (d, DI) \\
r'_2 &: (CE, 0.7) \wedge (FI, 0.47) \wedge (EI, 0.88) \rightarrow (d, NI) \\
r'_3 &: (CE, 0.3) \wedge (FI, 0.41) \wedge (EI, 0.56) \rightarrow (d, NI) \\
r'_4 &: (CE, 0.2) \wedge (FI, 0.455) \wedge (EI, 0.58) \rightarrow (d, I) \\
r'_5 &: (CE, 0.25) \wedge (FI, 0.26) \wedge (EI, 0.21) \rightarrow (d, DI) \\
r'_6 &: (CE, 0.51) \wedge (FI, 0.38) \wedge (EI, 0.26) \rightarrow (d, DI) \\
r'_7 &: (CE, 0.35) \wedge (FI, 0.55) \wedge (EI, 0.29) \rightarrow (d, I) \\
r'_8 &: (CE, 0.65) \wedge (FI, 0.6) \wedge (EI, 0.24) \rightarrow (d, NI)
\end{aligned}$$

This set of decision rules will be denoted as $Dec(S')$, where S' is the decision table obtained from S removing the attribute SB . Considering the same family of separable $[0, 1]$ -fuzzy tolerance relations T that in the previous study, all the degrees of satisfaction for each antecedent in $Dec(S')$ and object are shown in Table 20.

Notice that, the degrees of satisfaction for the consequents do not change due to the same decision attribute is considered. Table 21 shows the T -support, T -strength, T -certainty and T -coverage of the decision rules of $Dec(S')$.

Comparing Tables 17 and 21 we notice very similar values, emphasizing the increment of the T -support of some rules, given that we have considered three attributes for the antecedents of the rules instead of the four, increasing the degrees of satisfaction for the antecedents.

Analogously, we compute the relationship between each pair of antecedents of the rules of $Dec(S')$. The results are shown in Table 22.

Table 22
Relation between each pair of antecedents of decision rules of $Dec(S')$.

R_{Fd}	Φ'_1	Φ'_2	Φ'_3	Φ'_4	Φ'_5	Φ'_6	Φ'_7	Φ'_8
Φ'_1	1	0.42	0.74	0.7	0.75	0.82	0.65	0.6
Φ'_2	0.42	1	0.6	0.5	0.33	0.38	0.41	0.36
Φ'_3	0.74	0.6	1	0.9	0.65	0.7	0.73	0.65
Φ'_4	0.7	0.5	0.9	1	0.63	0.68	0.71	0.55
Φ'_5	0.75	0.33	0.65	0.63	1	0.74	0.71	0.6
Φ'_6	0.82	0.38	0.7	0.68	0.74	1	0.83	0.78
Φ'_7	0.65	0.41	0.73	0.71	0.71	0.83	1	0.7
Φ'_8	0.6	0.36	0.65	0.55	0.6	0.78	0.7	1

Once again, we consider the value $\alpha = 0.75$ to compute the $R_{Fd} - \alpha$ -blocks of the antecedents, which are the following:

$$\begin{aligned}
 [\Phi'_1]_{0.75} &= \{\Phi'_1, \Phi'_5, \Phi'_6\} & [\Phi'_4]_{0.75} &= \{\Phi'_3, \Phi'_4\} & [\Phi'_7]_{0.75} &= \{\Phi'_6, \Phi'_7\} \\
 [\Phi'_2]_{0.75} &= \{\Phi'_2\} & [\Phi'_5]_{0.75} &= \{\Phi'_1, \Phi'_5\} & [\Phi'_8]_{0.75} &= \{\Phi'_6, \Phi'_8\} \\
 [\Phi'_3]_{0.75} &= \{\Phi'_3, \Phi'_4\} & [\Phi'_6]_{0.75} &= \{\Phi'_1, \Phi'_6, \Phi'_7, \Phi'_8\}
 \end{aligned}$$

It can be checked that $Dec_{0.75}^+(S') = \{r'_1, r'_2, r'_5\}$ and that $Dec(S')$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{0.75}$ -decision algorithm with $\alpha_1 \geq 0.9$, $\alpha_2 \geq 0.5$, $\alpha_3 < 3.09$ and $\alpha_4 \geq 0.21$. These bounds are the same that in the algorithm $Dec(S)$. This fact shows that we do not lose relatively much information considering in the antecedents of the rules the attributes $\{CE, FI, EI\}$ instead of $\{CE, FI, SB, EI\}$. As a consequence, it is possible not to take into account the attribute SB to decide whether to invest or not in these kinds of projects, taking into account the tolerance value $\alpha = 0.75$, which can be increased to reduce this tolerance. Moreover, recall that $\{CE, FI, EI\}$ is not a fuzzy m -decision reduct to degree 1, so that from this point of view, some information can also be lost. Although this loss of information has not impact in the obtained decision algorithms. In the future, the relationships between the degree of the m -reducts, the tolerance value and the decision algorithms will be studied in depth.

6. Conclusions and future work

Decision rules and decision algorithms have been studied in the fuzzy environment in this paper. In particular, a generalization to the fuzzy framework of the classic notions related to decision rules, including the support, strength, certainty and coverage have been introduced. Furthermore, the fuzzy notion of decision algorithms have also been presented. This notion will allow us to study the efficiency of a decision algorithm in this framework, which will be developed in future works. In addition, some examples have been introduced in order to illustrate all these new definitions.

In the future, we are interested in studying a generalized notion of efficiency which can be used to compare different decision algorithms. In addition, the results shown in this work can be applied to the decision rules obtained from mixed contexts [27]. However, the obtained algorithms would not take into account the attribute dependency given in this kind of contexts. Therefore, the consideration of mixed contexts will also be studied in the future. Moreover, we will apply the obtained theoretical developments in real examples whose decision attribute is not boolean. In order to evaluate the usefulness of the new notion of decision algorithm we will explore diverse examples that contain a large number of decision rules. In addition, it will be compared and complemented with other notions in the literature frameworks, such as logic program in fuzzy logic programming [28,29] and attribute implicational system in formal concept analysis [30,31].

Data availability

No data was used for the research described in the article

References

- [1] Z. Pawlak, Information systems theoretical foundations, *Inf. Syst.* 6 (3) (1981) 205–218.
- [2] Z. Pawlak, Rough sets, *Int. J. Comput. Inf. Sci.* 11 (1982) 341–356.
- [3] M.J. Benítez-Caballero, J. Medina, Characterizing clarify and reduce method by means of left-sided formal concept analysis, *Stud. Comput. Intell.* X (2023) 1–6.
- [4] M.J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, Attribute reduction in rough set theory and formal concept analysis, in: L. Polkowski, Y. Yao, P. Artiemjew, D. Ciucci, D. Liu, D. Ślęzak, B. Zielosko (Eds.), *Rough Sets*, Springer International Publishing, Cham, 2017, pp. 513–525.
- [5] M.J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, D. Ślęzak, Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis, *Fuzzy Sets and Systems* (2019).
- [6] J. Medina, Relating attribute reduction in formal, object-oriented and property-oriented concept lattices, *Comput. Math. Appl.* 64 (6) (2012) 1992–2002.
- [7] M. Qi, Y. Gu, Q. Wang, Internet financial risk management and control based on improved rough set algorithm, *J. Comput. Appl. Math.* 384 (2021) 113179.

- [8] R.G. Aragón, J. Medina, E. Ramírez-Poussa, Impact of local congruences in variable selection from datasets, *J. Comput. Appl. Math.* (2022) 113416, <http://dx.doi.org/10.1016/j.cam.2021.113416>.
- [9] Z. Xiao, K. Gong, Y. Zou, A combined forecasting approach based on fuzzy soft sets, *J. Comput. Appl. Math.* 228 (1) (2009) 326–333.
- [10] Q.-E. Wu, T. Wang, X.-M. Pang, Y.-X. Huang, J.-S. Li, Research on two different mathematical theories on control, *J. Comput. Appl. Math.* 222 (2) (2008) 675–689.
- [11] H. Bai, Y. Ge, J. Wang, D. Li, Y. Liao, X. Zheng, A method for extracting rules from spatial data based on rough fuzzy sets, *Knowl.-Based Syst.* 57 (2014) 28–40.
- [12] J. Li, C. Mei, Y. Lv, Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction, *Internat. J. Approx. Reason.* 54 (1) (2013) 149–165.
- [13] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning About Data*, Kluwer Academic Publishers, Norwell, MA, USA, 1992.
- [14] Z. Pawlak, Rough sets and decision algorithms, in: W. Ziarko, Y. Yao (Eds.), *Rough Sets and Current Trends in Computing*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2001, pp. 30–45.
- [15] S. Stawicki, D. Ślęzak, A. Janusz, S. Widz, Decision bireducts and decision reducts - a comparison, *Internat. J. Approx. Reason.* 84 (2017) 75–109.
- [16] Y. Yao, Three-way decisions with probabilistic rough sets, *Inform. Sci.* 180 (3) (2010) 341–353.
- [17] C. Cornelis, J. Medina, N. Verbiest, Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection, *Internat. J. Approx. Reason.* 55 (2014) 412–426.
- [18] J. Medina, Towards Multi-Adjoint Property-Oriented Concept Lattices, in: *Lecture Notes in Artificial Intelligence*, vol. 6401, 2010, pp. 159–166.
- [19] J. Medina, Multi-adjoint property-oriented and object-oriented concept lattices, *Inform. Sci.* 190 (2012) 95–106.
- [20] J. Qu, X. Bai, J. Gu, F. Taghizadeh-Hesary, J. Lin, Assessment of rough set theory in relation to risks regarding hydraulic engineering investment decisions, *Mathematics* 8 (8) 2020.
- [21] X. Zhang, J. Li, W. Li, A new mechanism of rule acquisition based on covering rough sets, *Appl. Intell.* 52 (2022) 12369–12381.
- [22] Z. Meng, Z. Shi, On rule acquisition methods for data classification in heterogeneous incomplete decision systems, *Knowl.-Based Syst.* 193 (2020) 105472.
- [23] M.E. Cornejo, J. Medina, E. Ramírez-Poussa, Algebraic structure and characterization of adjoint triples, *Fuzzy Sets and Systems* 425 (2021) 117–139, *Mathematics*.
- [24] M.E. Cornejo, J. Medina, E. Ramírez-Poussa, Multi-adjoint algebras versus non-commutative residuated structures, *Internat. J. Approx. Reason.* 66 (2015) 119–138.
- [25] J. Medina, M. Ojeda-Aciego, J. Ruiz-Calviño, Formal concept analysis via multi-adjoint concept lattices, *Fuzzy Sets and Systems* 160 (2) (2009) 130–144.
- [26] L. Zedam, H. Bouremel, B. De Baets, Left- and right-compatibility of order relations and fuzzy tolerance relations, *Fuzzy Sets and Systems* 360 (2019) 65–81, Theme: Fuzzy Relations.
- [27] X. Ren, D. Li, Y. Zhai, Research on mixed decision implications based on formal concept analysis, *Int. J. Cogn. Comput. Eng.* 4 (2023) 71–77.
- [28] M.E. Cornejo, D. Lobo, J. Medina, Extended multi-adjoint logic programming, *Fuzzy Sets and Systems* 388 (2020) 124–145, *Logic*.
- [29] P. Julián-Iranzo, J. Medina, M. Ojeda-Aciego, On reductants in the framework of multi-adjoint logic programming, *Fuzzy Sets and Systems* 317 (2017) 27–43.
- [30] F. Chacón-Gómez, M.E. Cornejo, J. Medina, Relating decision rules and attribute implications, in: *The 16th International Conference on Concept Lattices and their Applications, CLA 2022*, 2022 (in press).
- [31] F. Pérez-Gámez, D. López-Rodríguez, P. Cordero, Á. Mora, M. Ojeda-Aciego, Simplifying implications with positive and negative attributes: A logic-based approach, *Mathematics* 10 (4) 2022.

Capítulo 6

Toma de decisiones en la teoría de conjuntos rugosos difusos

Título:	Decision-making in Fuzzy Rough Set Theory
Autores:	Fernando Chacón-Gómez, M. Eugenia Cornejo, Jesús Medina
Revista:	Mathematics
Factor de impacto JCR:	2,4. Posición 23 de 329 (Q1)
Año:	2023
Categoría:	Mathematics
Publicación:	Octubre 2023
DOI:	10.3390/math11194187

Article

Decision Making in Fuzzy Rough Set Theory

Fernando Chacón-Gómez [†], M. Eugenia Cornejo [†] and Jesús Medina ^{*,†}

Department of Mathematics, University of Cádiz, 11510 Puerto Real, Cádiz, Spain;
fernando.chacon@uca.es (F.C.-G.); mariaeugenia.cornejo@uca.es (M.E.C.)

* Correspondence: jesus.medina@uca.es

† These authors contributed equally to this work.

Abstract: Decision rules are powerful tools to manage information and to provide descriptions of data sets; as a consequence, they can acquire a useful role in decision-making processes where fuzzy rough set theory is applied. This paper focuses on the study of different methods to classify new objects, which are not considered in the starting data set, in order to determine the best possible decision for them. The classification methods are supported by the relevance indicators associated with decision rules, such as support, certainty, and credibility. Specifically, the first one is based on how the new object matches decision rules that describe the data set, while the second one also takes into account the representativeness of these rules. Finally, the third and fourth methods take into account the credibility of the rules compared with the new object. Moreover, we have shown that these methods are richer alternatives or generalize other approaches given in the literature.

Keywords: fuzzy rough set theory; decision rules; classification methods

MSC: 03E72; 28E10; 94D05; 06A15



Citation: Chacón-Gómez, F.; Cornejo, M.E.; Medina, J. Decision Making in Fuzzy Rough Set Theory. *Mathematics* **2023**, *11*, 4187. <https://doi.org/10.3390/math11194187>

Received: 10 August 2023

Revised: 30 September 2023

Accepted: 3 October 2023

Published: 6 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fuzzy rough set theory (FRST) [1,2] arises as an extension of the theoretical approach introduced by Pawlak [3,4] in order to analyze relational data sets with imprecise or incomplete information. One of the most flexible approaches considers the multi-adjoint paradigm [5–7]. In FRST, data sets are represented by using decision tables containing information about objects of interest characterized by some attributes. Rule generation is an interesting task in FRST since it allows us to classify, simplify, and acquire knowledge from the data set. Decision rules, together with their relevance indicators, that is, the support, the certainty, the credibility, etc., have been widely studied from a Boolean perspective, such as in the classical framework [8,9], in incomplete decision contexts [10], taking into account bireducts [11], and in three-way rough set theory [12]. However, the use of relevance indicators on decision rules in FRST is really limited [13]. In this paper, we are interested in exploiting the powerful advantages provided by decision rules in FRST for decision-making methods.

A fundamental goal in decision-making methods is the discovery of patterns and the application of this knowledge to new scenarios, which have not been considered in the starting data set, in order to determine their behavior. This task has already been addressed in different fields such as artificial intelligence for dynamic classification decision making [14], for rule-driven algorithms [15], for multi-criteria decision-making methods [16,17], for clinical practice [18], and for psychotherapy [19]. In particular, the classification of new data has also been studied by the rough set theory framework due to its potential to deal with imprecise information. For instance, in [20,21], the new objects to classify are compared with the antecedents of the considered decision rules. This philosophy is similar to the clustering in machine learning [22] as it considers each rule as a cluster and studies the belonging of the new object to each of them. In [20], a vote is made among the different

decisions to be made for the new object. For that purpose, the support of the decision rules whose values in the condition attributes are exactly the same those in the new object is taken into account. However, this restrictive consideration of matching exact values makes it difficult to apply this method in real data sets. In [21], this issue is solved, since, thanks to the use of tolerance relations, a value in the unit interval is obtained, which means how much the new object resembles each of the rules. Thus, the best decision to be made for the new object is given by the highest value obtained from the tolerance relation. In addition, in [21], the support of the rules is also taken into account when it is difficult to make a decision. Note that the method presented in [23] provides a similar classification to the ones given in [20,21]. This aforementioned paper considers two kinds of attributes, categorical and numerical attributes, and it requires an exact match in the comparison between the new object and the decision rules for categorical attributes and the use of the Euclidean distance for numerical attributes. Nevertheless, in the fuzzy case, tolerance relations can be considered instead of similarity ones, and, with respect to the categorical attributes of the objects, they do not need to be boolean, but they can have a truth degree and take into account tolerance relations instead of exact matches. In this paper, we will incorporate the added benefits of considering separable tolerance relations to compare new objects with decision rules. Furthermore, the support of the rules will enrich this comparison, which will provide the possibility of obtaining extra knowledge for making a final selection when the values of the possible decisions are very close.

On the other hand, [24] presents the notion of credibility of a decision rule in terms of t-norms and implications and a method to classify new objects based on that credibility. In this paper, we will extend that notion of credibility, taking advantage of the multi-adjoint framework, that is, taking adjoint pairs into consideration. As a consequence, our classification method based on this notion will be also more general. Another notion of credibility is used in [25] based on particular operators, whose relationship with the previous notion is not clear. Hence, we will show that this notion of credibility is a particular case of the one given in [24], taking into account the Gödel t-norm and a particular fuzzy implication operator. Therefore, we will perform a thorough comparison of the methods given in [24,25] with the classification methods presented in this paper based on the credibility notion.

Specifically, this paper presents four different original methods to classify new objects, after a previous analysis of the data set from a FRST perspective. The two first methods are supported by the notions introduced in [13], which are the (1) degree of satisfiability, to measure how much the new object resembles the antecedent of decision rules extracted from the decision table, and the (2) support in the fuzzy setting, to take into account the representativeness of the rules whose antecedents are more similar to the new object. A third method, based on the notion of credibility [23,25], and a fourth one, consisting of a combination of the three previous ones, are also proposed. In summary, the main benefits of our contribution will be the definition of these methods in FRST, taking advantage of the great level of flexibility provided by the multi-adjoint framework. Furthermore, an example is worked through after presenting each method in order to illustrate them and to analyze the obtained results. We will also compare our first proposed method with the one given in [21] and the third one with the one given in [24,25], showing that our contribution presents relevant advantages, such as the flexibility inherent in the multi-adjoint framework.

The outline followed in the paper is shown below. Section 2 recalls some preliminary notions of FRST, together with an illustrative example. Section 3 introduces the four methods devoted to classifying new objects in FRST. In Section 4, we make a detailed comparison between the third classification method proposed in this paper and the ones presented in [24,25]. Next, in Section 5, a real case is studied to expose the effectiveness of each method. Finally, Section 6 presents some conclusions and prospects for future works.

2. Decision Rules in Fuzzy Rough Set Theory

In this section, important notions of FRST [5,13] are recalled. First of all, the multi-adjoint framework [5] allows us to set preferences among the objects and attributes of the decision table in order to give them more or less importance. These different considerations are based on adjoint pairs.

Definition 1 ([5]). Let $(P_1, \leq_1)$, $(P_2, \leq_2)$ and $(P_3, \leq_3)$ be partially ordered sets (posets) and $\&: P_1 \times P_2 \rightarrow P_3, I: P_2 \times P_3 \rightarrow P_1$ be mappings. Then, $(\&, I)$ is an adjoint pair with respect to P_1, P_2 , and P_3 if these mappings satisfy the adjoint property; that is,

$$\&(x, y) \leq_3 z \quad \text{iff} \quad x \leq_1 I(y, z)$$

for all $x \in P_1, y \in P_2, z \in P_3$.

Some examples of adjoint pairs are the Gödel, product, and Łukasiewicz t-norms, together with their corresponding residuated implications [26]. Next, the notion of a multi-adjoint property-oriented frame is presented.

Definition 2 ([5]). The tuple $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ is a multi-adjoint property-oriented frame where $(P_1, \leq_1)$ is a poset, $(L_2, \leq_2)$ and $(L_3, \leq_3)$ are two complete lattices, and $(\&_i, I_i)$ are adjoint pairs with respect to P_1, L_2, L_3 , for all $i \in \{1, \dots, n\}$.

In the following definition, we recall the notion of context.

Definition 3 ([5]). Let A and B be non-empty sets and $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ be a multi-adjoint property-oriented frame. A context is a tuple (A, B, R, τ) , where R is a P_1 -fuzzy relation $R: A \times B \rightarrow P_1$ and $\tau: A \times B \rightarrow \{1, \dots, n\}$ is a mapping that associates any pair of elements in $A \times B$ with a particular adjoint pair in the frame.

Now, we introduce the generalization of the notion of upper and lower approximation of FRST given in [5]. Before that, we remark that, given two sets L and X , L^X will denote the set of mappings from X to L , that is, $L^X = \{h \mid h: X \rightarrow L\}$.

Definition 4 ([5]). Let $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ be a multi-adjoint property-oriented frame and (A, B, R, τ) be a context. Given $g \in L_2^B$ and $f \in L_3^A$, we define the possibility and necessity operators, $g^{\uparrow\pi}: L_2^B \rightarrow L_3^A$ and $f^{\downarrow N}: L_3^A \rightarrow L_2^B$, respectively, as,

$$\begin{aligned} g^{\uparrow\pi}(a) &= \sup\{\&_{\tau(a,b)}(R(a,b), g(b)) \mid b \in B\} \\ f^{\downarrow N}(b) &= \inf\{I_{\tau(a,b)}(R(a,b), f(a)) \mid a \in A\} \end{aligned}$$

where $g^{\uparrow\pi}$ generalizes the classical notion of upper approximation in RST. Similarly, $f^{\downarrow N}$ generalizes the notion of lower approximation in RST.

From now on, we will focus on a multi-adjoint property-oriented frame $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$ and a context (U, U, R_A, τ) . In a proper rough set framework, the related data set is interpreted as an information or decision table instead of a formal context. In this paper, we will consider decision tables.

Definition 5 ([13]). Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A decision table is a tuple $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with the attribute a over U , and $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a}: U \rightarrow V_a\}$. In this case, the attributes of $\mathcal{A}$ are called condition attributes and d is called a decision attribute.

From a decision table, an indiscernibility relation $R_A: U \times U \rightarrow [0, 1]$ is defined on $\mathcal{A}$ and a highlighted relation $R_d: U \times U \rightarrow [0, 1]$, which shows how two objects are related with respect to the attribute d . Finally, we recall the notion of a positive region in the fuzzy environment when R_d only takes boolean values, which was also introduced in [5].

Definition 6 ([5]). Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and R_d be a boolean relation. The multi-adjoint fuzzy $\mathcal{A}$ -positive region is defined, for each $y \in U$, as

$$POS_{\mathcal{A}}^f(y) = (R_d y)^{\downarrow^N}(y) = \inf\{I_{\tau(x,y)}(R_{\mathcal{A}}(x, y), (R_d y)(x)) \mid x \in U\}$$

where $R_d y: U \rightarrow \{0, 1\}$ is defined as $(R_d y)(x) = R_d(y, x)$ and $I_{\tau(x,y)}$ is the residuated fuzzy implication of $\&_{\tau(x,y)}$ associated with the pair of objects x, y .

It is also necessary to recall a particular type of tolerance relation [27], which will be used in the rest of this work.

Definition 7 ([27]). Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, and let $a \in \mathcal{A}$ and $T_a: V_a \times V_a \rightarrow [0, 1]$ be a tolerance relation. If $T_a(v, w) = 1$ implies $v = w$ for each $v, w \in V_a$, then T_a is called a separable tolerance relation.

Thanks to the reflexivity of tolerance relations, we can conclude that a separable tolerance relation T_a satisfies that $T_a(v, w) = 1$ if and only if $v = w$ for all $v, w \in V_a$.

Now, we recall the notions related to decision rules in the fuzzy framework [13] required for this paper. The first one presents the formulas, which are indispensable to express the information contained in decision tables in logical terms. Moreover, this notion will also allow us to compute how much an object satisfies a formula.

Definition 8 ([13]). Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}_d$. The set of formulas associated with B , denoted as $For(B)$, is built from attribute-value pairs (a, v) , where $a \in B$ and $v \in V_a$, by means of the conjunction logical connective $\wedge$.

Given $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$, a family of separable $[0, 1]$ -fuzzy tolerance relations, the mapping $\|\cdot\|_S^T: For(B) \rightarrow [0, 1]^U$ is inductively defined as

$$\|\Phi\|_S^T(x) = T_a(\bar{a}(x), v)$$

for all $x \in U$ and $\Phi = (a, v)$, where $a \in B$ and $v \in V_a$. Therefore, $\|\Phi\|_S^T(x)$ is the degree of satisfiability to the formula Φ of the object x , through the relationships between the values of the attributes in the object x and the values of the attributes in the formula Φ . Moreover, we will need the α -cuts of these fuzzy sets, which are defined as

$$(\|\Phi\|_S^T)_{\alpha} = \{x_i \in U \mid \|\Phi\|_S^T(x_i) > \alpha\}$$

Finally, for every $\Phi, \Psi \in For(B)$, the conjunction of formulas is defined, for each $x \in U$, as follows:

$$\|\Phi \wedge \Psi\|_S^T(x) = \min\{\|\Phi\|_S^T(x), \|\Psi\|_S^T(x)\}$$

Now, we introduce the notion of a decision rule, which is supported by Definition 8.

Definition 9 ([13]). Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $B \subseteq \mathcal{A}$ and $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. A decision rule in S is an expression $\Phi \rightarrow \Psi$, with $\Phi \in For(B)$, $\Psi \in For(\{d\})$, where Φ and Ψ are the antecedent and the consequent of the decision rule, respectively.

In addition, we will say that an object $x \in U$ satisfies the decision rule $\Phi \rightarrow \Psi$ if $\|\Phi \wedge \Psi\|_S^T(x) = 1$. In this way, we can say that the object x induces the decision rule $\Phi \rightarrow \Psi$, which can be denoted as $\Phi_x \rightarrow \Psi_x$.

Decision rules are very important to summarize the information contained in decision tables. However, they must be analyzed in order to be interpreted. With this purpose, we will describe them by using some measures. First of all, we must recall the cardinal of a fuzzy set.

Definition 10 ([28]). Given a universe U and $V \subseteq U$, a fuzzy set characterized by a membership function $g: U \rightarrow [0, 1]$, the cardinal of V is defined as

$$\text{card}_F(V) = \sum_{x \in U} g(x)$$

Now, we introduce a pair of measures used to study the decision rules. They determine the representativeness of the rules in the considered decision table.

Definition 11. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\text{Dec}(S)$ be a set of decision rules, and $\Phi \rightarrow \Psi \in \text{Dec}(S)$ and $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. We call

- T -support of the decision rule $\Phi \rightarrow \Psi$ to the value

$$\text{supp}_S^T(\Phi, \Psi) = \text{card}_F(\|\Phi \wedge \Psi\|_S^T)$$

- Normalized T -support (NT-support, in short) of the decision rule $\Phi \rightarrow \Psi$ to the value

$$\text{supp}_S^{\text{NT}}(\Phi, \Psi) = \frac{\text{supp}_S^T(\Phi, \Psi)}{\max\{\text{supp}_S^T(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in \text{Dec}(S)\}}$$

Another important measure to describe decision rules is the credibility, already defined in [24,25,29]. The credibility of a decision rule can be seen as the degree to which, if an object satisfies the antecedent of the rule, or the antecedent of similar rules, then it also satisfies the consequent of the given rule. In this way, it provides the consistency of the rule in the table, so that the higher the credibility, the more reliable the decision according to the conditions. Now, we present the extension of this notion to the multi-adjoint framework.

Definition 12. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\text{Dec}(S)$ be a set of decision rules, $\Phi_x \rightarrow \Psi_x \in \text{Dec}(S)$ be the decision rule induced by an object $x \in U$, and $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of separable $[0, 1]$ -fuzzy tolerance relations. Given the multi-adjoint property oriented frame $([0, 1], [0, 1], [0, 1], \&_1, \dots, \&_n)$, where $\&_i$ is an associative conjunctive for all $i \in \{1, \dots, n\}$; we call α - T -credibility of the decision rule $\Phi_x \rightarrow \Psi_x$ to the value

$$\mu_{S,\tau}^\alpha(\Phi_x, \Psi_x) = \&_i\{I_{\tau(x,y)}(\|\Phi_x\|_S^T(y), \mu_{\Psi_{x,S,\tau}}(y)) \mid y \in (\|\Phi_x\|_S^T)_\alpha\}$$

with

$$\mu_{\Psi_{x,S,\tau}}(y) = \&_{j_y}\{I_{\tau(y,y')}\|\Phi_y\|_S^T(y'), \|\Psi_x\|_S^T(y')\} \mid y' \in (\|\Phi_y\|_S^T)_\alpha \text{ and } \Phi_y \rightarrow \Psi_y \in \text{Dec}(S)\}$$

where $i, j_y \in \{1, \dots, n\}$ and $\Phi_y \rightarrow \Psi_y \in \text{Dec}(S)$ is a decision rule induced by $y \in (\|\Phi_x\|_S^T)_\alpha$, if it exists in $\text{Dec}(S)$.

Note that the conjunctors $\&_{j_y}$ can be the adjoints of the implications $I_{\tau(x,y)}$ considered in the computation of $\mu_{S,\tau}^\alpha(\Phi_x, \Psi_x)$. It is important to emphasize that the α -cuts $(\|\Phi\|_S^T)_\alpha$ are considered in order to compare only the objects most similar between them given α . Moreover, the multi-adjoint framework allows us to consider different adjoint pairs for each pair of objects of the table, providing a higher level of flexibility in the studies.

In the next example, we sequentially illustrate the notions presented in this section.

Example 1. We consider the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ represented in Table 1, where the set of objects is $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, the set of attributes is $\mathcal{A} = \{a_1, a_2, a_3\}$, and $V_a = [0, 1]$ for all $a \in \mathcal{A}$ and $V_d = \{0, 1\}$; that is, d is a boolean attribute.

Table 1. Table associated with $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 1.

	a_1	a_2	a_3	d
x_1	0.34	0.31	0.75	0
x_2	0.21	0.71	0.5	1
x_3	0.52	0.92	1	0
x_4	0.85	0.65	1	1
x_5	0.43	0.89	0.5	0
x_6	0.21	0.47	0.25	1
x_7	0.09	0.93	0.25	0

In order to define decision rules to describe S , we take into account the conjunction of all the attributes of $\mathcal{A}$ for the antecedents of these rules, since they will provide us with more accurate information and will not imply many computations given that $\text{card}(\mathcal{A}) = 3$. On the other hand, the consequent obviously takes into account the only decision attribute, d . As a consequence, the next decision rules are obtained.

$$\begin{aligned}
 r_1 &: (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75) \rightarrow (d, 0) \\
 r_2 &: (a_1, 0.21) \wedge (a_2, 0.71) \wedge (a_3, 0.5) \rightarrow (d, 1) \\
 r_3 &: (a_1, 0.52) \wedge (a_2, 0.92) \wedge (a_3, 1) \rightarrow (d, 0) \\
 r_4 &: (a_1, 0.85) \wedge (a_2, 0.65) \wedge (a_3, 1) \rightarrow (d, 1) \\
 r_5 &: (a_1, 0.43) \wedge (a_2, 0.89) \wedge (a_3, 0.5) \rightarrow (d, 0) \\
 r_6 &: (a_1, 0.21) \wedge (a_2, 0.47) \wedge (a_3, 0.25) \rightarrow (d, 1) \\
 r_7 &: (a_1, 0.09) \wedge (a_2, 0.93) \wedge (a_3, 0.25) \rightarrow (d, 0)
 \end{aligned}$$

Although each decision rule is induced by an object, we denote them as $\Phi_i \rightarrow \Psi_i$ for each $i \in \{1, \dots, 7\}$ instead of $\Phi_{x_i} \rightarrow \Psi_{x_i}$ to simplify the notation. Now, $\|\Phi_i\|_S^T(x_j)$ with $i, j \in \{1, 2, \dots, 7\}$ is computed, that is, the degree of satisfiability to each antecedent of each object. For that purpose, we consider the family of separable $[0, 1]$ -fuzzy tolerance relation $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ given $T_a(\bar{a}(x), v) = 1 - |\bar{a}(x) - v|$, for all $a \in \mathcal{A}_d$, $x \in U$ and $v \in [0, 1]$. All the results are shown in Table 2.

Table 2. Degree of satisfiability of each antecedent of the decision rules of Example 1.

	$\ \Phi_1\ _S^T$	$\ \Phi_2\ _S^T$	$\ \Phi_3\ _S^T$	$\ \Phi_4\ _S^T$	$\ \Phi_5\ _S^T$	$\ \Phi_6\ _S^T$	$\ \Phi_7\ _S^T$
x_1	1	0.6	0.39	0.49	0.42	0.5	0.38
x_2	0.6	1	0.5	0.36	0.78	0.75	0.75
x_3	0.39	0.5	1	0.67	0.5	0.25	0.25
x_4	0.49	0.36	0.67	1	0.5	0.25	0.24
x_5	0.42	0.78	0.5	0.5	1	0.58	0.66
x_6	0.5	0.75	0.25	0.25	0.58	1	0.54
x_7	0.38	0.75	0.25	0.24	0.66	0.54	1

On the other hand, since d is a Boolean attribute, we deduce that

$$\|\Psi_i\|_S^T(x_j) = \|(d, v_i)\|_S^T(x_j) = \begin{cases} 1 & \text{if } \bar{d}(x_j) = v_i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In order to compute the α -T-credibility of the previous decision rules, we set the multi-adjoint property-oriented frame $([0, 1], [0, 1], [0, 1], \&_L, \&_P, \&_G)$, where $(\&_L, I_L)$, $(\&_P, I_P)$ and $(\&_G, I_G)$ are the Łukasiewicz, product and Gödel adjoint pairs, respectively, and the context (U, U, R_A, τ) with $\tau: U \times U \rightarrow \{L, P, G\}$ defined for each $x, y \in U$ as

$$\tau(x, y) = \begin{cases} L & \text{if } \bar{d}(x) = \bar{d}(y) = 1 \\ P & \text{if } \bar{d}(x) = \bar{d}(y) = 0 \\ G & \text{otherwise} \end{cases}$$

We have considered this mapping to reinforce the computation of the credibility when the objects have the same decision (implication I_G is smaller than I_P and I_L) and model the preferences of the user, who is more confident with the decision 1 than with the decision 0 (implication I_P is smaller than I_L). On the other hand, we consider the adjoint pair $(\&_L, I_L)$ and the 0.76 cut in order to compute the 0.76-T-credibility of each decision rule belonging to $\text{Dec}(S)$. This threshold has been chosen to minimize the number of objects related via the antecedent but not via the consequent in the sets $(\|\Phi\|_S^T)_\alpha$ with $\Phi \rightarrow \Psi \in \text{Dec}(S)$, while some flexibility is allowed in the comparisons. In this way, thanks to this choice of α in Table 2, we deduce that the only inconsistencies in these sets are caused by the objects x_2 and x_5 .

By using Table 2, Equation (1), and Definitions 11 and 12, we compute the T support, NT support, and 0.76-T-credibility of each decision rule. All these measures are shown in Table 3.

Table 3. T support, NT support, and 0.76-T-credibility of the decision rules of Example 1.

Rule	supp_S^T	supp_S^{NT}	$\mu_{S,\tau}^{0.76}$
r_1	2.19	0.85	1
r_2	2.11	0.82	0
r_3	2.14	0.83	1
r_4	1.61	0.62	1
r_5	2.58	1	0
r_6	2	0.78	1
r_7	2.29	0.89	1

The decision rules $\Phi \rightarrow \Psi \in \text{Dec}(S)$ with 0.76-T-credibility equal to 1 satisfy that $(\|\Phi\|_S^T)_{0.76} = \{x\}$, where $x \in U$ is the object that induces them. On the other hand, $\mu_{S,\tau}^{0.76}(\Phi_2, \Psi_2) = \mu_{S,\tau}^{0.76}(\Phi_5, \Psi_5) = 0$ since $\|\Phi_2\|_S^T(x_5) = \|\Phi_5\|_S^T(x_2) = 0.78 > 0.76$; that is, $(\|\Phi_2\|_S^T)_{0.76} = (\|\Phi_5\|_S^T)_{0.76} = \{x_2, x_5\}$, whereas $\|\Psi_2\|_S^T(x_5) = \|\Psi_5\|_S^T(x_2) = 0$. Therefore, rules r_2 and r_5 are very related in the antecedent but not at all in the consequent. As a consequence, these rules do not have any credibility.

In conclusion, in this section, we have recalled important notions to study data sets in the FRST framework. In particular, we must emphasize the relevance of the notions related to decision rules, since they will be indispensable to classify objects outside of the data set. This task is carried out in the following section.

3. Classification of New Objects

In this section, we analyze different methods to make decisions in the FRST framework, and study their properties and which method is the most suitable, depending on the case. With this purpose, we consider a set of decision rules obtained from a decision table

that represents it, and we will try to classify new objects not belonging to that table. For that purpose, we will know the values of these objects based on the condition attributes and we will detect the most suitable value of the decision attribute, which is originally unknown. Hence, we are considering an extended universe $\mathcal{U}$ where we have two different classes of objects, the objects in U from which the decision value is known and the objects in $\mathcal{U} \setminus U = U^c$ from which the decision value is unknown. That is, the main goal is to determine the value of the decision attribute in these objects in order to make predictions about them. For that purpose, we will study each possible decision that can be made for the new objects by comparing the values given in the condition attributes of the new objects and the values given in the rules of the set considered. From this study, a truth value is obtained for each possible decision that can be made. Finally, the recommended decision will be determined by paying attention to the highest truth value among all of them. We will study four classification methods with different features.

3.1. Suitability Method

This section is devoted to presenting the first method we will use to make a decision for the new objects, which extends the one given in [21]. The comparison with this method will be given at the end of the section. This classification method will take into account Definition 8, since we will compute how much the new objects satisfy the antecedents of the rules considered in the study. In particular, the next notion focuses only on the maximum degree of satisfiability of the antecedents for each possible consequent so that a truth value is provided for all the possible values of the decision attribute.

Definition 13. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $Dec(S)$ be a set of decision rules. For each new object $x \in U^c$ and $v \in V_d$, the degree of suitability of the decision $\bar{d}(x) = v$ according to $Dec(S)$ is given via the mapping $\sigma_{Dec(S)}: U^c \times V_d \rightarrow [0, 1]$, defined as

$$\sigma_{Dec(S)}(x, v) = \max\{\|\Phi\|_S^T(x) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

Notice that, in Definition 8, $\|\Phi\|_S^T$ is evaluated in objects $x \in U$ to compute how much each object of U satisfies the formula Φ , while in Definition 13, $\|\Phi\|_S^T(x)$ is calculated with $x \in U^c$. However, this computation is carried out to compare new objects with those present in the original decision table. Moreover, the degree of suitability is a truth value, since $0 \leq \sigma_{Dec(S)}(x, v) \leq 1$ for all new objects x , decisions $\bar{d}(x) = v$, and sets of decision rules $Dec(S)$ according to Definition 8.

Once the degree of suitability of a given decision has been defined, we present the method used to classify new objects according to Definition 13, which will be called the *suitability method* or *S-method* for short.

- S1. First of all, given a new object to study $x \in U^c$, we compute the degrees of suitability of each possible decision $\bar{d}(x) = v$, with $v \in V_d$.
- S2. Then, all the obtained degrees of suitability are compared. If the degree of suitability of a decision $\bar{d}(x) = v$ is high and different enough from the degrees of suitability of the rest of the possible decisions, then we will consider the decision $\bar{d}(x) = v$ for the object x .
- S3. If there exist some values of the decision attribute $v \in V \subset V_d$ with similar degrees of suitability, then we must consider more degrees of satisfiability for these decisions instead of only the maximum depending on a threshold $\beta \in [0, 1]$. In particular, denoting as $\mathcal{F}_x^\beta(v) = \{\Phi \in For(\mathcal{A}) \mid \Phi \rightarrow \Psi \in Dec(S), \beta \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}$, we will consider for each decision $\bar{d}(x) = v$, with $v \in V$, the value

$$\mathcal{M}_{\mathcal{F}_x^\beta(v)} = \sum_{\Phi \in \mathcal{F}_x^\beta(v)} \frac{\|\Phi\|_S^T(x)}{\text{card}(\mathcal{F}_x^\beta(v))}$$

Once this computation is carried out for each $v \in V$, the obtained values are compared, and we come back to the second step.

- S4. If it is not possible to make a decision according to the second step, then the set $\mathcal{F}_x^\beta(v)$ is modified, either by changing the value β or by considering another aggregation operator instead of the arithmetic mean.

We illustrate this first method in the following example.

Example 2. Coming back to the environment of Example 1, we will study three new objects, x_8 , x_9 , and x_{10} , in order to determine the best decision for them. In Table 4, we show these objects together with the objects previously presented in Table 1.

Table 4. Table 1 together with the new objects to classify x_8 , x_9 , and x_{10} .

	a_1	a_2	a_3	d
x_1	0.34	0.31	0.75	0
x_2	0.21	0.71	0.5	1
x_3	0.52	0.92	1	0
x_4	0.85	0.65	1	1
x_5	0.43	0.89	0.5	0
x_6	0.21	0.47	0.25	1
x_7	0.09	0.93	0.25	0
x_8	0.05	0.95	0.1	
x_9	0.89	0.5	0.95	
x_{10}	0.4	0.67	0.31	

First of all, we compute the degree of satisfiability to each antecedent of the decision rules of Example 1 of each new object in Table 4 considering again the $[0, 1]$ -fuzzy tolerance relation given as $T_a(\bar{a}(x), v) = 1 - |\bar{a}(x) - v|$ for each $a \in \mathcal{A}$. These results are presented in Table 5.

Table 5. Degree of satisfiability of each antecedent of decision rules from Example 1.

	$\ \Phi_1\ _S^T$	$\ \Phi_2\ _S^T$	$\ \Phi_3\ _S^T$	$\ \Phi_4\ _S^T$	$\ \Phi_5\ _S^T$	$\ \Phi_6\ _S^T$	$\ \Phi_7\ _S^T$
x_8	0.35	0.6	0.1	0.1	0.6	0.52	0.85
x_9	0.45	0.32	0.58	0.85	0.54	0.3	0.2
x_{10}	0.56	0.81	0.31	0.31	0.78	0.8	0.69

Now, we follow the method presented above to make a decision for the objects x_8 , x_9 and x_{10} .

- S1. We begin computing the degrees of suitability of each possible decision for these objects. Then, using Definition 13, we obtain that

$$\sigma_{Dec(S)}(x_8, 0) = \|\Phi_7\|_S^T(x_8) = 0.85$$

$$\sigma_{Dec(S)}(x_8, 1) = \|\Phi_2\|_S^T(x_8) = 0.6$$

$$\sigma_{Dec(S)}(x_9, 0) = \|\Phi_3\|_S^T(x_9) = 0.58$$

$$\sigma_{Dec(S)}(x_9, 1) = \|\Phi_4\|_S^T(x_9) = 0.85$$

$$\sigma_{Dec(S)}(x_{10}, 0) = \|\Phi_5\|_S^T(x_{10}) = 0.78$$

$$\sigma_{Dec(S)}(x_{10}, 1) = \|\Phi_2\|_S^T(x_{10}) = 0.81$$

- S2. Regarding to the object x_8 , there is a remarkable difference between the values $\sigma_{Dec(S)}(x_8, 0) = 0.85$ and $\sigma_{Dec(S)}(x_8, 1) = 0.6$. Consequently, we consider the decision $\bar{d}(x_8) = 0$ for the object x_8 with a degree of suitability of 0.85.

The object x_9 is also easy to decide according to this first method, since both degrees of suitability are significantly different, considering that $\sigma_{Dec(S)}(x_9, 0) = 0.58$ and $\sigma_{Dec(S)}(x_9, 1) = 0.85$. Hence, we consider the decision $\bar{d}(x_9) = 1$ for the object x_9 with a degree of suitability of 0.85. Finally, with respect to the object x_{10} , it is not possible to make a decision for it due to both degrees of suitability, namely $\sigma_{Dec(S)}(x_{10}, 0) = 0.78$ and $\sigma_{Dec(S)}(x_{10}, 1) = 0.81$, are not sufficiently different. As a consequence, we need to carry out more computations to be able to determine the best decision for this object.

- S3. First of all, we set the threshold $\beta = 0.7$ and we compute the set $\mathcal{F}_{x_{10}}^{0.7}(v)$ for $v \in V_d = \{0, 1\}$. Based on Table 5, we have that $\mathcal{F}_{x_{10}}^{0.7}(0) = \{\Phi_5\}$ and $\mathcal{F}_{x_{10}}^{0.7}(1) = \{\Phi_2, \Phi_6\}$. Thus, we obtain that

$$\begin{aligned}\mathcal{M}_{\mathcal{F}_{x_{10}}^{0.7}(0)} &= \|\Phi_5\|_S^T(x_{10}) \\ &= 0.78 \\ \mathcal{M}_{\mathcal{F}_{x_{10}}^{0.7}(1)} &= \frac{\|\Phi_2\|_S^T(x_{10}) + \|\Phi_6\|_S^T(x_{10})}{2} = \frac{0.81 + 0.8}{2} \\ &= 0.81\end{aligned}$$

We can conclude that it is still difficult to make a decision for the object x_{10} . Therefore, we must continue to study this object.

- S4. Now, we set the threshold $\beta = 0.5$. Hence, it is deduced that $\mathcal{F}_{x_{10}}^{0.5}(0) = \{\Phi_1, \Phi_5, \Phi_7\}$ and $\mathcal{F}_{x_{10}}^{0.5}(1) = \{\Phi_2, \Phi_6\}$. Thus, we obtain that

$$\begin{aligned}\mathcal{M}_{\mathcal{F}_{x_{10}}^{0.5}(0)} &= \frac{\|\Phi_1\|_S^T(x_{10}) + \|\Phi_5\|_S^T(x_{10}) + \|\Phi_7\|_S^T(x_{10})}{3} = \frac{0.56 + 0.78 + 0.69}{3} \\ &= 0.68 \\ \mathcal{M}_{\mathcal{F}_{x_{10}}^{0.5}(1)} &= \frac{\|\Phi_2\|_S^T(x_{10}) + \|\Phi_6\|_S^T(x_{10})}{2} = \frac{0.81 + 0.8}{2} \\ &= 0.81\end{aligned}$$

It is easy to check that the difference between both values has increased. As a result, following this method, we will adopt the decision $\bar{d}(x_{10}) = 1$ for the object x_{10} with a degree of suitability of 0.81. However, we should take into account that we have notably decreased the threshold β .

Thus, we must remark that, while the decision $\bar{d}(x_8) = 0$ and $\bar{d}(x_9) = 1$ for the objects x_8 and x_9 , respectively, seems quite clear, the object x_{10} has been more difficult to classify, needing more steps of the presented method to be applied. However, even then, the decision for the object x_{10} is not as clear as for the objects x_8 and x_9 .

Hence, the difficulty in classifying the object x_{10} in the previous example has shown that it is necessary to find another approach to the decision-making process to obtain a more confident result, which is presented in the next section.

It is important to mention that the two first steps of this classification method are similar to the ones carried out in [21]. However, we have presented this method in the fuzzy framework by using separable $[0, 1]$ fuzzy tolerance relations to compute the degree of satisfiability to antecedents of decision rules of the objects. We have chosen this kind of relation because we consider it indispensable that a degree of satisfiability to an antecedent is 1 only in the cases that the values taken by the attributes in the new object are the same as those present in that antecedent, because otherwise, we would be losing information. On the other hand, in [21], tolerance relations are considered for attributes $a \in \mathcal{A}$ such that $V_a \subseteq \mathbb{R}$ and separable tolerance relations for the rest of attributes. Moreover, they consider a discordance relation to make a penalty in the degree of satisfiability to the considered antecedent of the new object due to the existence of some attributes in that antecedent with a very different value to that taken by these attributes in the new object. In our work, we

have discarded this penalization because it is already implicit in the degree of satisfiability. Furthermore, the inclusion of the two last steps of the classification method provides more flexibility to the study, allowing us to make decisions in those cases where the method in [21] is unable to offer a clear decision.

3.2. ε -Representativeness Method

As we have shown in the previous section, the suitability method cannot provide good enough results in some cases, such as in the study of the object x_{10} in Example 2. However, in that method, all the rules have the same importance, and it is not taken into account how many objects satisfy them (that is, the T -support). Therefore, the method presented in this section will be based on both the degrees of satisfiability to the antecedents of the rules compared with the new object and the T -support of those rules.

Definition 14. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $Dec(S)$ be a set of decision rules. Given $\varepsilon \in [0, 1]$, for each new object $x \in U^c$ and $v \in V_d$ the degree of ε -representativeness of the decision $\bar{d}(x) = v$ according to $Dec(S)$ is given by the mapping $\rho_{Dec(S)}^\varepsilon : U^c \times V_d \rightarrow [0, 1]$ defined as

$$\rho_{Dec(S)}^\varepsilon(x, v) = \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}$$

If $\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\} = \emptyset$, the object x is ε -undecidable.

It is easy to check that $0 \leq \rho_{Dec(S)}^\varepsilon(x, v) \leq 1$, regardless of the object x to classify, the decision $\bar{d}(x) = v$ studied, and the set of decision rules $Dec(S)$ considered, so it is a truth value. We must emphasize the importance of the threshold ε in Definition 14. If a high value is fixed, then only the most similar antecedents to the new object are considered, so a more restrictive study is carried out. On the other hand, if a low value is considered, then many antecedents are compared with the new object, so the classification is less accurate. As a consequence, different values of ε can be considered depending on the flexibility chosen by the user. In addition, we have deemed indispensable the inclusion of the factor $\|\Phi\|_S^T(x)$ to weigh the T -supports of the decision rules and thus give more importance to the rules with the most similar antecedents to the new object x .

The next result shows a pair of interesting properties relating to Definitions 13 and 14.

Proposition 1. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $Dec(S)$ be a set of decision rules, $x \in U^c$ be a new object, and $\varepsilon \in [0, 1]$. Then

- $\sum_{v \in V_d} \rho_{Dec(S)}^\varepsilon(x, v) = 1$.
- $\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d\} < \varepsilon$ if and only if the object x is ε -undecidable.

Proof. The first item is immediate since

$$\begin{aligned} \sum_{v \in V_d} \rho_{Dec(S)}^\varepsilon(x, v) &= \sum_{v \in V_d} \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &\stackrel{(1)}{=} \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)} = 1 \end{aligned}$$

where (1) is obtained due to

$$\sum_{v \in V_d} \sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x), \Psi = (d, v)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi) = \sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\}} \|\Phi\|_S^T(x) \cdot \text{supp}_S^T(\Phi, \Psi)$$

For the second item, we suppose that $\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d\} < \varepsilon$. Then, we have that $\sigma_{Dec(S)}(x, v) < \varepsilon$ for each $v \in V_d$. Hence, $\|\Phi\|_S^T(x) < \varepsilon$ for each $\Phi \rightarrow \Psi \in Dec(S)$. Therefore,

$$\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\} = \emptyset$$

Consequently, the object x is ε -undecidable.

Now, we suppose that the object x is ε -undecidable. Using *reductio ad absurdum*, we assume that $\varepsilon \leq \max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d\}$. Therefore, there exists $v \in V_d$ such that $\varepsilon \leq \sigma_{Dec(S)}(x, v)$. As a consequence, by Definition 13, there exists $\Phi \rightarrow \Psi \in Dec(S)$ with $\Psi = (d, v)$ such that $\varepsilon \leq \|\Phi\|_S^T(x)$, which leads to a contradiction, given that x is ε -undecidable. $\square$

From the first item of Proposition 1, it is possible to clearly differentiate a possible decision from the others by choosing an appropriate threshold ε , as the next result shows.

Lemma 1. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $Dec(S)$ be a set of decision rules, and $x \in U^c$ be a new object. If there exists $v_k \in V_d$ and $\varepsilon \in [0, 1]$ satisfying

$$\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d \setminus \{v_k\}\} < \varepsilon \leq \sigma_{Dec(S)}(x, v_k)$$

then

$$\begin{aligned} \rho_{Dec(S)}^\varepsilon(x, v_k) &= 1 \\ \rho_{Dec(S)}^\varepsilon(x, v_j) &= 0 \end{aligned}$$

for all $v_j \in V_d \setminus \{v_k\}$.

Proof. First of all, since $\varepsilon \leq \sigma_{Dec(S)}(x, v_k)$, via Proposition 1, we deduce that x is not ε -undecidable. Consequently,

$$\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x)\} \neq \emptyset$$

On the other hand, given that $\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d \setminus \{v_k\}\} < \varepsilon$, we obtain that $\sigma_{Dec(S)}(x, v_j) < \varepsilon$ for all $v_j \in V_d \setminus \{v_k\}$. Thus, using Definition 13, we deduce that $\{\Phi \rightarrow \Psi \in Dec(S) \mid \varepsilon \leq \|\Phi\|_S^T(x), \Psi = (d, v_j)\} = \emptyset$ for all $v_j \in V_d \setminus \{v_k\}$. Therefore, using Definition 14, for each $v_j \in V_d \setminus \{v_k\}$, we conclude that $\rho_{Dec(S)}^\varepsilon(x, v_j) = 0$. Finally, using Proposition 1, it is obtained that $\rho_{Dec(S)}^\varepsilon(x, v_k) = 1 - \sum_{v \in V_d \setminus \{v_k\}} \rho_{Dec(S)}^\varepsilon(x, v) = 1$. $\square$

This result presents a great advance with respect to the suitability method, since it allows one to select a decision above the rest by choosing an adequate value of ε . However, this value may not be satisfactory for the user since it may be too high. Therefore, we present a classification method based on Definition 14, called the ε -representativeness method, or the *R-method* for short, in which the threshold ε can be chosen by the user.

- R1. Given a new object to classify $x \in U^c$ and once the threshold ε is fixed, we compute the degrees of ε -representativeness $\rho_{Dec(S)}^\varepsilon(x, v)$ for all $v \in V_d$.
- R2. After that, all the obtained degrees of ε -representativeness are compared. If the degree of ε -representativeness of a decision $\bar{d}(x) = v$ is high and different enough from the

rest of the degrees of ε -representativeness, then we will consider the decision $\bar{d}(x) = v$ for the object x .

- R3. If there exist some values of the decision attribute with similar degrees of ε -representativeness, then the value ε must be changed to be more or less flexible in the study, depending on the case.

It is important to emphasize that if an object x is ε -undecidable, then it is impossible to make a decision for it following this method, given that $\rho_{Dec(S)}^\varepsilon(x, v)$ cannot be computed for any $v \in V_d$. In order to solve this problem, the value ε must be changed.

In the following example, we classify the new objects in Table 4 by using Definition 14, Proposition 1 and Lemma 1.

Example 3. Coming back to the environment of Example 1 and taking into account the degrees of suitability obtained in Example 2, we will determine the best value to be taken by the decision attribute in the new objects x_8 , x_9 and x_{10} according to the ε -representativeness method. Since this method has been introduced to improve the first one, it is expected that it will classify x_8 and x_9 as the suitability method and determine a clearer decision for x_{10} .

- R1. First of all, we set the value $\varepsilon = 0.7$. Notice that it is close to 1, so the antecedents taken into account in the study are significantly similar to the new object, and it allows some flexibility in the comparatives. Notice that, for the object x_8 , the value $v = 0$ and $\varepsilon = 0.7$ verifies the condition of Lemma 1, that is

$$\max\{\sigma_{Dec(S)}(x, v) \mid v \in V_d \setminus \{0\}\} = \sigma_{Dec(S)}(x_8, 1) = 0.6 < \varepsilon \leq 0.85 = \sigma_{Dec(S)}(x_8, 0)$$

As a consequence, $\rho_{Dec(S)}^{0.7}(x_8, 1) = 0$ and $\rho_{Dec(S)}^{0.7}(x_8, 0) = 1$.

In a similar way, for the object x_9 , we have that

$$\sigma_{Dec(S)}(x_9, 0) = 0.58 < \varepsilon \leq 0.85 = \sigma_{Dec(S)}(x_9, 1)$$

Thus, from Lemma 1, $\rho_{Dec(S)}^{0.7}(x_9, 0) = 0$ and $\rho_{Dec(S)}^{0.7}(x_9, 1) = 1$.

Finally, we study the classification of the object x_{10} . From Definition 14 and taking into account Tables 3 and 5, we obtain that

$$\begin{aligned} \rho_{Dec(S)}^{0.7}(x_{10}, 0) &= \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid 0.7 \leq \|\Phi\|_S^T(x_{10}), \Psi = (d, 0)\}} \|\Phi\|_S^T(x_{10}) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) \mid 0.7 \leq \|\Phi\|_S^T(x_{10})\}} \|\Phi\|_S^T(x_{10}) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{\|\Phi_5\|_S^T(x_{10}) \cdot \text{supp}_S^T(\Phi_5, \Psi_5)}{\sum_{\Phi \rightarrow \Psi \in \{\Phi_2 \rightarrow \Psi_2, \Phi_5 \rightarrow \Psi_5, \Phi_6 \rightarrow \Psi_6\}} \|\Phi\|_S^T(x_{10}) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{0.78 \cdot 2.58}{0.81 \cdot 2.11 + 0.78 \cdot 2.58 + 0.8 \cdot 2} = \frac{2.01}{1.71 + 2.01 + 1.6} \\ &= \frac{2.01}{5.32} = 0.38 \end{aligned}$$

On the other hand, from Proposition 1, we have that

$$\rho_{Dec(S)}^{0.7}(x_{10}, 0) + \rho_{Dec(S)}^{0.7}(x_{10}, 1) = 1$$

As a result, we deduce that

$$\rho_{Dec(S)}^{0.7}(x_{10}, 1) = 1 - 0.38 = 0.62$$

- R2. With respect to the object x_8 , the decision to be made is clear, given that $\rho_{Dec(S)}^{0.7}(x_8, 1) = 0$ and $\rho_{Dec(S)}^{0.7}(x_8, 0) = 1$. In this way, the recommended decision for the object x_8 following the ε -representativeness method is $\bar{d}(x_8) = 0$ with a degree of 0.7 representativeness of 1.

For the object x_9 , we obtain similar conclusions. Since $\rho_{Dec(S)}^{0.7}(x_9, 0) = 1$ and $\rho_{Dec(S)}^{0.7}(x_9, 1) = 0$, this method recommends the decision $\bar{d}(x_9) = 1$ for the object x_9 with a degree of 0.7-representativeness of 1.

Finally, regarding the object x_{10} , we obtained the results that $\rho_{Dec(S)}^{0.7}(x_{10}, 1) = 0.62$ and $\rho_{Dec(S)}^{0.7}(x_{10}, 0) = 0.38$. Although the decision for the object x_{10} is not as clear as for the other objects, this method recommends the decision $\bar{d}(x_{10}) = 1$ with a degree of 0.7 representativeness of 0.62.

Now, we present some remarks about the classification of these objects, taking advantage of Proposition 1 and Lemma 1. First of all, for the object x_8 , if $\varepsilon \in (\sigma_{Dec(S)}(x_8, 1), \sigma_{Dec(S)}(x_8, 0)] = (0.6, 0.85]$ from Lemma 1, we have that $\rho_{Dec(S)}^{0.7}(x_8, 1) = 0$ and $\rho_{Dec(S)}^{0.7}(x_8, 0) = 1$. As a consequence, for many values ε , we have a clear classification for it. On the other hand, from Proposition 1, if we consider a value ε such that

$$\max\{\sigma_{Dec(S)}(x_8, v) \mid v \in \{0, 1\}\} = \sigma_{Dec(S)}(x_8, 0) = 0.85 < \varepsilon$$

then the object x_8 is ε -undecidable. Finally, we have studied the degree of ε -representativeness of both possible decisions for each $\varepsilon \in [0, 0.85]$, obtaining that

$$\rho_{Dec(S)}^{\varepsilon}(x_8, 1) < \rho_{Dec(S)}^{\varepsilon}(x_8, 0) \quad \text{if } \varepsilon \leq 0.85$$

As a consequence, for any $\varepsilon \in [0, 0.85]$ the most recommended decision for the object x_8 is $\bar{d}(x_8) = 0$.

Now, for the object x_9 , if $\varepsilon \in (\sigma_{Dec(S)}(x_9, 0), \sigma_{Dec(S)}(x_9, 1)] = (0.58, 0.85]$ from Lemma 1, $\rho_{Dec(S)}^{0.7}(x_9, 0) = 0$ and $\rho_{Dec(S)}^{0.7}(x_9, 1) = 1$. Once again, if

$$\max\{\sigma_{Dec(S)}(x_9, v) \mid v \in \{0, 1\}\} = \sigma_{Dec(S)}(x_9, 1) = 0.85 < \varepsilon$$

the object x_9 is ε -undecidable according to Proposition 1. Moreover, it can be checked that

$$\begin{aligned} \rho_{Dec(S)}^{\varepsilon}(x_9, 1) &< \rho_{Dec(S)}^{\varepsilon}(x_9, 0) \quad \text{if } 0 \leq \varepsilon \leq 0.54 \\ \rho_{Dec(S)}^{\varepsilon}(x_9, 0) &< \rho_{Dec(S)}^{\varepsilon}(x_9, 1) \quad \text{if } 0.54 < \varepsilon \leq 0.85 \end{aligned}$$

Although the degree of ε -representativeness changes with respect to the value ε , it is clear that the decision $\bar{d}(x_9) = 1$ should be made, since the degree of ε -representativeness of this decision is higher when higher values of ε are considered and, as we mentioned above, the study is more accurate in these cases.

Finally, for the object x_{10} , if $\varepsilon \in (\sigma_{Dec(S)}(x_{10}, 0), \sigma_{Dec(S)}(x_{10}, 1)] = (0.78, 0.81]$ then by Lemma 1 $\rho_{Dec(S)}^{0.7}(x_{10}, 0) = 0$ and $\rho_{Dec(S)}^{0.7}(x_{10}, 1) = 1$. Moreover, from Proposition 1, the object x_{10} is ε -undecidable if

$$\max\{\sigma_{Dec(S)}(x_{10}, v) \mid v \in \{0, 1\}\} = \sigma_{Dec(S)}(x_{10}, 1) = 0.81 < \varepsilon$$

which is a lower threshold than the present in the previous objects, allowing us to choose fewer values of ε to classify this object. On the other hand, it can be checked that

$$\begin{aligned} \rho_{Dec(S)}^{\varepsilon}(x_{10}, 1) &< \rho_{Dec(S)}^{\varepsilon}(x_{10}, 0) \quad \text{if } 0 \leq \varepsilon \leq 0.69 \\ \rho_{Dec(S)}^{\varepsilon}(x_{10}, 0) &< \rho_{Dec(S)}^{\varepsilon}(x_{10}, 1) \quad \text{if } 0.69 < \varepsilon \leq 0.81 \end{aligned} \quad (2)$$

In this case, the recommended decision for the object x_{10} changes with a higher value of ε and the object is ε -undecidable with a lower ε . This fact also exposes the higher difficulty of classifying the object x_{10} because Equation (2) takes place for fewer values ε than in the object x_9 . Notice that this problem of classification was also reflected in Example 2, when we had to apply the third step of the suitability method to be able to decide for this object. Nevertheless, the decision provided by the

ε -representativeness method is more accurate when ε is higher, as we commented above. In this way, the decision $\bar{d}(x_{10}) = 1$ is more confident since, in these cases, $\rho_{Dec(S)}^\varepsilon(x_{10}, 0) < \rho_{Dec(S)}^\varepsilon(x_{10}, 1)$.

Notice that we have obtained more conclusive results than in Example 2 thanks to the consideration of the T support of the decision rules. This is due to the fact that the difference between the degrees of 0.7 representativeness of both possible decisions for each new object is greater than the corresponding differences between the degrees of suitability.

3.3. α -Credibility Method

Some authors [24,25] have considered the measure of the credibility to classify new objects to be important. In this section, we will propose a method in which the α - T -credibility plays an indispensable role. The corresponding method will be based on the following definition.

Definition 15. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $Dec(S)$ be a set of decision rules, $\alpha \in [0, 1)$, and $(\&, I)$ be an adjoint pair. For each new object $x \in U^c$ and $v \in V_d$, the degree of α -credibility of the decision $\bar{d}(x) = v$ according to $Dec(S)$ is given by the mapping $\theta_{Dec(S)}^\alpha: U^c \times V_d \rightarrow [0, 1]$, defined as

$$\theta_{Dec(S)}^\alpha(x, v) = \max\{\&(\|\Phi\|_S^T(x), \mu_{S,\tau}^\alpha(\Phi, \Psi)) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

As in the previous notions, the degree of α -credibility is a truth value satisfying that $0 \leq \theta_{Dec(S)}^\alpha(x, v) \leq 1$ for all new objects x , decisions $\bar{d}(x) = v$, $\alpha \in [0, 1)$ and sets of decision rules $Dec(S)$. The third proposed classification method, called α -credibility method, or C -method for short, which is based on Definition 15, is presented below.

- C1. Given a new object to classify $x \in U^c$, we compute the degrees of α -credibility $\theta_{Dec(S)}^\alpha(x, v)$ for all $v \in V_d$.
- C2. Next, we compare all the obtained degrees of α -credibility. If there exists a decision $w \in V_d$ such that $\theta_{Dec(S)}^\alpha(x, w)$ is high and different enough from $\theta_{Dec(S)}^\alpha(x, v)$, for each $v \in V_d \setminus \{w\}$, then the decision $\bar{d}(x) = w$ will be chosen for the object x .
- C3. If there exist some values of the decision attribute with similar degrees of α -credibility, then we must re-examine the relationship between the variables involved in the computation. Then, it is possible to select another adjoint pair $(\&, I)$ from the considered multi-adjoint property framework in order to aggregate the variables in a better way.

Now, we classify the new objects for Example 2 by using this method.

Example 4. Coming back to the environment of Example 2, we study the best decision to make for the objects x_8 , x_9 , and x_{10} according to the α -credibility method.

- C1. First of all, we compute the degrees of 0.76 credibility of each decision for the new objects. With this purpose, we consider the Łukasiewicz adjoint pair $(\&_L, I_L)$, since we assume that the degree of satisfiability and the α - T -credibility are practically independent. For the object x_8 and decision $\bar{d}(x_8) = 0$, via Tables 3 and 5 and Definition 15, we obtain that

$$\begin{aligned} \theta_{Dec(S)}^{0.76}(x_8, 0) &= \max\{\&_L(\|\Phi_1\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_1, \Psi_1)), \&_L(\|\Phi_3\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3)), \\ &\quad \&_L(\|\Phi_5\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_5, \Psi_5)), \&_L(\|\Phi_7\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_7, \Psi_7))\} \\ &= \max\{\&_L(0.35, 1), \&_L(0.1, 1), \&_L(0.6, 0), \&_L(0.85, 1)\} \\ &= \max\{0.35, 0.1, 0, 0.85\} = 0.85 \end{aligned}$$

Analogously, the rest of the degrees of 0.76 credibility are computed as follows:

$$\begin{aligned}\theta_{Dec(S)}^{0.76}(x_8, 1) &= \&_L(\|\Phi_6\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_6, \Psi_6)) = 0.52 \\ \theta_{Dec(S)}^{0.76}(x_9, 0) &= \&_L(\|\Phi_3\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3)) = 0.58 \\ \theta_{Dec(S)}^{0.76}(x_9, 1) &= \&_L(\|\Phi_4\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_4, \Psi_4)) = 0.85 \\ \theta_{Dec(S)}^{0.76}(x_{10}, 0) &= \&_L(\|\Phi_7\|_S^T(x_{10}), \mu_{S,\tau}^{0.76}(\Phi_7, \Psi_7)) = 0.69 \\ \theta_{Dec(S)}^{0.76}(x_{10}, 1) &= \&_L(\|\Phi_6\|_S^T(x_{10}), \mu_{S,\tau}^{0.76}(\Phi_6, \Psi_6)) = 0.8\end{aligned}$$

- C2. For the object x_8 , there is a big difference between the degrees of 0.76 credibility of both possible decisions. Therefore, the α -credibility method recommends the decision $\bar{d}(x_8) = 0$ for the object x_8 with a degree of 0.76-credibility of 0.85.

On the other hand, the case of the object x_9 is very similar. Since $\theta_{Dec(S)}^{0.76}(x_9, 0) = 0.58$ and $\theta_{Dec(S)}^{0.76}(x_9, 1) = 0.85$, the recommended decision for the object x_9 according to the α -credibility method is $\bar{d}(x_9) = 1$ with a degree of 0.76 credibility of 0.85.

Finally, with respect to the object x_{10} , the decision is less clear than in the objects x_8 and x_9 , as the degrees of 0.76 credibility of both possible decisions are more similar. Nevertheless, we consider that the difference between both degrees of 0.76-credibility is sufficient to conclude that the best decision for the object x_{10} is $\bar{d}(x_{10}) = 1$ with a degree of 0.76 credibility of 0.8. Anyway, the support decision system incorporating this procedure will show that the other decision has a 0.76 credibility with a degree of 0.69.

Now, we make some remarks about the results obtained. First of all, notice that the decisions suggested by the α -credibility method are the same as those recommended by the methods presented in Sections 3.1 and 3.2. In fact, the degrees of 0.76 credibility of the decisions $\bar{d}(x_8) = 0$ and $\bar{d}(x_9) = 1$ in Example 4 coincide with their degrees of suitability in Example 2. This is due to the fact that the decision rules involved in the computation of the degrees of suitability of these decisions, which are r_7 and r_4 , respectively, have a 0.76-T-credibility equal to 1. Finally, the decision for the object x_{10} has been slightly more difficult to make, as in Example 2, where the T support of the rules was also not taken into account. Therefore, we can conclude that the inclusion of this measure helps to classify new objects, so it is an aspect to take into account in the last method that we will present in this paper.

The main advantage of this method is the inclusion of the α -T-credibility, which is useful for checking the reliability of the rules. In this way, the decision given by this method will be supported by a decision rule with a high α -T-credibility, which strengthens the decision to be made. However, a disadvantage of this method is that, as we have mentioned above, we have not taken into account the T support of the rules, which is also a relevant measure. In this way, the decision may be driven by a decision rule with a low representativeness in the decision table. This problem is addressed with the method proposed in the next section.

3.4. Support-Driven α -Credibility Method

In this section, we present the last proposed classification method. It is very similar to the one presented in Section 3.3, including the NT-support. The corresponding method is based on the following definition.

Definition 16. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $Dec(S)$ be a set of decision rules, $\alpha \in [0, 1)$, and $(\&, I)$ be an adjoint pair, where $\&$ is an associative conjunctor. For each new object $x \in U^c$ and $v \in V_d$, the degree of support-driven α -credibility of the decision $\bar{d}(x) = v$ according to $Dec(S)$ is given by the mapping $\delta_{Dec(S)}^\alpha : U^c \times V_d \rightarrow [0, 1]$ defined as

$$\delta_{Dec(S)}^\alpha(x, v) = \max\{\&(\|\Phi\|_S^T(x), \mu_{S,\tau}^\alpha(\Phi, \Psi), \text{supp}_S^{NT}(\Phi, \Psi)) \mid \Phi \rightarrow \Psi \in Dec(S), \Psi = (d, v)\}$$

We must emphasize that $0 \leq \delta_{Dec(S)}^\alpha(x, v) \leq 1$ for all new objects x , decisions $\bar{d}(x) = v$, $\alpha \in [0, 1)$ and sets of decision rules $Dec(S)$, so $\delta_{Dec(S)}^\alpha$ is a membership function of the new object on the decision value set V_d . The steps of the fourth proposed classification method, called *support-driven α -credibility method* or *D-method* for short, are the following:

- D1. First of all, given a new object to classify $x \in U^c$, we compute the degrees of support-driven α -credibility $\delta_{Dec(S)}^\alpha(x, v)$ for all $v \in V_d$.
- D2. After that, all the obtained degrees of support-driven α -credibility are compared. If there exists a decision $w \in V_d$ such that $\delta_{Dec(S)}^\alpha(x, w)$ is high and sufficiently different from $\delta_{Dec(S)}^\alpha(x, v)$, for each $v \in V_d \setminus \{w\}$, then the decision $\bar{d}(x) = w$ will be chosen for the object x .
- D3. If there exist some values of the decision attribute with similar degrees of support-driven α -credibility, then we will check the relationship between the variables involved in the computation. Then, another adjoint pair $(\&, I)$ from the considered multi-adjoint property framework can be chosen to represent more appropriately the relationship and to obtain a better aggregation of the corresponding values.

In the next example, we will use this method to classify the new objects considered in Table 4.

Example 5. Coming back to the environment of Example 2, the goal is to determine a decision for the objects x_8, x_9 , and x_{10} according to the support-driven α -credibility method.

- D1. We begin computing the degrees of support-driven 0.76 credibility of each decision for the new objects. Now, we suppose that there exists a close relationship between the considered variables, so we use the Gödel adjoint pair $(\&_G, I_G)$. Using Tables 3 and 5 and Definition 16 for the object x_8 and decision $\bar{d}(x_8) = 0$, we obtain that

$$\begin{aligned} \delta_{Dec(S)}^{0.76}(x_8, 0) &= \max\{\&_G(\|\Phi_1\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_1, \Psi_1), \text{supp}_S^{NT}(\Phi_1, \Psi_1)), \\ &\quad \&_G(\|\Phi_3\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3), \text{supp}_S^{NT}(\Phi_3, \Psi_3)), \\ &\quad \&_G(\|\Phi_5\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_5, \Psi_5), \text{supp}_S^{NT}(\Phi_5, \Psi_5)), \\ &\quad \&_G(\|\Phi_7\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_7, \Psi_7), \text{supp}_S^{NT}(\Phi_7, \Psi_7))\} \\ &= \max\{\&_G(0.35, 1, 0.85), \&_G(0.1, 1, 0.83), \\ &\quad \&_G(0.6, 0, 1), \&_G(0.85, 1, 0.89)\} \\ &= \max\{0.35, 0.1, 0, 0.85\} = 0.85 \end{aligned}$$

In a similar way, we compute the rest of the degrees of support-driven 0.76 credibility, whose results are the following ones:

$$\begin{aligned} \delta_{Dec(S)}^{0.76}(x_8, 0) &= \&_G(\|\Phi_7\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_7, \Psi_7), \text{supp}_S^{NT}(\Phi_7, \Psi_7)) = 0.85 \\ \delta_{Dec(S)}^{0.76}(x_8, 1) &= \&_G(\|\Phi_6\|_S^T(x_8), \mu_{S,\tau}^{0.76}(\Phi_6, \Psi_6), \text{supp}_S^{NT}(\Phi_6, \Psi_6)) = 0.52 \\ \delta_{Dec(S)}^{0.76}(x_9, 0) &= \&_G(\|\Phi_3\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3), \text{supp}_S^{NT}(\Phi_3, \Psi_3)) = 0.58 \\ \delta_{Dec(S)}^{0.76}(x_9, 1) &= \&_G(\|\Phi_4\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_4, \Psi_4), \text{supp}_S^{NT}(\Phi_4, \Psi_4)) = 0.62 \\ \delta_{Dec(S)}^{0.76}(x_{10}, 0) &= \&_G(\|\Phi_7\|_S^T(x_{10}), \mu_{S,\tau}^{0.76}(\Phi_7, \Psi_7), \text{supp}_S^{NT}(\Phi_7, \Psi_7)) = 0.69 \\ \delta_{Dec(S)}^{0.76}(x_{10}, 1) &= \&_G(\|\Phi_6\|_S^T(x_{10}), \mu_{S,\tau}^{0.76}(\Phi_6, \Psi_6), \text{supp}_S^{NT}(\Phi_6, \Psi_6)) = 0.78 \end{aligned}$$

- D2. First of all, notice that the values obtained for the object x_8 are the same as in Example 4. Consequently, the support-driven α -credibility method also recommends the decision $\bar{d}(x_8) = 0$ with a degree of support-driven 0.76 credibility of 0.85.

However, with respect to the object x_9 , the degrees of support-driven 0.76 credibility of both decisions are quite similar, making it difficult to decide about this object. In order to solve this,

we will use other adjoint pairs in the next step of the method that consider the three values in the computation.

Finally, the case of the object x_{10} is very similar to the one in Example 4, recommending once again the decision $\bar{d}(x_{10}) = 1$ with a degree of support-driven 0.76 credibility of 0.78.

- D3. Now, assuming that there is less of a relationship between the variables, we compute the degree of support-driven 0.76 credibility of the decisions $\bar{d}(x_9) = 0$ and $\bar{d}(x_9) = 1$ by using the adjoint pair $(\&_P, I_P)$, which multiplies the three values equitably. We obtained

$$\begin{aligned}\delta_{Dec(S)}^{0.76}(x_9, 0) &= \&_P(\|\Phi_3\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3), \text{supp}_S^{NT}(\Phi_3, \Psi_3)) = 0.48 \\ \delta_{Dec(S)}^{0.76}(x_9, 1) &= \&_P(\|\Phi_4\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_4, \Psi_4), \text{supp}_S^{NT}(\Phi_4, \Psi_4)) = 0.53\end{aligned}$$

We can observe similar difficulties in classifying the object x_9 . Finally, we suppose that there is almost no relationship between the considered variables. Hence, we consider the adjoint pair $(\&_L, I_L)$, obtaining

$$\begin{aligned}\delta_{Dec(S)}^{0.76}(x_9, 0) &= \&_L(\|\Phi_3\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_3, \Psi_3), \text{supp}_S^{NT}(\Phi_3, \Psi_3)) = 0.41 \\ \delta_{Dec(S)}^{0.76}(x_9, 1) &= \&_L(\|\Phi_4\|_S^T(x_9), \mu_{S,\tau}^{0.76}(\Phi_4, \Psi_4), \text{supp}_S^{NT}(\Phi_4, \Psi_4)) = 0.47\end{aligned}$$

We conclude that, independently of the adjoint pair considered in the multi-adjoint property framework $([0, 1], [0, 1], [0, 1], \&_L, \&_P, \&_G)$, it is difficult to make a decision for the object x_9 by using the support-driven α -credibility method. Nevertheless, since in all the cases we have obtained that $\delta_{Dec(S)}^{0.76}(x_9, 0) < \delta_{Dec(S)}^{0.76}(x_9, 1)$, the support-driven α -credibility method recommends the decision $\bar{d}(x_9) = 1$ for the object x_9 .

This difficult classification for the object x_9 has been caused by the inclusion of the NT support in Definition 16. As shown in Table 3, the NT support of the decision rules whose consequent is $(d, 1)$ is lower than the NT support of the rules with consequent $(d, 0)$.

We must emphasize that all the methods presented in this paper are important since each one approaches the decision-making process from a different point of view. Hence, they should be considered together in order to make decisions. In particular, the first one focuses on the degree of satisfiability with respect to the antecedents of the rules of the new object to classify in order to discard rules whose antecedents are very different from the new object. With respect to the ε -representativeness method, it is presented as an alternative to the suitability method to solve some situations in which this method is unable to provide a satisfactory decision. For that purpose, the ε -representativeness method also considers the T -support of the decision rules, given that it shows the representativeness of the rule in the decision table. On the other hand, the α -credibility method is proposed since the credibility of the decision rules has been considered in other studies [20,25] to classify new objects. The main novelty of our contribution with this third method is the consideration of the multi-adjoint framework. In addition, the support-driven α -credibility method is based on the third one but also considers the NT support of the rules compared with the new object for the same reasons as the ε -representativeness method was introduced. Furthermore, it is also important to remark again that these classification methods are not mutually exclusive but can complement each other, as we have mentioned above, to classify objects when one of the methods does not give a sufficiently conclusive result.

Finally, Table 6 presents a summary of the different notions considered in the methods proposed in this paper. In this table, “satisfiability” represents whether the method takes into account in the computation the degree to which the new object satisfies the given decision rules. It is clear that this degree is very relevant to classifying new objects, and so it is considered in every method.

Table 6. Summary of the notions considered in each method.

Method	Satisfiability	Support	Credibility
Suitability	X		
ε -representativeness	X	X	
α -credibility	X		X
Support-driven α -credibility	X	X	X

4. Comparison with Other Methods

In the introduction section, the differences between the introduced methods and the ones given in [20,21,23] were enumerated. This section will be focused on the methods introduced in [24,25], whose comparison is not trivial.

A detailed study on the classification of new data is carried out in [25]. The authors sequentially present several notions, such as the upper and lower approximations of sets of attributes or the degree of credibility of a decision rule. They play an indispensable role in the classification method proposed, which is given as an algorithm. Since we have detected some similarities between our α -credibility method and the one given in [25], we have considered it important to compare both studies.

First of all, in [25], the following specific fuzzy tolerance relation is considered to compare pairs of objects according to each attribute $a \in \mathcal{A}_d$:

$$R_a(x, y) = \frac{\max\{0, \min\{\bar{a}(x), \bar{a}(y)\} + k_a - \max\{\bar{a}(x), \bar{a}(y)\}\}}{k_a}$$

where $k_a > 0$ for all $a \in \mathcal{A}_d$. Notice that it is also a separable relation. In addition, if $k_a = 1$ and $V_a = [0, 1]$ for each $a \in \mathcal{A}_d$, then the expression of the relation R_a coincides with one of the mappings T_a considered in Example 1 since

$$\begin{aligned} R_a(x, y) &= \frac{\max\{0, \min\{\bar{a}(x), \bar{a}(y)\} + 1 - \max\{\bar{a}(x), \bar{a}(y)\}\}}{1} \\ &= \min\{\bar{a}(x), \bar{a}(y)\} + 1 - \max\{\bar{a}(x), \bar{a}(y)\} \\ &= 1 - (\max\{\bar{a}(x), \bar{a}(y)\} - \min\{\bar{a}(x), \bar{a}(y)\}) \\ &= 1 - |\bar{a}(x) - \bar{a}(y)| \\ &= T_a(\bar{a}(x), \bar{a}(y)) \end{aligned}$$

Moreover, in [25], the relation between objects according to the total set of attributes $\mathcal{A}$ is defined for each $x, y \in U$ as

$$R_{\mathcal{A}}(x, y) = \min\{R_a(x, y) \mid a \in \mathcal{A}\}$$

Notice that $R_{\mathcal{A}}$ coincides with the degree of satisfiability to a conjunction of formulas of Definition 8. Indeed, given $\mathcal{A} = \{a_1, \dots, a_n\}$ and $\Phi = \bigwedge_{i=1}^n (a_i, v_i) \in \text{For}(\mathcal{A})$ induced by an object $y \in U$, that is, $v_i = \bar{a}_i(y)$ for all $i \in \{1, \dots, n\}$, by Definition 8, we have

$$\begin{aligned} \|\Phi\|_S^T(x) &= \left\| \bigwedge_{i=1}^n (a_i, v_i) \right\|_S^T(x) = \min\{\|(a_i, v_i)\|_S^T(x) \mid a_i \in \mathcal{A}\} \\ &= \min\{\|(a_i, \bar{a}_i(y))\|_S^T(x) \mid a_i \in \mathcal{A}\} = \min\{T_{a_i}(\bar{a}_i(x), \bar{a}_i(y)) \mid a_i \in \mathcal{A}\} \\ &= \min\{R_{a_i}(x, y) \mid a_i \in \mathcal{A}\} \\ &= R_{\mathcal{A}}(x, y) \end{aligned}$$

This relation is used to compare each new object with the antecedents of the rules of the decision table, as we have considered in Section 3.1. In order to carry out the classification of new objects, [25] also takes into account the credibility of the decision rule $r_i : \Phi_i \rightarrow \Psi_i$

compared with the new object and whose antecedent is composed of all the condition attributes of the table, defined as

$$\mu_{\mathcal{A}}(\Phi_i, \Psi_i) = \min\{\max\{1 - R_{\mathcal{A}}(x, x_i), \mu_{\bar{d}(x_i)\mathcal{A}}(x)\} \mid x \in R_{\mathcal{A}}(x_i)\}$$

where $x_i \in U$ is the object that induces the rule r_i and

$$\begin{aligned}\mu_{\bar{d}(x_i)\mathcal{A}}(x) &= \min\{\max\{1 - R_{\mathcal{A}}(x, y), R_d(x_i, y)\} \mid y \in R_{\mathcal{A}}(x)\} \\ R_{\mathcal{A}}(x_i) &= \{x \in U \mid R_{\mathcal{A}}(x, x_i) > 0\}\end{aligned}$$

In our terms, these notions are given as

$$\begin{aligned}\mu_{S,\tau}^0(\Phi_i, \Psi_i) &= \min\{\max\{1 - \|\Phi_i\|_S^T(x), \mu_{\Psi_{i,S,\tau}}(x)\} \mid x \in (\|\Phi_i\|_S^T)_0\} \\ \mu_{\Psi_{i,S,\tau}}(x) &= \min\{\max\{1 - \|\Phi\|_S^T(y), \|\Psi_i\|_S^T(y)\} \mid y \in (\|\Phi\|_S^T)_0\} \\ (\|\Phi_i\|_S^T)_0 &= \{x \in U \mid \|\Phi_i\|_S^T(x) > 0\}\end{aligned}$$

where $\Phi \rightarrow \Psi \in Dec(S)$ is the decision rule induced by x considering all the condition attributes for the antecedent. Notice that $\mu_{\mathcal{A}}(\Phi_i, \Psi_i)$ can be expressed as

$$\mu_{S,\tau}^0(\Phi_i, \Psi_i) = \&_G(I(\|\Phi_i\|_S^T(x), \&_G(I(\|\Phi\|_S^T(y), \|\Psi_i\|_S^T(y)))))$$

where $x \in (\|\Phi_i\|_S^T)_0$, $y \in (\|\Phi\|_S^T)_0$, $\&_G$ is the Gödel t-norm and the I is implication residuated of the conjunctive $\&$, which are both defined for each $x, y \in [0, 1]$ as

$$\begin{aligned}\&(x, y) &= \begin{cases} 0 & \text{if } x \leq 1 - y \\ x & \text{otherwise} \end{cases} \\ I(x, y) &= \max\{1 - x, y\}\end{aligned}$$

On the other hand, it is also possible to obtain a relationship between the credibility and Definition 4. First of all, notice that

$$\&_G(\{I(R_{\mathcal{A}}(x, y), R_d(x_i, y)) \mid y \in R_{\mathcal{A}}(x)\}) = \&_G(\{I(R_{\mathcal{A}}(x, y), R_d(x_i, y)) \mid y \in U\})$$

holds because if $y \notin R_{\mathcal{A}}(x)$ then

$$\begin{aligned}I(R_{\mathcal{A}}(x, y), R_d(x_i, y)) &= \max\{1 - R_{\mathcal{A}}(x, y), R_d(x_i, y)\} \\ &= \max\{1 - 0, R_d(x_i, y)\} \\ &= 1\end{aligned}$$

and $\&_G(v, 1) = v$ for each $v \in [0, 1]$ due to the boundary property satisfied by the t-norms. As a consequence, we deduce that

$$\begin{aligned}\mu_{\bar{d}(x_i)\mathcal{A}}(x) &= \&_G(\{I(R_{\mathcal{A}}(x, y), R_d(x_i, y)) \mid y \in R_{\mathcal{A}}(x)\}) \\ &= \min\{I(R_{\mathcal{A}}(x, y), R_d(x_i, y)) \mid y \in U\} \\ &= (R_d x_i)^{\downarrow N}(x) \\ \mu_{\bar{d}(x_i)\mathcal{A}}(x_i) &= (R_d x_i)^{\downarrow N}(x_i) = \text{POS}_{\mathcal{A}}^f(x_i) \\ \mu_{\mathcal{A}}(\Phi_i, \Psi_i) &= \&_G(\{I(R_{\mathcal{A}}(x, x_i), \mu_{\bar{d}(x_i)\mathcal{A}}(x)) \mid x \in R_{\mathcal{A}}(x_i)\}) \\ &= \min\{I(R_{\mathcal{A}}(x, x_i), \mu_{\bar{d}(x_i)\mathcal{A}}(x)) \mid x \in U\} \\ &= (\mu_{\bar{d}(x_i)\mathcal{A}})^{\downarrow N}(x_i)\end{aligned}$$

In conclusion, we must emphasize that the credibility $\mu_{\mathcal{A}}(\Phi_i, \Psi_i)$ is similar to the one given in Definition 12. However, in that definition, we have considered different

adjoint pairs for each pair of objects of the decision table, extending the definition given in [25] and allowing preferences to be set among the objects. Moreover, in the definition of $(\|\Phi\|_S^T)_\alpha$, we have set an α -cut $\|\Phi\|_S^T(x_i) > \alpha$ that can be changed depending on the study, providing a greater level of flexibility, while in [25], the threshold is $\|\Phi\|_S^T(x_i) > 0$. This main difference is caused by the definition of the tolerance relations used to compare the objects. We have considered tolerance relations T_a in general, while in [25] a specific one R_a is provided, in which a value k_a is included to distinguish objects. This choice is justified by the fact that, if $k_a \leq |\bar{a}(x) - \bar{a}(y)|$, then it is assumed that the objects x, y are not related according to a , obtaining $R_a(x, y) = 0$. Nevertheless, in this paper, we assume that a pair of objects x, y are not related if and only if $|\bar{a}(x) - \bar{a}(y)| = 1$, which is the motivation for the introduction of the α -cut. Despite this difference between both approaches, it is possible to relate them, so that $y \in R_A(x)$ if and only if $y \in (\|\Phi\|_S^T)_\alpha$, where $\Phi \rightarrow \Psi$ is the decision rule induced by x with $\Phi \in \text{For}(\mathcal{A})$. Indeed, defining $\alpha = 1 - \min\{k_a \mid a \in \mathcal{A}\}$ when $T_a(\bar{a}(x), \bar{a}(y)) = 1 - |\bar{a}(x) - \bar{a}(y)|$, we deduce that, if $\|\Phi\|_S^T(y) > \alpha$, then from Definition 8, $\min\{T_a(\bar{a}(x), \bar{a}(y)) \mid a \in \mathcal{A}\} > \alpha$. Hence, $|\bar{a}(x) - \bar{a}(y)| < 1 - \alpha = \min\{k_a \mid a \in \mathcal{A}\}$ for all $a \in \mathcal{A}$. In conclusion, the objects x, y are related in both environments.

Finally, the decision $d(x) = v$ to be made for the new object x according to [25] is given as

$$v = \arg \max \{ \min \{ \|\Phi\|_S^T(x), \mu_A(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in \text{Dec}(S) \} \mid v_d \in V_d, \Psi = (d, v_d) \}$$

where $\arg \max$ is the argument of the maximum of a function; that is, given two sets X, Y and a mapping $f: X \rightarrow Y$, the *argument of the maximum of f in X* is defined as $\arg \max f = \{x \in X \mid f(y) \leq f(x) \text{ for all } y \in X\}$. Notice that the α -credibility method extends this one since Definition 12 generalizes the credibility used in [25] due to the use of adjoint pairs, the consideration of a family of separable $[0, 1]$ -fuzzy tolerance relations to define the degree of satisfiability of a formula, and the replacement of the minimum with an adjoint conjunct in general.

With respect to the comparison between the α -credibility method and the given in [24], there are also some similarities. On the one hand, Definition 12 generalizes the credibility presented in [24], since we can consider different adjoint pairs depending on the objects compared, but they consider a t-norm and an implication for all the objects of the table. On the other hand, with the classification method, they consider a t-conorm to obtain the values to be compared to make a decision. This choice is due to the fact that they make a preprocess to discard those decision rules whose credibility or similarity of their antecedents to the new object studied is low. In this paper, we have considered the maximum from the beginning in order to avoid the scenario in which the user has to make a preprocess in which a threshold has to be fixed. Furthermore, they determine the best possible decision for the new object by means of the maximum. However, we believe that it is essential to compare all the values obtained so that the decision to be made is not only inferred from the maximum, but its value must be sufficiently different from the others obtained to provide a solid classification. Otherwise, as stated in the third step of all the methods presented, we must continue applying the corresponding method.

Thus, the introduced methods generalize the given one in the literature, such as [24,25], as we have detailed above, and [20,21,23], which were commented on in the introduction.

5. A Toy Example

This section is devoted to studying the classification of a new object in a real case following the different strategies presented in this paper. In particular, the considered data set concerns eight project evaluations described using the attributes construction expense (CE), financial income (FI), strategy benefit (SB), and external influence (EI) in order to decide whether to make an investment (I), a delayed investment (DI) or no investment (NI) in each project. This data set was introduced in [30] and is represented in Table 7, where every unit corresponds to 10.000 yuan.

Table 7. Knowledge system on water conservancy project investment decision making.

	CE	FI	SB	EI	Decision (d)
Project 1 (p_1)	500	200	690	150	DI
Project 2 (p_2)	700	470	650	440	NI
Project 3 (p_3)	300	410	500	280	NI
Project 4 (p_4)	200	455	550	290	I
Project 5 (p_5)	250	260	490	105	DI
Project 6 (p_6)	510	380	430	130	DI
Project 7 (p_7)	350	550	255	145	I
Project 8 (p_8)	650	600	570	120	NI

This data set was analyzed from an RST perspective in [30] and was studied in the fuzzy framework in [13]. Now, we use that study to obtain the decision rules considered to decide on the investment in a new project. First of all, the previous knowledge system is formally represented as the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, where the set of objects is $U = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\}$, the set of attributes is $\mathcal{A} = \{\text{CE}, \text{FI}, \text{SB}, \text{EI}\}$, $\mathcal{V}_a = [0, 1]$, for all $a \in \mathcal{A}$, $\mathcal{V}_d = \{\text{I}, \text{NI}, \text{DI}\}$, and $\overline{\mathcal{A}_d}$ is represented in Table 8. Notice that Table 8 provides a normalization of the values in Table 7, taking into account the possible range of values for each condition attribute, which we have considered to be 1000 for the attributes CE, FI, and SB and 500 for the attribute EI.

Table 8. Decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ extracted from Table 7.

	CE	FI	SB	EI	d
p_1	0.5	0.2	0.69	0.3	DI
p_2	0.7	0.47	0.65	0.88	NI
p_3	0.3	0.41	0.5	0.56	NI
p_4	0.2	0.455	0.55	0.58	I
p_5	0.25	0.26	0.49	0.21	DI
p_6	0.51	0.38	0.43	0.26	DI
p_7	0.35	0.55	0.255	0.29	I
p_8	0.65	0.6	0.57	0.24	NI

The decision rules $r_i: \Phi_i \rightarrow \Psi_i$ for each $i \in \{1, \dots, 8\}$ used to classify the new project are the following:

$$\begin{aligned}
 r_1 &: (\text{CE}, 0.5) \wedge (\text{FI}, 0.2) \wedge (\text{SB}, 0.69) \wedge (\text{EI}, 0.3) \rightarrow (d, \text{DI}) \\
 r_2 &: (\text{CE}, 0.7) \wedge (\text{FI}, 0.47) \wedge (\text{SB}, 0.65) \wedge (\text{EI}, 0.88) \rightarrow (d, \text{NI}) \\
 r_3 &: (\text{CE}, 0.3) \wedge (\text{FI}, 0.41) \wedge (\text{SB}, 0.5) \wedge (\text{EI}, 0.56) \rightarrow (d, \text{NI}) \\
 r_4 &: (\text{CE}, 0.2) \wedge (\text{FI}, 0.455) \wedge (\text{SB}, 0.55) \wedge (\text{EI}, 0.58) \rightarrow (d, \text{I}) \\
 r_5 &: (\text{CE}, 0.25) \wedge (\text{FI}, 0.26) \wedge (\text{SB}, 0.49) \wedge (\text{EI}, 0.21) \rightarrow (d, \text{DI}) \\
 r_6 &: (\text{CE}, 0.51) \wedge (\text{FI}, 0.38) \wedge (\text{SB}, 0.43) \wedge (\text{EI}, 0.26) \rightarrow (d, \text{DI}) \\
 r_7 &: (\text{CE}, 0.35) \wedge (\text{FI}, 0.55) \wedge (\text{SB}, 0.255) \wedge (\text{EI}, 0.29) \rightarrow (d, \text{I}) \\
 r_8 &: (\text{CE}, 0.65) \wedge (\text{FI}, 0.6) \wedge (\text{SB}, 0.57) \wedge (\text{EI}, 0.24) \rightarrow (d, \text{NI})
 \end{aligned}$$

We denote this set of decision rules as $\text{Dec}(S)$. In order to compute the measures of Definition 11 for each rule of $\text{Dec}(S)$, we consider the family of separable $[0, 1]$ -fuzzy

tolerance relations $T = \{T_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ defined as $T_a(v, w) = 1 - |v - w|$, for all $a \in \mathcal{A}$, $v, w \in [0, 1]$ and

$$T_d(v, w) = \begin{cases} 1 & \text{if } v = w \\ 0.5 & \text{if } v = \text{DI and } w \neq \text{DI} \\ 0 & \text{otherwise} \end{cases}$$

since we consider that DI is an intermediate decision between I and NI. For this study, we will consider the multi-adjoint property-oriented frame $([0, 1], [0, 1], [0, 1], \&_L, \&_P, \&_G)$ and the context $(U, U, R_{\mathcal{A}}, \tau)$, where $\tau: U \times U \rightarrow \{L, P, G\}$ is defined for each $x, y \in U$ as

$$\tau(x, y) = \begin{cases} L & \text{if } \bar{d}(x) = \bar{d}(y) \\ P & \text{if } \bar{d}(x) = \text{DI and } \bar{d}(y) \neq \text{DI} \\ G & \text{otherwise} \end{cases}$$

As in Example 1, we have considered this mapping to incorporate some preferences among the objects. Specifically, we have also reinforced the computation of the credibility when the objects have the same decision (implication I_L is greater than I_P and I_G). Moreover, we will assign more relevance to the objects with the intermediate decision DI (implication I_P is greater than I_G).

Furthermore, we will use the adjoint pair $(\&_L, I_L)$, and we will fix the α cut to be equal to 0.79 in order to compute the 0.79- T -credibility of each decision rule belonging to $\text{Dec}(S)$ to reduce the number of inconsistencies in the sets $(\|\Phi\|_S^T)_\alpha$ with $\Phi \rightarrow \Psi \in \text{Dec}(S)$, while some flexibility is allowed in the comparisons.

Now, the T -support, NT -support, and 0.79- T -credibility of each decision rule are computed by using Definitions 11 and 12. The corresponding results are shown in Table 9.

Table 9. T support, NT support, and 0.79- T -credibility of each rule in $\text{Dec}(S)$.

Rule	supp_S^T	supp_S^{NT}	$\mu_{S,\tau}^{0.79}$
r_1	4.91	1	1
r_2	3.09	0.63	1
r_3	3.75	0.76	0
r_4	3.21	0.65	0
r_5	4.82	0.98	1
r_6	4.86	0.99	0.37
r_7	3.21	0.65	0.37
r_8	3.51	0.71	1

As in Example 1, the decision rules $\Phi \rightarrow \Psi \in \text{Dec}(S)$ with 0.79 T -credibility equal to 1 satisfy $(\|\Phi\|_S^T)_{0.79} = \{x\}$ where $x \in U$, the objects that induce them. Analogously, $\mu_{S,\tau}^{0.79}(\Phi_3, \Psi_3) = \mu_{S,\tau}^{0.79}(\Phi_4, \Psi_4) = 0$ since $\|\Phi_3\|_S^T(x_4) = \|\Phi_4\|_S^T(x_3) = 0.9 > 0.79$ and $\|\Psi_3\|_S^T(x_4) = \|\Psi_4\|_S^T(x_3) = 0$. The case of r_6 and r_7 is similar, but $\|\Phi_6\|_S^T(x_7) = \|\Phi_7\|_S^T(x_6) = 0.83$, whereas $\|\Psi_6\|_S^T(x_7) = \|\Psi_7\|_S^T(x_6) = 0.5$; that is, both rules are less related in the antecedent and more related in the consequent than the rules r_3 and r_4 , making the 0.79 T -credibility higher.

Now, we use these rules and each method presented in this paper to classify the new project p_9 given in Table 10.

Table 10. New project p_9 to evaluate.

	CE	FI	SB	EI
p_9	500	300	800	300

As can be seen in Table 10, the project p_9 presents a low financial income, a high strategy benefit, and intermediate construction expense and external influence. In order to decide whether or not to invest in it, we normalize these values according to the criteria established above. In Table 11, the project p_9 is shown together with the rest of the projects.

Table 11. Table 8 together with the new project to evaluate p_9 .

	CE	FI	SB	EI	d
p_1	0.5	0.2	0.69	0.3	DI
p_2	0.7	0.47	0.65	0.88	NI
p_3	0.3	0.41	0.5	0.56	NI
p_4	0.2	0.455	0.55	0.58	I
p_5	0.25	0.26	0.49	0.21	DI
p_6	0.51	0.38	0.43	0.26	DI
p_7	0.35	0.55	0.255	0.29	I
p_8	0.65	0.6	0.57	0.24	NI
p_9	0.5	0.3	0.8	0.6	

We begin computing the degree of satisfiability for each antecedent of the decision rules of $Dec(S)$ of the new project p_9 . These results are shown in Table 12.

Table 12. Degree of satisfiability for each antecedent of decision rules of the new project p_9 .

	$\ \Phi_1\ _S^T$	$\ \Phi_2\ _S^T$	$\ \Phi_3\ _S^T$	$\ \Phi_4\ _S^T$	$\ \Phi_5\ _S^T$	$\ \Phi_6\ _S^T$	$\ \Phi_7\ _S^T$	$\ \Phi_8\ _S^T$
p_9	0.7	0.72	0.7	0.7	0.61	0.63	0.455	0.64

Notice that, by Definitions 13–16, the degree of satisfiability to antecedents for the new object plays a key role in each of the presented methods. Taking into account that, from Table 12, most degrees of satisfiability are quite similar, we can anticipate relevant difficulties to provide an accurate classification for the project p_9 .

Next, we are going to apply the classification methods presented in Section 3 to determine the best possible decision for the project p_9 according to the set of decision rules $Dec(S)$.

5.1. Application of the Suitability Method

We begin studying the new project p_9 by means of the suitability method.

- S1. First of all, we compute the degree of suitability of each possible decision to be made. Thus, from Definition 13, we obtain that

$$\begin{aligned}\sigma_{Dec(S)}(p_9, I) &= \|\Phi_4\|_S^T(p_9) = 0.7 \\ \sigma_{Dec(S)}(p_9, NI) &= \|\Phi_2\|_S^T(p_9) = 0.72 \\ \sigma_{Dec(S)}(p_9, DI) &= \|\Phi_1\|_S^T(p_9) = 0.7\end{aligned}$$

- S2. We can immediately check that all the degrees of suitability are very similar. As a result, we cannot make a decision for the project p_9 . To solve that, we apply the third step of the method.

- S3. Since many degrees of satisfiability are very similar, we will consider all of them in the computations by setting the threshold $\beta = 0.4$. Consequently, we obtain $\mathcal{F}_{p_9}^{0.4}(\text{I}) = \{\Phi_4, \Phi_7\}$, $\mathcal{F}_{p_9}^{0.4}(\text{NI}) = \{\Phi_2, \Phi_3, \Phi_8\}$ and $\mathcal{F}_{p_9}^{0.4}(\text{DI}) = \{\Phi_1, \Phi_5, \Phi_6\}$. Therefore, it is

$$\begin{aligned}\mathcal{M}_{\mathcal{F}_{p_9}^{0.4}(\text{I})} &= \frac{\|\Phi_4\|_S^T(p_9) + \|\Phi_7\|_S^T(p_9)}{2} = \frac{0.7 + 0.455}{2} \\ &= 0.58 \\ \mathcal{M}_{\mathcal{F}_{p_9}^{0.4}(\text{NI})} &= \frac{\|\Phi_2\|_S^T(p_9) + \|\Phi_3\|_S^T(p_9) + \|\Phi_8\|_S^T(p_9)}{3} = \frac{0.72 + 0.7 + 0.64}{3} \\ &= 0.69 \\ \mathcal{M}_{\mathcal{F}_{p_9}^{0.4}(\text{DI})} &= \frac{\|\Phi_1\|_S^T(p_9) + \|\Phi_5\|_S^T(p_9) + \|\Phi_6\|_S^T(p_9)}{3} = \frac{0.7 + 0.61 + 0.63}{3} \\ &= 0.65\end{aligned}$$

Notice that the values obtained for the decisions $\vec{d}(p_9) = \text{NI}$ and $\vec{d}(p_9) = \text{DI}$ are quite similar. As a consequence, we cannot ensure which is the best decision for the new project according to the suitability method, resulting in the slight recommendation not to invest in it. In any case, we can discard the decision $\vec{d}(p_9) = \text{I}$ because the value obtained for it is significantly lower than for the other decisions.

Since we were unable to make a decision for project p_9 according to the suitability method, we must address this classification problem by using other methods proposed in this paper.

5.2. Application of the ε -Representativeness Method

Now, we will try to make a satisfactory decision for the project p_9 by applying the ε -representativeness method.

- R1. Once again, we set the threshold $\varepsilon = 0.7$ for the reasons presented in Example 3. Hence, by Definition 14 and Proposition 1, and taking into account Tables 9 and 12, we have that

$$\begin{aligned}\rho_{Dec(S)}^{0.7}(p_9, \text{I}) &= \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | 0.7 \leq \|\Phi\|_S^T(p_9), \Psi = (d, \text{I})\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | 0.7 \leq \|\Phi\|_S^T(p_9)\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{\|\Phi_4\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi_4, \Psi_4)}{\sum_{\Phi \rightarrow \Psi \in \{\Phi_1 \rightarrow \Psi_1, \Phi_2 \rightarrow \Psi_2, \Phi_3 \rightarrow \Psi_3, \Phi_4 \rightarrow \Psi_4\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{0.7 \times 3.21}{0.7 \times 4.91 + 0.72 \times 3.09 + 0.7 \times 3.75 + 0.7 \times 3.21} \\ &= \frac{2.24}{10.53} = 0.21 \\ \rho_{Dec(S)}^{0.7}(p_9, \text{NI}) &= \frac{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | 0.7 \leq \|\Phi\|_S^T(p_9), \Psi = (d, \text{NI})\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)}{\sum_{\{\Phi \rightarrow \Psi \in Dec(S) | 0.7 \leq \|\Phi\|_S^T(p_9)\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{\|\Phi_2\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi_2, \Psi_2) + \|\Phi_3\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi_3, \Psi_3)}{\sum_{\Phi \rightarrow \Psi \in \{\Phi_1 \rightarrow \Psi_1, \Phi_2 \rightarrow \Psi_2, \Phi_3 \rightarrow \Psi_3, \Phi_4 \rightarrow \Psi_4\}} \|\Phi\|_S^T(p_9) \cdot \text{supp}_S^T(\Phi, \Psi)} \\ &= \frac{0.72 \times 3.09 + 0.7 \times 3.75}{0.7 \times 4.91 + 0.72 \times 3.09 + 0.7 \times 3.75 + 0.7 \times 3.21} \\ &= \frac{4.85}{10.53} = 0.46 \\ \rho_{Dec(S)}^{0.7}(p_9, \text{DI}) &= 1 - \rho_{Dec(S)}^{0.7}(p_9, \text{I}) - \rho_{Dec(S)}^{0.7}(p_9, \text{NI}) \\ &= 1 - 0.21 - 0.46 = 0.33\end{aligned}$$

- R2. Since relevance differences exist between all the degrees of 0.7 representativeness, we conclude that this method recommends the decision $\bar{d}(p_9) = \text{NI}$ with a degree of 0.7 representativeness of 0.46; that is, it is suggested that this new project is not invested in. Therefore, the ε -representativeness method solves some situations in which the first one is not conclusive enough for the decision to be taken. Notice that, by definition, the values are “complementary”, so they are usually low but should be different.

Moreover, we can extract some more conclusions. From Lemma 1, we obtain the result that if the value ε satisfies

$$\max\{\sigma_{Dec(S)}(p_9, v) \mid v \in V_d \setminus \{\text{NI}\}\} = 0.7 < \varepsilon \leq \sigma_{Dec(S)}(p_9, \text{NI}) = 0.72$$

then $\rho_{Dec(S)}^\varepsilon(p_9, \text{I}) = \rho_{Dec(S)}^\varepsilon(p_9, \text{DI}) = 0$ and $\rho_{Dec(S)}^\varepsilon(p_9, \text{NI}) = 1$. Hence, few values of the threshold ε provide a clear classification for the project p_9 , showing the existing difficulty of deciding about it explained above. In addition, from Proposition 1 the project p_9 is ε -undecidable if ε satisfies

$$\max\{\sigma_{Dec(S)}(p_9, v) \mid v \in \{\text{I}, \text{NI}, \text{DI}\}\} = \max\{\sigma_{Dec(S)}(p_9, \text{NI}) < \varepsilon$$

Finally, it can be checked that

$$\begin{aligned} \rho_{Dec(S)}^\varepsilon(p_9, \text{I}) < \rho_{Dec(S)}^\varepsilon(p_9, \text{NI}) < \rho_{Dec(S)}^\varepsilon(p_9, \text{DI}) & \quad \text{if } 0 \leq \varepsilon \leq 0.61 \\ \rho_{Dec(S)}^\varepsilon(p_9, \text{I}) < \rho_{Dec(S)}^\varepsilon(p_9, \text{DI}) < \rho_{Dec(S)}^\varepsilon(p_9, \text{NI}) & \quad \text{if } 0.61 < \varepsilon \leq 0.72 \end{aligned}$$

Therefore, given that the study is more accurate when ε is higher, these inequalities strengthen the idea that the best decision for the project p_9 is $\bar{d}(p_9) = \text{NI}$.

Although we have provided a solid decision for project p_9 , we consider it important to continue testing the other two methods to identify potential benefits and drawbacks.

5.3. Application of the α -Credibility Method

Next, the α -credibility method will be applied.

- C1. We assume the variables are related somehow. Hence, we consider the adjoint pair $(\&_P, I_P)$ to compute the degree of 0.79 credibility of each possible decision. In this way, from Definition 15 and Tables 9 and 12, for the decision $\bar{d}(p_9) = \text{I}$, it is obtained that

$$\begin{aligned} \theta_{Dec(S)}^{0.79}(p_9, \text{I}) &= \max\{\&_P(\|\Phi_4\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_4, \Psi_4)), \\ &\quad \&_P(\|\Phi_7\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_7, \Psi_7))\} \\ &= \max\{\&_P(0.7, 0), \&_P(0.455, 0.37)\} \\ &= 0.17 \end{aligned}$$

In a similar way, the degree of 0.79 credibility of the decisions $\bar{d}(p_9) = \text{NI}$ and $\bar{d}(p_9) = \text{DI}$ are computed.

$$\begin{aligned} \theta_{Dec(S)}^{0.79}(p_9, \text{NI}) &= \&_P(\|\Phi_2\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_2, \Psi_2)) = 0.72 \\ \theta_{Dec(S)}^{0.79}(p_9, \text{DI}) &= \&_P(\|\Phi_1\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_1, \Psi_1)) = 0.7 \end{aligned}$$

- C2. As a result, since $\theta_{Dec(S)}^{0.79}(p_9, \text{NI})$ and $\theta_{Dec(S)}^{0.79}(p_9, \text{DI})$ are very similar, we cannot make any decision for the project p_9 according to the α -credibility method. Moreover, it can be shown that the same degrees of 0.79 credibility of both decisions are obtained if the adjoint pairs $(\&_G, I_G)$ and $(\&_L, I_L)$ are considered, since from Table 3, $\mu_{S,\tau}^{0.79}(\Phi_2, \Psi_2) = \mu_{S,\tau}^{0.79}(\Phi_1, \Psi_1) = 1$. In conclusion, following this method, we can only discard the decision $\bar{d}(p_9) = \text{I}$ since its degree of 0.79 credibility is much lower than the others.

Consequently, we cannot make a decision for project p_9 by using the α -credibility method due to the existing similarity between the degrees of α -credibility of the two decisions. Therefore, a further study is needed.

5.4. Application of Support-Driven α -Credibility Method

Finally, we conclude the analysis of the classification of the new project p_9 by applying the support-driven α -credibility method.

- D1. First of all, we compute the degree of support-driven 0.79 credibility of the decisions $\bar{d}(p_9) = I$, $\bar{d}(p_9) = NI$ and $\bar{d}(p_9) = DI$. As in Example 5, we first suppose that the variables are related. Therefore, we will use the adjoint pair $(\&_G, I_G)$. Consequently, from Definition 16 and Tables 9 and 12, we obtain for the decision $\bar{d}(p_9) = I$ that

$$\begin{aligned}\delta_{Dec(S)}^{0.79}(p_9, I) &= \max\{\&_G(\|\Phi_4\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_4, \Psi_4), supp_S^{NT}(\Phi_4, \Psi_4)), \\ &\quad \&_G(\|\Phi_7\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_7, \Psi_7), supp_S^{NT}(\Phi_7, \Psi_7))\} \\ &= \max\{\&_G(0.7, 0, 0.65), \&_G(0.455, 0.37, 0.65)\} \\ &= 0.37\end{aligned}$$

Analogously, we compute the degree of support-driven 0.79 credibility of the decisions $\bar{d}(p_9) = NI$ and $\bar{d}(p_9) = DI$, which are the following:

$$\begin{aligned}\delta_{Dec(S)}^{0.79}(p_9, NI) &= \&_G(\|\Phi_8\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_8, \Psi_8), supp_S^{NT}(\Phi_8, \Psi_8)) = 0.64 \\ \delta_{Dec(S)}^{0.79}(p_9, DI) &= \&_G(\|\Phi_1\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_1, \Psi_1), supp_S^{NT}(\Phi_1, \Psi_1)) = 0.7\end{aligned}$$

- D2. Following the support-driven α -credibility method, although the decision to be made for the project p_9 is not very clear, the suggested decision contradicts the previous methods, since $\delta_{Dec(S)}^{0.79}(p_9, NI) < \delta_{Dec(S)}^{0.79}(p_9, DI)$, meaning that a delayed investment for the project p_9 is now recommended. In order to examine this change in more detail, we will assume that the variables are less related. As a consequence, we will consider another adjoint pair taking into account all the input values in the computation to obtain the degree of support-driven 0.79 credibility of these two decisions.
- D3. Now, following the same procedure as in Example 5, we suppose that the variables are partially related. With this purpose, we use the adjoint pair $(\&_P, I_P)$, obtaining

$$\begin{aligned}\delta_{Dec(S)}^{0.79}(p_9, NI) &= \&_P(\|\Phi_8\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_8, \Psi_8), supp_S^{NT}(\Phi_8, \Psi_8)) = 0.45 \\ \delta_{Dec(S)}^{0.79}(p_9, DI) &= \&_P(\|\Phi_1\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_1, \Psi_1), supp_S^{NT}(\Phi_1, \Psi_1)) = 0.7\end{aligned}$$

Now, the difference between $\delta_{Dec(S)}^{0.79}(p_9, NI)$ and $\delta_{Dec(S)}^{0.79}(p_9, DI)$ has increased notably, allowing us to conclude that the support-driven α -credibility method suggests the decision $\bar{d}(p_9) = DI$ for the project p_9 with a degree of support-driven 0.79 credibility of 0.7. Nevertheless, for a more detailed study, we will also study the classification of project p_9 by means of the adjoint pair $(\&_L, I_L)$, assuming that the variables are very independent. Thus, we obtain

$$\begin{aligned}\delta_{Dec(S)}^{0.79}(p_9, NI) &= \&_L(\|\Phi_8\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_8, \Psi_8), supp_S^{NT}(\Phi_8, \Psi_8)) = 0.35 \\ \delta_{Dec(S)}^{0.79}(p_9, DI) &= \&_L(\|\Phi_1\|_S^T(p_9), \mu_{S,\tau}^{0.79}(\Phi_1, \Psi_1), supp_S^{NT}(\Phi_1, \Psi_1)) = 0.7\end{aligned}$$

The difference between both degrees of support-driven 0.79 credibility has increased even more so that the decision $\bar{d}(p_9) = DI$ is the best to be made for the project p_9 . This fact is due to the fact that decision rules with higher NT support take the decision DI in the consequent due to the tolerance relation T_d , as can be checked in Table 9.

Consequently, and taking into account that similar conclusions were extracted in Example 5, we have to pay special attention to the introduction of the NT support in the

third classification method. For instance, in Definition 14, we included it in a weighted average that also depends on a threshold ε , giving more satisfactory results.

On the other hand, the α - T -credibility is a great notion for detecting inconsistencies in the rules and thus in the decision table, since it measures how reliable a decision rule is by comparing the objects similar to the antecedent of this rule with the objects similar to the first ones. Nevertheless, its consideration in the study of new objects may not be useful in decision tables in which the condition attributes are more granulated than the decision attributes, since small changes in the α cut may lead to drastic changes in the α - T -credibility. We will keep studying this notion in decision tables with a fuzzy decision attribute in the future.

6. Conclusions and Future Work

In this paper, we have faced the decision-making problem from the FRST framework. We have presented four different methods to classify new objects, each of them based on a novel notion, such as the degree of suitability of a decision, the degree of ε -representativeness of a decision, the degree of α -credibility of a decision, and the degree of support-driven α -credibility of a decision. We have introduced a detailed discussion about the advantages and disadvantages of each method, showing that each one solves problems of the previous method and that they can be combined and applied to the same data set. We have provided examples for illustrating each classification method and the developed theory, as well as carrying out a comparison between the proposed methods and the ones given in the literature. In particular, this comparison allows us to ensure that the introduced classification methods are richer alternatives or generalize the ones considered in [20,21,23–25]. Finally, we have studied the classification of a new object in a real case associated with a knowledge system for decision making on investments in a water conservancy project.

As future work, we will continue studying the decision-making process in the multi-adjoint framework, given its usefulness in studying complex data sets and proposing new methods to classify objects. Furthermore, we will study the notion of the credibility of a decision rule in decision tables in which all the attributes are fuzzy, including the decision attribute d . We will also study the credibility of decision rules considering fuzzy reducts [5] in their antecedents, given their relevant properties, allowing us to analyze possible relationships between the credibility and these reducts. Finally, we are also interested in studying this classification problem in other frameworks, such as Formal Concept Analysis.

Author Contributions: All authors contributed equally to this work. All authors read and agreed to the published version of the manuscript.

Funding: This research has been partially supported by the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in projects PID2019-108991GB-I00 and PID2022-137620NB-I00, with the Ecological and Digital Transition Projects 2021 of the Ministry of Science and Innovation in project TED2021-129748B-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nakamura, A. Fuzzy rough sets. Note on Multiple-Valued Logic in Japan. *Comput. Sci.* **1988**, *9*, 1–8.
2. Nanda, S.; Majumdar, S. Fuzzy rough sets. *Fuzzy Sets Syst.* **1992**, *45*, 157–160. [[CrossRef](#)]
3. Pawlak, Z. Information systems theoretical foundations. *Inf. Syst.* **1981**, *6*, 205–218. [[CrossRef](#)]
4. Pawlak, Z. Rough sets. *Int. J. Comput. Inf. Sci.* **1982**, *11*, 341–356. [[CrossRef](#)]
5. Cornelis, C.; Medina, J.; Verbiest, N. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *Int. J. Approx. Reason.* **2014**, *55*, 412–426. [[CrossRef](#)]

6. Medina, J. Towards multi-adjoint property-oriented concept lattices. *Lect. Notes Artif. Intell.* **2010**, *6401*, 159–166.
7. Medina, J. Multi-adjoint property-oriented and object-oriented concept lattices. *Inf. Sci.* **2012**, *190*, 95–106. [\[CrossRef\]](#)
8. Pawlak, Z. *Rough Sets: Theoretical Aspects of Reasoning About Data*; Kluwer Academic Publishers: Norwell, MA, USA, 1992.
9. Pawlak, Z. Rough sets and decision algorithms. In *Rough Sets and Current Trends in Computing*; Ziarko, W., Yao, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2001; pp. 30–45.
10. Li, J.; Mei, C.; Lv, Y. Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction. *Int. J. Approx. Reason.* **2013**, *54*, 149–165. [\[CrossRef\]](#)
11. Stawicki, S.; Ślęzak, D.; Janusz, A.; Widz, S. Decision bireducts and decision reducts—A comparison. *Int. J. Approx. Reason.* **2017**, *84*, 75–109. [\[CrossRef\]](#)
12. Yao, Y. Three-way decisions with probabilistic rough sets. *Inf. Sci.* **2010**, *180*, 341–353. [\[CrossRef\]](#)
13. Chacón-Gómez, F.; Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Rough set decision algorithms for modeling with uncertainty. *J. Comput. Appl. Math.* **2023**, *437*, 115413. [\[CrossRef\]](#)
14. Mi, Y.; Wang, Z.; Liu, H.; Qu, Y.; Yu, G.; Shi, Y. Divide and conquer: A granular concept-cognitive computing system for dynamic classification decision making. *Eur. J. Oper. Res.* **2022**, *308*, 255–273. [\[CrossRef\]](#)
15. Wang, G.; Guo, Y.; Zhang, W.; Xie, S.; Chen, Q. What type of algorithm is perceived as fairer and more acceptable? A comparative analysis of rule-driven versus data-driven algorithmic decision-making in public affairs. *Gov. Inf. Q.* **2023**, *40*, 101803. [\[CrossRef\]](#)
16. Selvan, S.U.; Saroglou, S.T.; Joschinski, J.; Calbi, M.; Vogler, V.; Barath, S.; Grobman, A.Y.J. Toward multi-species building envelopes: A critical literature review of multi-criteria decision-making for design support. *Build. Environ.* **2023**, *231*, 110006. [\[CrossRef\]](#)
17. Wu, J.J.; Mazzuchi, T.A.; Sarkani, S. Comparison of multi-criteria decision-making methods for online controlled experiments in a launch decision-making framework. *Inf. Softw. Technol.* **2023**, *155*, 107115. [\[CrossRef\]](#)
18. Harding, S. Supported decision making: Skills and tools for clinical practice. *J. Nurse Pract.* **2023**, *19*, 104497. [\[CrossRef\]](#)
19. Hendrikse, S.C.; Kluiver, S.; Treur, J.; Wilderjans, T.F.; Dikker, S.; Koole, S.L. How virtual agents can learn to synchronize: An adaptive joint decision-making model of psychotherapy. *Cogn. Syst. Res.* **2022**, *79*, 138–155. [\[CrossRef\]](#)
20. Nguyen, H.S. Scalable classification method based on rough sets. In *Rough Sets and Current Trends in Computing*; Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 433–440.
21. Słowiński, R.; Stefanowski, J. Rough classification with valued closeness relation. In *New Approaches in Classification and Data Analysis*; Diday, E., Lechevallier, Y., Schader, M., Bertrand, P., Burtschy, B., Eds.; Springer: Berlin/Heidelberg, Germany, 1994; pp. 482–489.
22. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. [\[CrossRef\]](#)
23. Hamed, A.; Sobhy, A.; Nassar, H. Accurate classification of COVID-19 based on incomplete heterogeneous data using a knn variant algorithm. *Arab. J. Sci. Eng.* **2021**, *46*, 8261–8272. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Stefanowski, J.; Tsoukiás, A. Induction of decision rules and classification in the valued tolerance approach. In *Rough Sets and Current Trends in Computing*; Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 271–278.
25. Duong-Bao, N.; He, J.; Thi, L.N.; Nguyen-Huu, K.; Lee, S.-W. A novel valued tolerance rough set and decision rules method for indoor positioning using wifi fingerprinting. *Sensors* **2022**, *22*, 5709. [\[CrossRef\]](#)
26. Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Multi-adjoint algebras versus non-commutative residuated structures. *Int. J. Approx. Reason.* **2015**, *66*, 119–138. [\[CrossRef\]](#)
27. Zedam, L.; Bouremel, H.; Baets, B. Left- and right-compatibility of order relations and fuzzy tolerance relations. *Fuzzy Sets Syst.* **2019**, *360*, 65–81. [\[CrossRef\]](#)
28. Luca, A.; Termini, S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf. Control* **1972**, *20*, 301–312. [\[CrossRef\]](#)
29. Stefanowski, J.; Tsoukiás, A. Valued tolerance and decision rules. In *Rough Sets and Current Trends in Computing*; Ziarko, W., Yao, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2001; pp. 212–219.
30. Qu, J.; Bai, X.; Gu, J.; Taghizadeh-Hesary, F.; Lin, J. Assessment of rough set theory in relation to risks regarding hydraulic engineering investment decisions. *Mathematics* **2020**, *8*, 1308. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Capítulo 7

Conclusiones y trabajo futuro

Este capítulo está dedicado a exponer las conclusiones obtenidas en esta tesis doctoral, en primer lugar a nivel global, y después detallando las conclusiones derivadas de cada artículo publicado en revistas científicas de alto impacto que sustentan la investigación realizada. Finalmente, se presentarán las perspectivas de futuro para continuar trabajando las temáticas tratadas en este estudio.

7.1. Conclusiones alcanzadas

El principal objetivo de esta tesis doctoral ha sido el análisis de conjuntos de datos con información imprecisa o incompleta usando las herramientas matemáticas proporcionadas por RST y FRST. Esta investigación se inició mediante un estudio comparativo de diversas nociones de RST centradas en la reducción de información, las cuales facilitan el tratamiento de los datos. Posteriormente, debido a la presencia de información imperfecta en las bases de datos reales, se hizo necesario su estudio desde el enfoque difuso de FRST, por lo que fue imprescindible la generalización de diversas nociones de RST a FRST. Las conclusiones derivadas de este

enfoque difuso, son de especial utilidad en dos temáticas de gran importancia en la actualidad: la predicción del comportamiento de nuevos datos y la toma de decisiones, en las que también hemos centrado este trabajo. A continuación, se detallan las conclusiones de cada uno de los estudios que han contribuido al desarrollo de la presente tesis doctoral.

Conclusiones alcanzadas en *Value reducts and bireducts: a comparative study*, artículo publicado en *Math. Methods in the Applied Sciences*

Este artículo se centra en el estudio comparativo de las nociones de reducto y reducto valor con los bireductos, tanto en tablas de información como de decisión. En primer lugar, se ha mostrado que el subconjunto de atributos de un bireducto de información, proporciona un recubrimiento del conjunto de objetos de la tabla de información en subtablas no redundantes y maximales en el conjunto de objetos. Esta propiedad asegura la posibilidad de reconstruir la tabla original a partir de cada una de las subtablas. Posteriormente, se ha mostrado que es posible obtener reductos a partir de bireductos y viceversa. Por tanto, el usuario puede seleccionar el método de reducción que piense más adecuado para su estudio y obtener la reducción correspondiente, y mediante estos resultados, disponer sencillamente de otra reducción aportada por los otros mecanismos.

Además, se han introducido las tablas de información reducidas por clases para eliminar las redundancias de la tabla de información original. Este hecho ha permitido obtener relaciones entre reductos y bireductos que no se satisfacen en las tablas de información. La segunda parte de este trabajo se ha dedicado al estudio de relaciones entre los bireductos y los reductos valor, tanto en tablas de información como de decisión. Nuevamente, a pesar de haber sido nociones definidas de forma independiente, se ha mostrado la estrecha relación entre ambas, pudiendo obtener unos a partir de otros y viceversa. Como consecuencia, los cálculos para la obten-

ción de los reductos valor y los birreductos se simplifican notablemente.

Conclusiones alcanzadas en *Rough set decision algorithms for modeling with uncertainty*, artículo publicado en *Journal of Computational and Applied Mathematics*

En este trabajo se han extendido multitud de nociones relacionadas con las reglas de decisión en RST al marco de trabajo de FRST, con el fin de abordar estudios desde este enfoque, y aprovechar así las numerosas ventajas del mismo. En primer lugar, se introduce la noción de fórmula en el marco de FRST, con la diferencia de que el conjunto de objetos que satisface dicha fórmula ahora es un conjunto difuso. De aquí, la evaluación de cada objeto en dicha fórmula da como resultado un valor en el intervalo unidad, interpretado como el grado de satisfacibilidad del objeto estudiado respecto a la fórmula considerada por medio de una relación de tolerancia separable. Este grado de satisfacibilidad fue fundamental para presentar los principales indicadores de relevancia de las reglas de decisión de RST en el entorno de FRST, lo que permite analizar la información dada por las reglas de decisión desde el punto de vista difuso.

Por otra parte, se define una relación de tolerancia difusa separable específica para comparar pares de fórmulas, dando lugar a α -bloques de tolerancia que agrupan aquellas fórmulas más similares de acuerdo a un umbral α en el intervalo unidad. Teniendo en cuenta la relación anterior y las reglas de decisión consistentes con respecto al umbral α , se presentan los algoritmos de decisión en el marco de trabajo de FRST. Esta noción es esencial para extraer conjuntos de reglas de decisión representativas de una tabla de decisión desde el enfoque difuso. Además, se ha probado que tanto los algoritmos de decisión difusos como el resto de nociones introducidas en este trabajo generalizan las dadas por Pawlak en RST. Por último, se estudia un conjunto de datos reales asociado a la inversión económica

en proyectos de construcción de ingeniería hidráulica a gran escala, apoyándonos en las nociones presentadas. Este hecho ha permitido obtener varias conclusiones relevantes, como la gran inconsistencia en dicha tabla, la cual no fue detectada en [76], puesto que los autores optaron por discretizar los datos y aplicar las nociones booleanas de RST. Por tanto, estudiar las bases de datos en el entorno de FRST contribuye significativamente a lograr análisis detallados, exhaustivos y enriquecedores.

Conclusiones alcanzadas en *Decision Making in Fuzzy Rough Set Theory*, artículo publicado en *Mathematics*

Tomando como punto de partida los avances llevados a cabo en el artículo previo, en esta publicación se aborda el problema de la toma de decisiones en el marco de trabajo de FRST multiadjunto. El estudio comienza introduciendo una generalización de la noción de credibilidad al marco de trabajo multiadjunto, dotándola de la flexibilidad propia de este entorno. Esta noción ha sido indispensable para los dos últimos métodos de clasificación de nuevos objetos de los cuatro presentados en este trabajo. Cada uno de estos métodos afronta el problema de la toma de decisiones desde un enfoque diferente por medio de nuevas nociones. Estas nociones son los grados de idoneidad, de ε -representatividad, de α -credibilidad y de α -credibilidad orientado por el soporte. Además, se incluye una discusión detallada acerca de las ventajas y desventajas de cada método, mostrando que cada uno de ellos soluciona problemas que tienen el resto de los métodos, por lo que pueden ser combinados y aplicados al mismo conjunto de datos para realizar un estudio más exhaustivo. Por otra parte, se lleva a cabo una comparativa entre los métodos de clasificación propuestos y los dados en la literatura. En particular, esta comparativa permite asegurar que los métodos de clasificación introducidos son alternativas más ricas y que generalizan las consideradas en [41, 45, 68, 87, 91]. Finalmente, se

estudia la clasificación de un nuevo objeto en el conjunto de datos reales del trabajo anterior usando cada uno de los métodos presentados. Los resultados obtenidos muestran la aplicabilidad de la propuesta a bases de datos reales, siendo de gran ayuda dada su trazabilidad, adaptabilidad y flexibilidad, y presentando una alternativa a las técnicas ofrecidas por otras herramientas que pueden utilizarse para un aprendizaje automático (*machine learning*) y que no son trazables ni sus resultados explicables formalmente.

7.2. Perspectiva

Respecto a los trabajos futuros que se derivan de esta tesis doctoral, haremos una distinción entre los objetivos a corto, medio y largo plazo. Para contextualizar el primer objetivo a cumplir es importante destacar que Pawlak presentó la noción de eficiencia de algoritmo de decisión [73] en el entorno clásico, con el fin de cuantificar la facultad del algoritmo de decisión seleccionado para describir la tabla de decisión. De esta forma, mediante un valor perteneciente al intervalo unidad, la eficiencia proporciona la capacidad de clasificación de dicho algoritmo y permite comparar diferentes algoritmos entre sí. En un futuro próximo, estamos interesados en continuar con los trabajos [16, 19] para introducir una noción generalizada de eficiencia en función de un umbral para estudiar los algoritmos de decisión considerados en el marco de trabajo de FRST. Por otra parte, usaremos las nociones de la reducción de atributos, como la de reducto y reducto valor, con el objetivo de obtener algoritmos de decisión difusos más simples sin perder información relevante. Para confirmar este último objetivo, analizaremos cómo impactan estas reducciones en los umbrales asociados a los algoritmos de decisión. Además, mediante la nueva noción de eficiencia los compararemos con el algoritmo de decisión original (teniendo en cuenta todos los atributos de la tabla de decisión estudiada).

Asimismo, también compararemos los umbrales de los algoritmos de decisión y el umbral de la eficiencia para analizar la relación entre ellos.

En cuanto a los objetivos planteados a medio plazo, estamos interesados en el estudio formal de la lógica de RST [32] para analizar en detalle las nociones de este entorno desde un punto de vista lógico. También proseguiremos investigando la clasificación de nuevos objetos para estudiar el impacto de su incorporación en la tabla de decisión de partida, así como el estudio general de bases de datos dinámicas [37]. Por otro lado, continuaremos estudiando el vínculo existente entre FCA y RST para el enriquecimiento de esta última teoría. Además, también incorporaremos la teoría de posibilidad [39, 40] en FCA para el manejo de información incompleta y con incertidumbre, y abordaremos el estudio de los atributos negativos [74, 79, 80] para analizar implicaciones de atributos en contextos mixtos [78]. Todos estos estudios se trasladarán al entorno de RST aprovechando la estrecha relación entre las implicaciones de atributos de FCA y las reglas de decisión de RST [14].

Finalmente, como objetivo a largo plazo quisiéramos implementar y trasladar todas las nociones y procedimientos presentados en esta tesis doctoral, es decir, los reductos, reductos valor, birreductos, reglas y algoritmos de decisión difusos y los métodos de clasificación de nuevos objetos, al diseño de un sistema inteligente. Este debe ser robusto, fiable, trazable y adaptable a los datos de partida introducidos para su correcto desempeño en la toma de decisiones. Mediante el uso de los reductos y reductos valor se simplificarán notablemente dichos datos, facilitando el tratamiento de los mismos, así como con las reglas y algoritmos de decisión el sistema podrá extraer conclusiones relevantes de los datos correspondientes, ayudando considerablemente a la clasificación de nuevos objetos. También estamos interesados en trasladar los estudios realizados en RST a otras teorías relacionadas. En particular, analizaremos la aplicación de los birreductos y reductos valor en el marco de FCA y en la teoría de las ecuaciones

de relaciones difusas [22, 23, 33, 82, 85]. Por otra parte, compararemos los algoritmos de decisión con los programas lógicos en programación lógica difusa [24, 53].

Anexos

Otras publicaciones relacionadas

Anexo A

Information Bireducts and Its Relationship with Reducts

Information Bireducts and Its Relationship with Reducts



Fernando Chacón-Gómez, M. Eugenia Cornejo, Jesús Medina,
and Eloísa Ramírez-Poussa

Abstract Information bireducts are useful tools in Rough Set Theory in order to simultaneously reduce the sets of objects and attributes in a dataset. Specifically, information bireducts provide non-redundant subtables preserving the original discernibilities. This paper presents different properties of information bireducts taking special interest in its relationship with reducts.

Keywords Rough Set Theory · Information tables · Reducts · Bireducts

1 Introduction

Rough Set Theory (RST) was introduced by Pawlak in the 1980s as an independent and formal tool for modeling and processing incomplete information contained in relational databases [10]. Nowadays, collected databases are usually composed of a huge amount of data, whose treatment becomes into a really difficult task. Hence, an useful, and many times necessary, preprocessing is to remove the contained redundant information.

Partially supported by the the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science and Technology (COST) Action CA17124.

F. Chacón-Gómez · M. E. Cornejo (✉) · J. Medina · E. Ramírez-Poussa
Department of Mathematics, University of Cádiz, Cádiz, Spain
e-mail: mariaeugenia.cornejo@uca.es

F. Chacón-Gómez
e-mail: fernando.chacon@uca.es

J. Medina
e-mail: jesus.medina@uca.es

E. Ramírez-Poussa
e-mail: eloisa.ramirez@uca.es

Reducts and bireducts play a key role in reduction mechanisms considered for information table in RST. On the one hand, reducts are minimal subset of attributes preserving the same information as the original set [5, 6, 9, 13, 14]. On the other hand, bireducts arise as an extension of the classical notions of reduct in order to reduce the number of attributes and objects by avoiding incompatibilities, redundancies and removing the existing noise in relational databases [2–4, 7, 8, 11, 12].

This work provides different properties associated with bireducts of an information table and its relationship with reducts. In particular, the following cases are analyzed: information bireducts whose subset of objects contains all objects, information bireducts whose subset of attributes contains all attributes, information bireducts whose subsets of objects or attributes are the empty set. Furthermore, fixed a subset of attributes B of a bireduct, we will prove that every objects of the original table is included in at least a non-redundant information subtable (information bireduct) of B . These results has taken special attention to the particular case of the subset of attributes of the bireduct is a reduct.

2 Preliminaries

This section presents the basic definitions and properties of RST which will play a crucial role in our study. First of all, we formally define information tables which represent databases in RST framework.

Definition 1 Let U be a non-empty set of objects, called *universe* and $\mathcal{A}$ be a non-empty set of attributes. An *information table* is a tuple $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ such that $\mathcal{V}_{\mathcal{A}} = \{\overline{V}_a \mid a \in \mathcal{A}\}$, where V_a is the set of values associated with the attribute a over U , and $\overline{\mathcal{A}} = \{\overline{a} : U \rightarrow V_a \mid a \in \mathcal{A}\}$.

Next, an equivalence relation is defined on the set of objects of an information table. This equivalence relation will be useful to compare objects according to a given subset of attributes.

Definition 2 Let $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ be an information table. The *indiscernibility relation* $I : \mathcal{P}(\mathcal{A}) \rightarrow \mathcal{P}(U \times U)$ is defined as the equivalence relation

$$I(B) = \{(x, y) \in U \times U \mid \overline{a}(x) = \overline{a}(y), \text{ for all } a \in B\}$$

for each $B \subset \mathcal{A}$, which is called *B-indiscernibility relation*. Each class of $I(B)$ can be written as $[x]_{I(B)} = \{y \in U \mid (x, y) \in I(B)\}$. The partition determined by $I(B)$ on the set of objects U is denoted as $U/I(B) = \{[x]_{I(B)} \mid x \in U\}$.

Notice that, when $B = \emptyset$, we have that $I(\emptyset) = U \times U$. Based on an indiscernibility relation, the notions of indiscernible and discernible objects are introduced as follows.

Definition 3 Given an information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ and a subset of attributes $B \subseteq \mathcal{A}$, we say that $x, y \in U$ are *B-indiscernible* if $y \in [x]_{I(B)}$. Otherwise, we say that x, y are *B-discernible*.

Now, we define a fundamental notion for the analysis and extraction of information from databases, that is, the notion of reduct. Reducts are introduced to cover the need to reduce information tables, removing their redundant and/or superfluous variables without losing the underlying information.

Definition 4 Given an information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ and a subset of attributes $B \subseteq \mathcal{A}$, we say that:

- $a \in B$ is *dispensable* in B if $I(B) = I(B \setminus \{a\})$. Otherwise, a is *indispensable* in B .
- B is *independent* if all its attributes are indispensable in B .
- $B' \subseteq B$ is a *reduct* of B if B' is independent and $I(B) = I(B')$.

Notice that, a reduct of $\mathcal{A}$ is a subset of attributes that preserves the partition and therefore, it classifies objects in the same way that if we consider the whole set of attributes. Consequently, attributes not belonging to a reduct can be removed from the information table preserving the original information (indiscernibilities). The notion of reduct will be illustrated in the following example.

Example 1 Let $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ be the information table represented in Table 1, where the set of objects $U = \{1, 2, 3, 4, 5\}$ represents customers and the set of attributes is given by $\mathcal{A} = \{\text{Height}(h), \text{Age}(a), \text{Gender}(g)\}$.

We will prove that $B = \{h, a\}$ is a reduct of $\mathcal{A}$. From Table 1 we obtain:

$$U/I(\mathcal{A}) = \{[1]_{I(\mathcal{A})}, [3]_{I(\mathcal{A})}, [4]_{I(\mathcal{A})}, [5]_{I(\mathcal{A})}\}$$

where $[1]_{I(\mathcal{A})} = \{1, 2\}$ and the other classes are the trivial classes. Clearly, it is easy to see that $U/I(\mathcal{A}) = U/I(B)$, or equivalent $I(\mathcal{A}) = I(B)$. Now, we will show that h and a are indispensable in B , or equivalently that the equalities $I(B) = I(B \setminus \{h\})$ and $I(B) = I(B \setminus \{a\})$ do not hold. This fact is deduced from the following chains:

$$\{3\} = [3]_{I(B)} \neq [3]_{I(\{a\})} = \{3, 4\}$$

Table 1 Table associated with $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ given in Example 1

Person	Height (h)	Age (a)	Gender (g)
1	Medium	Young	Woman
2	Medium	Young	Woman
3	Tall	Medium	Man
4	Short	Medium	Man
5	Short	Old	Woman

$$\{4\} = [4]_{I(B)} \neq [4]_{I(\{h\})} = \{4, 5\}$$

As all attributes in B are indispensable, we obtain that B is independent. Therefore, we can conclude that B is a reduct of $\mathcal{A}$. $\square$

3 Bireducts in Information Tables

Information bireducts generalize information reducts by operating with subsets of attributes and subsets of objects [4, 8, 12]. This section recalls the definition of information bireduct and provides interesting properties for a better understanding of this notion. Different results about extreme cases in information bireducts are analyzed, that is, when the considered subset of objects and attributes are either the universe or the empty set. Moreover, we introduce an interesting property which shows that, fixed a subset of attributes, all objects appear at least in a non-redundant information subtable (information bireduct).

From now on, an information table $(U, \mathcal{A}, \mathcal{V}_{\mathcal{A}}, \overline{\mathcal{A}})$ will be fixed. On this table, the notion of information bireduct is formally defined as follows.

Definition 5 Given $X \subseteq U$ and $B \subseteq \mathcal{A}$, we say that:

- B is *X-irreducible* if there is no $B' \subset B$ such that all pairs $x, y \in X$ are B' -discernible.
- X is *B-inextensible* if there is no $X' \subset U$ with $X \subset X'$ such that all pairs $x, y \in X'$ are B -discernible.
- The pair (X, B) is an *information bireduct* if all pairs $x, y \in X$ are B -discernible, B is X -irreducible and X is B -inextensible.

It is convenient to mention that this definition clarifies the usual one [4, 8, 12], since we need to fix the subset of objects to introduce the notion of irreducible subset of attributes and, similarly, to fix the subset of attributes in order to define an inextensible subset of objects. Definition 5 is illustrated in the following example.

Example 2 Coming back to the information table given in Example 1 and fixing the subset of attributes $B = \{h, a\} \subseteq \mathcal{A}$, we will compute all information bireducts related to this subset. Taking into account that the equality $U/I(\mathcal{A}) = U/I(B)$ was showed in Example 1, we can ensure that the pair of objects 1, 2 are B -indiscernible and the other pairs are B -discernible from it. Notice that, removing from U an object of the pair 1, 2, we obtain that the sets $U \setminus \{1\}$ and $U \setminus \{2\}$ are all possible B -inextensible subsets of objects, by construction. Consequently, by Definition 5, the candidates to information bireducts are the pairs $(U \setminus \{1\}, B)$ and $(U \setminus \{2\}, B)$.

On the other hand, B must be X -irreducible in order to satisfy the conditions required in the notion of information bireduct. It can be seen immediately that objects

3, 4 are $\{a\}$ -indiscernible and objects 4, 5 are $\{h\}$ -indiscernible. Hence, there is no $B' \subset \mathcal{A}$ with $B' \subset B$ such that all pairs $x, y \in X$ are B' -discernible, being $X \in \{U \setminus \{1\}, U \setminus \{2\}\}$. As a consequence, B is X -irreducible. Therefore, we can ensure that $(U \setminus \{1\}, B)$ and $(U \setminus \{2\}, B)$ are all possible information bireducts with B as subset of attributes. $\square$

In the following, we will introduce some results about particular cases of information bireducts. The first one, which relates the notion of reduct (Definition 4) to the notion of information bireduct (Definition 5), was already proven in [1] and it is recalled next.

Proposition 1 ([1]) *Given $B \subseteq \mathcal{A}$, if the pair (U, B) is an information bireduct, then B is a reduct of $\mathcal{A}$.*

This property indicates that if (U, B) is an information bireduct then we can distinguish all the objects of U by using B , so that B is a reduct. Notice that if B is a reduct of $\mathcal{A}$, then B distinguishes the same objects that $\mathcal{A}$, but there may be objects that $\mathcal{A}$ is not able to distinguish. Therefore, the reciprocal of the last property is not true in general.

Example 3 Considering again the information table given in Example 1, we obtained that $B = \{h, a\}$ is a reduct of $\mathcal{A}$ and $(U \setminus \{1\}, B)$ and $(U \setminus \{2\}, B)$ are all information bireducts with B as subset of attributes. Hence, (U, B) is not an information bireduct and, as a consequence, the reciprocal of Proposition 1 is not true. $\square$

Hence, as the previous example has advanced, it is possible to obtain information bireducts from a reduct, if a subset of objects $X \subseteq U$ is considered instead of U .

Proposition 2 *Given $B \subseteq \mathcal{A}$, if B is a reduct of $\mathcal{A}$ then there exists $X \subseteq U$ such that (X, B) is an information bireduct.*

It is important to emphasize the importance of this last result since it relates the notions of reduct and information bireduct. Hence, if a reduct B of an information table is known, it is possible to construct an information subtable from the information bireduct considered (X, B) where all the objects are B -discernible.

The following result shows the case when the subset of attributes of an information bireduct is $\mathcal{A}$, that is, all attributes are required to distinguish objects.

Corollary 1 *There exists $X \subseteq U$ such that $(X, \mathcal{A})$ is an information bireduct if and only if $\mathcal{A}$ is a reduct of $\mathcal{A}$.*

Notice that, Corollary 1 ensures that the reciprocal of Proposition 2 is only guaranteed when $B = \mathcal{A}$.

Now, we will study the cases when the subset of attributes or objects of an information bireduct is the empty set. First of all, we study information bireducts whose subset of attributes is the empty set, that is, there are no attribute in order to distinguish the elements in the subset of objects.

Proposition 3 *Let (X, B) be an information bireduct, with $X \subseteq U$, $B \subseteq \mathcal{A}$. Then $B = \emptyset$ if and only if $X = \{x\}$ for all $x \in U$.*

As a consequence, since there are no attribute to distinguish objects, the subset of objects of each information bireduct must be a singleton, so all the information subtables generated by information bireducts have a single object. For instance, this happens when there exists $B' \subseteq \mathcal{A}$ such that all objects of an information bireduct are B' -indiscernible. Since it is no possible to distinguish any pair of objects by using attributes of B' , there is no information bireducts with any subset of B' , excluding the empty set, as subset of attributes.

On the other hand, it is important to emphasize that there no exist information bireducts whose subset of objects is the empty set. It occurs because an information bireduct with at least one object can always be considered, as it is shown in the last result.

Finally, the following property guarantees that, given $B \subseteq \mathcal{A}$, if there exist information bireducts in the information table with B as subset of attributes, then each object belongs to some information bireduct.

Theorem 1 *Given a subset of attributes $B \subseteq \mathcal{A}$ and the subsets of objects $X_1, \dots, X_n \subseteq U$ such that the pairs $(X_1, B), \dots, (X_n, B)$ are all the information bireducts, then $U = \bigcup_{i=1}^n X_i$.*

The last property will be applied to the information table of Example 1.

Example 4 As we computed before, given the subset of attributes $B = \{h, a\} \subseteq \mathcal{A}$ and the subsets of objects $U \setminus \{1\}$ and $U \setminus \{2\}$, we have that $(U \setminus \{1\}, B)$ and $(U \setminus \{2\}, B)$ are all the information bireducts associated with B . Therefore, according to Theorem 1, it is straightforwardly obtained that:

$$U = U \setminus \{1\} \cup U \setminus \{2\}$$

As a result, we can ensure that each object belongs to some information bireduct. $\square$

As a consequence, when an information table is divided into non-redundant subtables by a subset of attributes B of an information bireduct, we have that all objects of the original table are represented in at least one of these subtables. This property can be useful for translating the obtained information from each bireduct to the original dataset. In addition, this fact takes on special importance when the subset of attributes of the computed information bireducts is a reduct of $\mathcal{A}$. In this case, the division given by a reduct contains the same information that the original table, providing a great representation of the database. Finally, notice that this division always exists because of Proposition 2.

4 Conclusions and Future Work

We have studied different properties of information bireducts, focusing on particular cases of the subsets of objects and attributes, respectively. We have proven that, when the subset of objects or attributes of information bireducts are the whole sets, we can straightforwardly obtain reducts of the information table. In addition, we have also studied the information bireducts whose subset of objects or attributes are the empty set. Finally, we have introduced an interesting property which guarantees that, given a subset of attributes B of a bireduct, each object in the universe belongs at least to an information bireduct associated with B .

As a future work, we will study properties of decision bireducts in order to relate them with reducts in decision tables framework complementing the ones given in [12]. In addition, we will study the relationship between information/decision bireducts and value reducts. These last reducts are another important notion in RST, due to they allow to remove some values of certain attributes for a specific object, when it is not possible to delete such attributes because they are necessary (for preserving the information contained in the considered database).

References

1. Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E.: Towards a classification of rough set bireducts. In: Lesot, M.J., Vieira, S., Reformat, M.Z., Carvalho, J.a.P., Wilbik, A., Bouchon-Meunier, B., Yager, R.R. (eds.) *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 759–770. Springer International Publishing, Cham (2020)
2. Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E., Ślęzak, D.: A computational procedure for variable selection preserving different initial conditions. *Int. J. Comput. Math.* **97**(1–2), 387–404 (2020)
3. Benítez-Caballero, M.J., Medina-Moreno, J., Ramírez-Poussa, E.: Bireducts in Formal Concept Analysis, pp. 191–198. Springer International Publishing, Cham (2020)
4. Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E., Ślęzak, D.: Bireducts with tolerance relations. *Inf. Sci.* **435**, 26–39 (2018)
5. Cornelis, C., Jensen, R., Hurtado, G., Ślęzak, D.: Attribute selection with fuzzy decision reducts. *Inf. Sci.* **180**, 209–224 (2010)
6. Janusz, A., Ślęzak, D.: Rough set methods for attribute clustering and selection. *Appl. Artif. Intell.* **28**(3), 220–242 (2014)
7. Janusz, A., Ślęzak, D., Nguyen, H.S.: Unsupervised similarity learning from textual data. *Fund. Inform.* **119**(319–336), 01 (2012)
8. Mac Parthalain, N., Jensen, R.: Simultaneous feature and instance selection using fuzzy-rough bireducts. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2013), pp. 1–8 (2013)
9. Medina, J.: Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Comput. & Math. Appl.* **64**(6), 1992–2002 (2012)
10. Pawlak, Z.: Rough sets. *Int. J. Comput. Inf. Sci.* **11**, 341–356 (1982)
11. Stawicki, S., Ślęzak, D.: Recent advances in decision bireducts: complexity, heuristics and streams. In: Lingras, P., Wolski, M., Cornelis, C., Mitra, S., Wasilewski, P. (eds.) *Rough Sets and Knowledge Technology. Lecture Notes in Computer Science*, vol. 8171, pp. 200–212. Springer, Berlin (2013)

12. Stawicki, S., Ślęzak, D., Janusz, A., Widz, S.: Decision bireducts and decision reducts - a comparison. *Int. J. Approx. Reason.* **84**, 75–109 (2017)
13. Yao, Y., Zhao, Y.: Attribute reduction in decision-theoretic rough set models. *Inf. Sci.* **178**(17), 3356–3373 (2008)
14. Zhao, Y., Yao, Y., Luo, F.: Data analysis based on discernibility and indiscernibility. *Inf. Sci.* **177**(22), 4959–4976 (2007)

Anexo B

Fuzzy Rough Set Decision Algorithms



Fuzzy Rough Set Decision Algorithms

Fernando Chacón-Gómez^(✉), Maria Eugenia Cornejo, Jesús Medina,
and Eloísa Ramírez-Poussa

Department of Mathematics, University of Cádiz, Cádiz, Spain
{fernando.chacon,mariaeugenia.cornejo,jesus.medina,eloisa.ramirez}@uca.es

Abstract. Decision rules are a powerful tool for the management of information from a relational database, allowing the extraction of conclusions. A decision algorithm collects a characteristic set of decision rules and its efficiency is one of its most interesting properties, since it allows us to know the usefulness of the decision algorithm and to compare it with other algorithms. This paper focuses on the generalization of the notions of decision algorithm and efficiency to the fuzzy case.

Keywords: Rough Set Theory · Fuzzy Rough Set Theory · Decision Rules · Decision Algorithm

1 Introduction

Rough set theory (RST) was introduced by Pawlak [17, 18] in the 1980s as a formal tool to deal with imprecise or incomplete information contained in databases. In this theory, a relational dataset is interpreted as a decision table, that is, a set of objects, a set of attributes and a collection of mappings characterizing the objects by the attributes. RST studies the relationship among the objects by using the attributes in order to summarize and extract information from the considered dataset.

In RST, the characterization of decision tables by a set of decision rules [8, 19–21, 24] is a powerful tool for the management of information from relational databases since they allow the extraction of significative conclusions. In addition, decision rules are accompanied by some measurements which provide us with valuable information of them. A characterization of decision tables from a decision algorithm is given in [20], which shows the interest of this notion. Another fundamental notion is the efficiency of a decision algorithm because it allows to measure the usefulness of the decision algorithm and compare with other algorithms as well.

Partially supported by the 2014-2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

The fuzzy paradigm has already been developed and applied to RST in different papers [3, 4, 6, 7, 10–15, 22], due to its great flexibility to deal with databases containing uncertainty information. A complete overview about different recent fuzzy generalisations of RST can be seen in [9]. In order to increase the potential of Fuzzy Rough Set Theory (FRST), this paper focuses on generalizing the classical notions related to decision rules to the fuzzy framework, paying special attention to the notions of decision algorithm and efficiency. Furthermore, different examples have been introduced through the paper to illustrate all the new definitions.

2 Basic Notions in RST

This section includes some preliminary definitions of RST [19, 20]. First of all, it is convenient to recall that databases are represented as decision tables in this framework.

Definition 1. Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A decision table is a tuple $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with attribute $a \in \mathcal{A}_d$ over U , and $\overline{\mathcal{A}_d} = \{\bar{a}: U \rightarrow V_a \mid a \in \mathcal{A}_d\}$. In this case, the attributes of $\mathcal{A}$ are called condition attributes and d is called decision attribute.

Notice that only one decision attribute is being considered in this definition. However, a similar study could be carried out when more than one decision attribute is worked.

Next, an equivalence relation is defined on the set of objects of a decision table. This equivalence relation will be useful to compare objects according to a given subset of attributes.

Definition 2. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table. The indiscernibility mapping $I: \mathcal{P}(\mathcal{A}_d) \rightarrow \mathcal{P}(U \times U)$ is defined, for each $B \subseteq \mathcal{A}_d$, as the equivalence relation

$$I(B) = \{(x, y) \in U \times U \mid \bar{a}(x) = \bar{a}(y), \text{ for all } a \in B\}$$

which is called B -indiscernibility relation. Each class of $I(B)$ can be written as $[x]_{I(B)} = \{y \in U \mid (x, y) \in I(B)\}$. The partition determined by $I(B)$ on the set of objects U is denoted as $U/I(B) = \{[x]_{I(B)} \mid x \in U\}$.

The positive region is another interesting notion of RST which is introduced below.

Definition 3. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. The positive region of the partition $U/I(\{d\})$ with respect to B is defined as

$$POS_B(\{d\}) = \bigcup_{X \in U/I(\{d\})} B_*(X)$$

where $B_*(X) = \{x \in U \mid [x]_{I(B)} \subseteq X\}$.

In the following, we recall some basic definitions of FRST which are required to study decision rules in the fuzzy environment [4]. The first one is essential to compare a pair of objects of a decision table according to an attribute.

Definition 4. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $a \in \mathcal{A}_d$. An a -indiscernibility relation is a $[0, 1]$ -fuzzy tolerance relation $R_a: U \times U \rightarrow [0, 1]$, that is, a reflexive and symmetrical fuzzy relation.

Next notion is useful to compare pairs of objects by using the total set of attributes $\mathcal{A}$ instead of a singleton. Aggregation operators [4] play a key role in the following definition.

Definition 5. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $R_a: U \times U \rightarrow [0, 1]$ be an a -indiscernibility relation, for all $a \in \mathcal{A}$. The $\mathcal{A}$ -indiscernibility relation $R_{\mathcal{A}}: U \times U \rightarrow [0, 1]$ is defined for each pair of objects $x, y \in X$ as

$$R_{\mathcal{A}}(x, y) = @ (R_{a_1}(x, y), \dots, R_{a_m}(x, y))$$

where $@: [0, 1]^m \rightarrow [0, 1]$ is an aggregation operator, that is, an increasing operator on each argument satisfying that $@(1, \dots, 1) = 1$ and $@(0, \dots, 0) = 0$.

Finally, we recall the generalization of positive region (Definition 3) to the fuzzy setting, from the notion of fuzzy implication [2].

Definition 6. Let $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $B \subseteq \mathcal{A}$ and $I: [0, 1] \times [0, 1] \rightarrow [0, 1]$ be a fuzzy implication, that is, it is increasing in the left argument (consequent), decreasing in the right argument (antecedent) and it satisfies $I(0, 0) = I(1, 1) = 1$ and $I(0, 1) = 0$. The fuzzy $\mathcal{A}$ -positive region is defined, for each $y \in U$, as:

$$POS_{\mathcal{A}}(y) = \inf \{ I(R_d(y, x), R_{\mathcal{A}}(x, y)) \mid x \in U \}$$

2.1 Decision Rules

As it was commented in [20], decision tables describe decisions in terms of conditions that must be satisfied in order to carry out the decision specified in the decision table. In addition, a set of decision rules can be associated with every decision table, which can be also seen as a logical description of approximation of decisions. This fact facilitates the interpretation of decision tables, so that the study of decision rules is justified. Before presenting the notion of decision rule, we need to define a formal language to describe approximations in logical terms.

Definition 7. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. The set of formulas associated with B , denoted as $For(B)$, is built from attribute-value pairs (a, v) , where $a \in B$ and $v \in V_a$, by means of the conjunction and disjunction logical connectives, $\wedge$ and $\vee$ respectively.

For each $\Phi \in For(B)$, the set of objects $x \in U$ that satisfies Φ in S is defined, for all $a \in B$ and $v \in V_a$, as:

$$\|\Phi\|_S = \|(a, v)\|_S = \{x \in U \mid \bar{a}(x) = v\}$$

Given $\Phi, \Psi \in \text{For}(B)$, the set of objects that satisfies $\Phi \wedge \Psi$ in S is defined as $\|\Phi \wedge \Psi\|_S = \|\Phi\|_S \cap \|\Psi\|_S$ and the set of objects that satisfies $\Phi \vee \Psi$ in S is defined as $\|\Phi \vee \Psi\|_S = \|\Phi\|_S \cup \|\Psi\|_S$.

Once we have presented the formal language associated with a subset of attributes, we can introduce the notion of decision rule.

Definition 8. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. A decision rule in S is an expression $\Phi \rightarrow \Psi$ where $\Phi \in \text{For}(B), \Psi \in \text{For}(\{d\})$ where $B, \{d\}$ are condition and decision attributes, respectively. In addition, we say that an object $x \in U$ satisfies a decision rule $\Phi \rightarrow \Psi$ if $x \in \|\Phi \wedge \Psi\|_S$.

Next, we introduce the definitions of *support*, *strength* and *certainty* of a decision rule. Specifically, the *support* takes into account the number of objects satisfying the decision rule, the *strength* represents the proportion of objects satisfying the decision rule and the *certainty* provides us with the probability of the consequent be fulfilled the antecedent. Hence, these measures offer interesting information about decision rules and they have already been applied to mining patterns in [5, 23].

Definition 9. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $\Phi \rightarrow \Psi$ be a decision rule in S , with $\Phi \in \text{For}(B)$ and $\Psi \in \text{For}(\{d\})$. We call

- support of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\text{supp}_S(\Phi, \Psi) = \|\Phi \wedge \Psi\|_S$$

- strength of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\sigma_S(\Phi, \Psi) = \frac{\text{supp}_S(\Phi, \Psi)}{|U|}$$

- certainty of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\text{cer}_S(\Phi, \Psi) = \frac{\|\Phi \wedge \Psi\|_S}{\|\Phi\|_S}$$

From the notion of certainty, we will say that $\Phi \rightarrow \Psi$ is a *true decision rule* if $\text{cer}_S(\Phi, \Psi) = 1$. Otherwise, except if $\text{cer}_S(\Phi, \Psi) = 0$, it will be called a *not entirely true decision rule*.

Notice that, the strength of a decision rule is called support of an association rule in formal concept analysis [1, 16]. In addition, the value of the certainty of a decision rule corresponds to the confidence factor of an association rule, although the meaning is slightly different.

These measures play a fundamental role in the study of decision algorithms and their efficiency, as it is shown in the next section.

2.2 Decision Algorithm

With every decision table, we can associate a decision algorithm which is a set of decision rules verifying certain conditions. Consequently, decision algorithms can be seen as a logical description of basic properties of the data, what favors the optimal treatment of information contained in decision tables. The formal definition of decision algorithm is given below [20].

Definition 10. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$. We say that:

1. $Dec(S)$ is a set of pairwise mutually exclusive (independent) decision rules, if each pair of decision rules $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisfies that $\Phi = \Phi'$ or $\|\Phi \wedge \Phi'\|_S = \emptyset$, and $\Psi = \Psi'$ or $\|\Psi \wedge \Psi'\|_S = \emptyset$.
2. $Dec(S)$ covers U , if $\|\bigvee_{i=1}^m \Phi_i\|_S = \|\bigvee_{i=1}^m \Psi_i\|_S = U$.
3. The decision rule $\Phi \rightarrow \Psi \in Dec(S)$ is admissible in S if $supp_S(\Phi, \Psi) \neq 0$.
4. $Dec(S)$ preserves the consistency of S if

$$POS_{\mathcal{A}}(\{d\}) = \bigvee_{\Phi \rightarrow \Psi \in Dec^+(S)} \|\Phi\|_S$$

where $Dec^+(S)$ is the set of true decision rules of $Dec(S)$.

The set of decision rules $Dec(S)$ satisfying the previous properties is called decision algorithm in S .

Notice that, from the first three items in Definition 10, we can deduce that the sum of the support of all decision rules is the cardinal of the set of objects, since each object is only represented in one decision rule. Consequently, the sum of the strength of all decision rules is 1.

Now, we present the notion of *efficiency* of a decision algorithm which is fundamental to compare decision algorithms and it will be one of the cornerstone of this paper.

Definition 11. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $B \subseteq \mathcal{A}$ and $Dec(S)$ a decision algorithm. We call efficiency of the decision algorithm $Dec(S)$ to the number

$$\eta(Dec(S)) = \sum_{\Phi \in For(B)} \eta_{\Phi}(Dec(S))$$

where

$$\eta_{\Phi}(Dec(S)) = \max\{\sigma_S(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in Dec(S)\}$$

According to Definition 11, we obtain that the efficiency of a decision algorithm only takes into account the decision rule with largest value of strength for each antecedent, this rule is called *representative decision rule of the antecedent*. Thus, the efficiency of a decision algorithm provides an accuracy degree of the

classification of the objects of the maximal subsets of representative decision rules of the decision table.

Notice that, a similar definition could be done using the certainty. However, the obtained value will not be bounded by 1. Hence, since the strength and certainty are related, the efficiency considers in some way the certainty but with the goal of obtaining a value between 0 and 1.

3 Decision Rules in Fuzzy Rough Set Theory

This section is devoted to the study of decision tables in the fuzzy environment, by using decision rules. To do this, we need to extend the notions presented in Sect. 2.1 to the fuzzy framework.

First of all, it is important to mention that the notion of set of formulas in FRST is the same as in Definition 7, that is attribute-value pairs (a, v) , due to the value v can be any element of an arbitrary set associated with the attribute a . This generalization provides enough flexibility to be considered in the fuzzy setting. However, the meaning of $\|\Phi\|_S$ is different.

Definition 12. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. Given $\Phi \in \text{For}(B)$, the mapping $\|\Phi\|_S: U \rightarrow [0, 1]$ is defined, for each object of the universe, as:

$$\|\Phi\|_S(x) = E_a(\bar{a}(x), v)$$

where $E_a: V_a \times V_a \rightarrow [0, 1]$ is a $[0, 1]$ -fuzzy tolerance relation. Therefore, $\|\Phi\|_S(x)$ represents how much the object x satisfies the formula Φ , through the relationship between the value of the attribute a in the object x and the value of the attribute a in the formula Φ .

Given $\Phi, \Psi \in \text{For}(B)$, $\|\Phi \wedge \Psi\|_S(x)$ and $\|\Phi \vee \Psi\|_S(x)$ represent how much the object x satisfies the formula $\Phi \wedge \Psi$ and how much the object x satisfies the formula $\Phi \vee \Psi$, respectively, both are defined as follows:

$$\begin{aligned} \|\Phi \wedge \Psi\|_S(x) &= \inf\{\|\Phi\|_S(x), \|\Psi\|_S(x)\} \\ \|\Phi \vee \Psi\|_S(x) &= \sup\{\|\Phi\|_S(x), \|\Psi\|_S(x)\} \end{aligned}$$

Notice that, the infimum and the supremum generalize the intersection and the union of sets, respectively, in the fuzzy environment, so that $\|\Phi \wedge \Psi\|_S(x)$ and $\|\Phi \vee \Psi\|_S(x)$ generalize the conjunction and disjunction of formulas in the classical environment given in Definition 7.

On the other hand, the definition of decision rule in FRST also coincides with the one given in Definition 8 the formulas are composed by conjunction or disjunction of attribute-value pairs (a, v) in both frameworks.

In the following, we will include an example in order to illustrate the formal language and the notions related to decision rules in FRST.

Example 1. Consider the decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ where the set of objects is $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ and the set of attributes is $\mathcal{A} = \{a_1, a_2, a_3\}$. This decision table is represented in Table 1. From these data, we will compute

the decision rules obtained by using the conjunction and considering the whole set of attributes. After that, we will calculate how much each object $x_j \in U$, with $j \in \{1, 2, \dots, 7\}$, satisfies each antecedent of the decision rules. We also make the corresponding computations in order to know the degree of satisfaction of each object $x_j \in U$ in all consequents of the decision rules.

Table 1. Decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ given in Example 1.

	a_1	a_2	a_3	d
x_1	0.34	0.31	0.75	0
x_2	0.21	0.71	0.5	1
x_3	0.52	0.92	1	0
x_4	0.85	0.65	1	1
x_5	0.43	0.89	0.5	0
x_6	0.21	0.47	0.25	1
x_7	0.09	0.93	0.25	0

First of all, we calculate decision rules considering the whole set of attributes by using the conjunction, and so we obtain one rule for each object.

$$\begin{aligned}
r_1 &: (a_1, 0.34) \wedge (a_2, 0.31) \wedge (a_3, 0.75) \rightarrow (d, 0) \\
r_2 &: (a_1, 0.21) \wedge (a_2, 0.71) \wedge (a_3, 0.5) \rightarrow (d, 1) \\
r_3 &: (a_1, 0.52) \wedge (a_2, 0.92) \wedge (a_3, 1) \rightarrow (d, 0) \\
r_4 &: (a_1, 0.85) \wedge (a_2, 0.65) \wedge (a_3, 1) \rightarrow (d, 1) \\
r_5 &: (a_1, 0.43) \wedge (a_2, 0.89) \wedge (a_3, 0.5) \rightarrow (d, 0) \\
r_6 &: (a_1, 0.21) \wedge (a_2, 0.47) \wedge (a_3, 0.25) \rightarrow (d, 1) \\
r_7 &: (a_1, 0.09) \wedge (a_2, 0.93) \wedge (a_3, 0.25) \rightarrow (d, 0)
\end{aligned}$$

On the one hand, to know the degree of satisfaction of each object in all antecedents of decision rules we compute $\|\Phi_i\|_S(x_j)$, where Φ_i is the antecedent of the decision rule r_i and $x_j \in U$, with $i, j \in \{1, 2, \dots, 7\}$. With this purpose, we have considered the $[0, 1]$ -fuzzy tolerance relation $E_a: V_a \times V_a \rightarrow [0, 1]$ defined as $E_a(\bar{a}(x), v) = 1 - |\bar{a}(x) - v|$, for all $a \in \mathcal{A}_d$, $x \in U$ and $v \in [0, 1]$. Since we have extracted the decision rules from each object, we obtain that $\|\Phi_i\|_S(x_j) = \|\Psi_i\|_S(x_j) = 1$, when $i = j$. Table 2 shows the rest of calculations.

On the other hand, we compute $\|\Psi_i\|_S(x_j)$ in order to know the degree of satisfaction of each object in all consequents of decision rules, where Ψ_i is the consequent of the decision rule r_i and $x_j \in U$, with $i, j \in \{1, 2, \dots, 7\}$. In an easier way, we obtain that:

$$\|\Psi_i\|_S(x_j) = \begin{cases} 1 & \text{if } \bar{d}(x_j) = v \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

□

Table 2. Degree of satisfaction of each object in all antecedents of decision rules.

	$\ \Phi_1\ _S$	$\ \Phi_2\ _S$	$\ \Phi_3\ _S$	$\ \Phi_4\ _S$	$\ \Phi_5\ _S$	$\ \Phi_6\ _S$	$\ \Phi_7\ _S$
x_1	1	0.6	0.39	0.49	0.42	0.5	0.38
x_2	0.6	1	0.5	0.36	0.78	0.75	0.75
x_3	0.39	0.5	1	0.67	0.5	0.25	0.25
x_4	0.49	0.36	0.67	1	0.5	0.25	0.24
x_5	0.42	0.78	0.5	0.5	1	0.58	0.66
x_6	0.5	0.75	0.25	0.25	0.58	1	0.54
x_7	0.38	0.75	0.25	0.24	0.66	0.54	1

In the following, we need to consider a family of $[0, 1]$ -fuzzy tolerance relations known as *separable relations* [25] in order to introduce the notions of *support*, *strength* and *certainty* of a decision rule in the fuzzy framework. The interpretation of these notions is analogous to the one given in the classical case.

Definition 13. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\Phi \rightarrow \Psi$ be a decision rule in S and $E = \{E_a : V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of $[0, 1]$ -fuzzy tolerance relations such that $E_a(\bar{a}(x), v) = 1$ if and only if $\bar{a}(x) = v$. We call:

- *E*-support of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\text{supp}_S^E(\Phi, \Psi) = \|\Phi \wedge \Psi\|_S = \sum_{x \in U} \|\Phi \wedge \Psi\|_S(x) = \sum_{x \in U} \|\Phi\|_S(x) \wedge \|\Psi\|_S(x)$$

- *E*-strength of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\sigma_S^E(\Phi, \Psi) = \frac{\text{supp}_S^E(\Phi, \Psi)}{|U|}$$

- *E*-certainty of the decision rule $\Phi \rightarrow \Psi$ to the number:

$$\text{cer}_S^E(\Phi, \Psi) = \frac{\|\Phi \wedge \Psi\|_S}{\|\Phi\|_S}$$

Analogously to the classical environment, we say that $\Phi \rightarrow \Psi$ is a *true decision rule* if $\text{cer}_S^E(\Phi, \Psi) = 1$. Otherwise, except if $\text{cer}_S^E(\Phi, \Psi) = 0$, it is a *not entirely true decision rule*.

Next, Table 3 shows *E*-support, *E*-strength and *E*-certainty for each decision rule computed in Example 1. The results are obtained by using Table 2, Eq. (1) and Definition 13.

4 Decision Algorithms in Fuzzy Rough Set Theory

This section focuses on the generalization of the notion of decision algorithm to the fuzzy case. To reach this goal, we need firstly to define a fuzzy tolerance relation which allows us to compare formulas, as the mapping given below.

Table 3. E -Support, E -strength and E -certainty of decision rules of Example 1.

Rule	$supp_S^E$	σ_S^E	cer_S^E
r_1	2.19	0.31	0.58
r_2	2.11	0.3	0.45
r_3	2.14	0.31	0.6
r_4	1.61	0.23	0.58
r_5	2.58	0.37	0.58
r_6	2	0.29	0.52
r_7	2.29	0.33	0.6

Definition 14. Given a decision table $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, we define the $[0, 1]$ -fuzzy tolerance relation $R_{Fd}: For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$, for all $\Phi, \Phi' \in For(\mathcal{A}_d)$, as follows:

$$R_{Fd}(\Phi, \Phi') = 1 - \max\{|\|\Phi\|_S(x) - \|\Phi'\|_S(x)| \mid x \in U\}$$

Given $\alpha \in [0, 1]$, the R_{Fd} - α -block is defined for each $\Phi \in For(\mathcal{A}_d)$ as:

$$[\Phi]_\alpha = \{\Phi' \in For(\mathcal{A}_d) \mid \alpha \leq R_{Fd}(\Phi, \Phi')\}$$

Notice that, R_{Fd} is not a transitive fuzzy tolerance relation and therefore, we have defined tolerance blocks instead of equivalence classes. In addition, according to the previous definition, we can deduce that α determines if two elements are related enough to belong to the same block.

Now, we can introduce the notion of decision algorithm to the fuzzy case.

Definition 15. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1]$, $\alpha_3 \geq 0$ and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$. We say that:

1. $Dec(S)$ is a set of $\alpha_1\alpha_2$ -pairwise mutually exclusive (independent) decision rules, if each pair of decision rules $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisfies that $\Phi = \Phi'$ or $\|\Phi \wedge \Phi'\|_S(x) \leq \alpha_1$ and $\Psi = \Psi'$ or $\|\Psi \wedge \Psi'\|_S(x) \leq \alpha_2$, for all $x \in U$.
2. $Dec(S)$ covers U , if $|\bigvee_{i=1}^m \Phi_i|_S| = |\bigvee_{i=1}^m \Psi_i|_S| = |U|$.
3. The decision rule $\Phi \rightarrow \Psi \in Dec(S)$ is α_3 -admissible in S if $\alpha_3 < supp_S(\Phi, \Psi)$.
4. $Dec(S)$ preserves the α -consistency of S with a degree α_4 if the next inequality holds for all $x \in U$:

$$|POS_{\mathcal{A}}(x) - \bigvee_{\Phi \rightarrow \Psi \in Dec_{\alpha}^+(S)} \|\Phi\|_S(x)| \leq \alpha_4$$

where $\alpha \in [0, 1]$ and

$$Dec_{\alpha}^+(S) = \{\Phi \rightarrow \Psi \in Dec(S) \mid \text{if there exists } \Phi' \rightarrow \Psi' \in Dec(S) \\ \text{such that } \Phi' \in [\Phi]_{\alpha} \text{ then } \Psi' \in [\Psi]_{\alpha}\}$$

The set of decision rules $Dec(S)$ satisfying the previous properties for the values $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ is called $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ -decision algorithm in S .

Observe that, none of the decision rules contained in the set $Dec_\alpha^+(S)$ lead to contradictions, that is, this set does not include decision rules for which there exists another decision rule whose antecedents are both included in the R_{Fd} - α -blocks of their antecedents but their consequents do not belong to both R_{Fd} - α -blocks of their consequents. Therefore, the set $Dec_\alpha^+(S)$ is generalizing to the set $Dec^+(S)$ included in Definition 10. As a consequence, when $\alpha_k = 0$, for each $k \in \{1, 2, 3, 4\}$, we obtain Definition 10 itself.

Finally, we introduce the generalization of the concept of efficiency.

Definition 16. Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $Dec(S)$ be an $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ -decision algorithm. Given $0 \leq \alpha \leq 1$, we call α -efficiency of the $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ -decision algorithm $Dec(S)$ to the number

$$\eta^\alpha(Dec(S)) = \sum_{\Phi \in For(\mathcal{A})} \eta_\Phi^\alpha(Dec(S))$$

where

$$\eta_\Phi^\alpha(Dec(S)) = \max\{\sigma_S(\Phi', \Psi') \mid \Phi' \rightarrow \Psi' \in Dec(S), \Phi' \in [\Phi]_\alpha\}$$

being Φ, Φ' two antecedents and Ψ' a consequent of the decision rules of $Dec(S)$.

Notice that, although $0 \leq R_{Fd}(\Phi, \Phi') \leq 1$, for all Φ, Φ' antecedents of decision rules of $Dec(S)$, the α -efficiency is not bounded by 1. It can be checked taking into account that the sum of the E -strength of all decision rules of a decision algorithm is not bounded by 1, as we will show in the following example. Thence, it is difficult to extract conclusions due to the loss of its probabilistic character. Moreover, the α -efficiency is a decreasing map due to the value of α gives a degree of flexibility. Finally, it is convenient to emphasize that, when the relations E_a are booleans and $\alpha > 0$, the α -efficiency coincides with Definition 11, because in this case $R_{Fd}(\Phi, \Phi') = 1$ if $\Phi = \Phi'$ and $R_{Fd}(\Phi, \Phi') = 0$ otherwise.

Finally, these notions are illustrated in the following example.

Example 2. Coming back to the environment of Example 1, we will compute the minimum/maximum values of α_k with $k \in \{1, 2, 3, 4\}$ in order to ensure that $Dec(S) = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$ forms a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ -decision algorithm. First of all, we calculate the value of α_1 by using Table 2 and Definition 15. Specifically, taking into account the first item of Definition 15, we compute the value

$$\max\{\|\Phi \wedge \Phi'\|_S(x) \mid \Phi, \Phi' \in For(\mathcal{A}), \Phi \neq \Phi', x \in U\} = 0.78$$

Therefore, any value $\alpha_1 \geq 0.78$ satisfies that $\|\Phi \wedge \Phi'\|_S(x) \leq \alpha_1$ for all $\Phi, \Phi' \in For(\mathcal{A})$, with $\Phi \neq \Phi'$, and $x \in U$. Clearly, $\alpha_2 \geq 0$ and we do not have

to consider more requirements because d is a boolean attribute. Now, for the computation of α_3 , we consider the third item of Definition 15 obtaining:

$$\min\{supp_S^E(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in Dec(S)\} = 1.61$$

Hence, for any value $\alpha_3 < 1.61$, the decision rules in $Dec(S)$ are α_3 -admissible in S .

Finally, for the computation of α_4 , we need to define the mappings involved in the fuzzy $\mathcal{A}$ -positive region $POS_{\mathcal{A}}$, as well as the set $Dec_{\alpha}^+(S)$, given $\alpha \in [0, 1]$.

On the one hand, in order to determine the fuzzy $\mathcal{A}$ -positive region $POS_{\mathcal{A}}$, we need to choose an a -indiscernibility relation, an $\mathcal{A}$ -indiscernibility relation and a fuzzy implication. We will consider the a -indiscernibility relation $R_a: U \times U \rightarrow [0, 1]$ defined, for all $a \in \mathcal{A}_d$, as:

$$R_a(x, y) = 1 - |\bar{a}(x) - \bar{a}(y)|$$

It is easy to see that $R_a(x, y) = 1$ if and only if $\bar{a}(x) = \bar{a}(y)$. In addition, we will consider the $\mathcal{A}$ -indiscernibility relation $R_{\mathcal{A}}: U \times U \rightarrow [0, 1]$ defined as:

$$\begin{aligned} R_{\mathcal{A}}(x, y) &= @ (R_{a_1}(x, y), R_{a_2}(x, y), R_{a_3}(x, y)) \\ &= \frac{R_{a_1}(x, y) + 2(R_{a_2}(x, y) + R_{a_3}(x, y))}{5} \end{aligned}$$

for all $x, y \in [0, 1]$. In this case, $R_{\mathcal{A}}(x, y) = 1$ if and only if $x = y$. It is important to mention that $R_{\mathcal{A}}$ is an aggregation operator and the attribute a_1 is the least important in the decision table since it has the least weight according to the definition of $R_{\mathcal{A}}$.

To sum up, R_a , $R_{\mathcal{A}}$ and the Łukasiewicz implication, defined as $z \leftarrow_L y = \min\{1, 1 + z - y\}$ for all $y, z \in [0, 1]$, will be the mappings used for computing the fuzzy $\mathcal{A}$ -positive region $POS_{\mathcal{A}}$.

On the other hand, we need to define the set $Dec_{\alpha}^+(S)$ to calculate α_4 . To do this, we will fix $\alpha = 0.75$ and, applying Definition 14, we calculate the relation between each pair of antecedents of all rules in $Dec(S)$, see Table 4.

Table 4. Relation between each pair of antecedents of decision rules of Example 1.

R_{Fd}	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
Φ_1	1	0.6	0.39	0.49	0.42	0.5	0.38
Φ_2	0.6	1	0.5	0.36	0.78	0.75	0.75
Φ_3	0.39	0.5	1	0.67	0.5	0.25	0.25
Φ_4	0.49	0.36	0.67	1	0.5	0.25	0.24
Φ_5	0.42	0.78	0.5	0.5	1	0.58	0.66
Φ_6	0.5	0.75	0.25	0.25	0.58	1	0.54
Φ_7	0.38	0.75	0.25	0.24	0.66	0.54	1

The R_{Fd} -0.75-block of each antecedent is calculated by using Definition 14, obtaining $[\Phi_1]_{0.75} = \{\Phi_1\}$, $[\Phi_2]_{0.75} = \{\Phi_2, \Phi_5, \Phi_6, \Phi_7\}$, $[\Phi_3]_{0.75} = \{\Phi_3\}$, $[\Phi_4]_{0.75} = \{\Phi_4\}$, $[\Phi_5]_{0.75} = \{\Phi_2, \Phi_5\}$, $[\Phi_6]_{0.75} = \{\Phi_2, \Phi_6\}$ and $[\Phi_7]_{0.75} = \{\Phi_2, \Phi_7\}$. In a similar way, the R_{Fd} -0.75-block of each consequent is given as $[\Psi_i]_{0.75} = \{1, 3, 5, 7\}$ if i is odd and $[\Psi_j]_{0.75} = \{2, 4, 6\}$ if j is even. Therefore, $Dec_{0.75}^+(S) = \{r_1, r_3, r_4, r_6\}$. Consider the fourth item of Definition 15, we obtain that:

$$\begin{aligned} & \max\{|POS_{\mathcal{A}}(x) - \bigvee_{\Phi \rightarrow \Psi \in Dec_{0.75}^+(S)} \|\Phi\|_S(x)| \mid x \in U\} \\ &= \max\{|0.29 - 1|, |0.12 - 0.75|, |0.17 - 1|, |0.17 - 1|, |0.12 - 0.58|, |0.21 - 1|, \\ & \quad |0.21 - -0.54|\} \\ &= 0.83 \end{aligned}$$

As a consequence, for any value $\alpha_4 \geq 0.83$, we have that $Dec(S)$ preserves the 0.75-consistency of S with a degree α_4 . In conclusion, $Dec(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ -decision algorithm.

Now, we will compute some α -efficiencies corresponding to this decision algorithm, by using Definition 16.

- The 1-efficiency of the decision algorithm $Dec(S)$ is determined by the R_{Fd} -1-blocks. Since $\Phi \neq \Phi'$ for all $\Phi, \Phi' \in For(\mathcal{A})$ it is obtained that $[\Phi]_1 = \{\Phi\}$ for each $\Phi \in For(\mathcal{A})$. As a consequence,

$$\eta^1(Dec(S)) = \sum_{\Phi \rightarrow \Psi \in Dec(S)} \sigma_S^E(\Phi, \Psi) = 2.14$$

Notice that the efficiency is greater than 1 because the tolerance relations are not boolean, which leads us to conclude that the sum of the E -support of all decision rules of $Dec(S)$ is greater than $|U|$.

- The 0.75-efficiency of the decision algorithm $Dec(S)$ is determined by the R_{Fd} -0.75-blocks, which have been computed previously. From Table 3, it is easy to check that the decision rules with antecedents whose R_{Fd} -0.75-blocks are not singleton and can be ranked in order of E -strength, from the greatest to the lowest, are r_5, r_7, r_2 and r_6 . As a consequence, the 0.75-efficiency of this decision algorithm $Dec(S)$ is:

$$\begin{aligned} \eta^{0.75}(Dec(S)) &= \sigma_S^E(\Phi_1, \Psi_1) + \sigma_S^E(\Phi_5, \Psi_5) + \sigma_S^E(\Phi_3, \Psi_3) + \sigma_S^E(\Phi_4, \Psi_4) \\ & \quad + \sigma_S^E(\Phi_5, \Psi_5) + \sigma_S^E(\Phi_2, \Psi_2) + \sigma_S^E(\Phi_7, \Psi_7) \\ &= 2.22 \end{aligned}$$

- Finally, we will compute the 0-efficiency. Since $R_{Fd}(\Phi, \Phi') \geq 0$, for each pair Φ, Φ' of antecedents of the decision rules of $Dec(S)$, we obtain that $[\Phi_i]_0 = \{\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \Phi_7\}$ for each $i \in \{1, 2, \dots, 7\}$. As a consequence,

$$\begin{aligned} \eta_\Phi^0(Dec(S)) &= \max\{\sigma_S(\Phi', \Psi') \mid \Phi' \rightarrow \Psi' \in Dec(S)\} \\ &= \sigma_S(\Phi_5, \Psi_5) \\ &= 0.37 \end{aligned}$$

for all $\Phi \in For(\mathcal{A})$. Therefore, the 0-efficiency of the decision algorithm $Dec(S)$ is

$$\begin{aligned}\eta^0(Dec(S)) &= \max\{\sigma_S^E(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in Dec(S)\} |U| \\ &= 0.37 \cdot 7 \\ &= 2.59\end{aligned}$$

□

From these computations, and taking into account that α -efficiency is a decreasing mapping in α , we can conclude that considering different levels of flexibility in the comparison of the antecedents of the decision rules of $Dec(S)$, given by the value of α , does not result in significant changes in the value of α -efficiency. The main reason is that the E -strength of all decision rules of $Dec(S)$ are similar, as it can be seen in Table 3. In addition, we have shown that the α -efficiency of a decision algorithm is not bounded by 1, as we remarked previously.

5 Conclusions and Further Work

In this paper, we have studied decision rules and decision algorithms in the fuzzy framework. For this purpose, we have introduced the generalization of classic notions of decision rule, support, strength and certainty. In addition, we have presented the fuzzy notions of decision algorithm and efficiency. Different examples have been worked in order to illustrate these new definitions.

As a future work, we will continue studying the relationships among these notions, both in the classical and fuzzy settings. Specifically, in the fuzzy case, we are interested in obtaining theoretical properties related to the given notions of support, strength, certainty and efficiency. For example, we will study new possible definitions of α -efficiency, in order to obtain as result a value in the unit interval. Furthermore, we will explore examples whose decision attribute is not boolean and real problems with many decision rules to check the usefulness of the introduced decision algorithms.

References

1. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. *SIGMOD Rec.* **22**(2), 207–216 (1993)
2. Baczyński, M., Jayaram, B.: An Introduction to Fuzzy Implications, pp. 1–35. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-69082-5_1
3. Cock, M.D., Cornelis, C., Kerre, E.E.: Fuzzy rough sets: the forgotten step. *IEEE Trans. Fuzzy Syst.* **15**(1), 121–130 (2007)
4. Cornelis, C., Medina, J., Verbiest, N.: Multi-adjoint fuzzy rough sets: definition, properties and attribute selection. *Int. J. Approximate Reasoning* **55**, 412–426 (2014)

5. de Sá, C.R., Azevedo, P., Soares, C., Jorge, A.M., Knobbe, A.: Preference rules for label ranking: mining patterns in multi-target relations. *Inf. Fusion* **40**, 112–125 (2018)
6. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. *Int. J. Gen Syst* **17**, 06 (1990)
7. Feng, T., Fan, H.-T., Mi, J.-S.: Uncertainty and reduction of variable precision multigranulation fuzzy rough sets based on three-way decisions. *Int. J. Approximate Reasoning* **85**, 36–58 (2017)
8. Li, J., Mei, C., Lv, Y.: Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction. *Int. J. Approximate Reasoning* **54**(1), 149–165 (2013)
9. Mardani, A., Nilashi, M., Antucheviciene, J., Tavana, M., Bausys, R., Ibrahim, O.: Recent fuzzy generalisations of rough sets theory: a systematic review and methodological critique of the literature. *Complexity* **2017**, 1–33 (2017)
10. Medina, J.: Towards multi-adjoint property-oriented concept lattices. In: Yu, J., Greco, S., Lingras, P., Wang, G., Skowron, A. (eds.) *RSKT 2010. LNCS (LNAI)*, vol. 6401, pp. 159–166. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16248-0_26
11. Medina, J.: Multi-adjoint property-oriented and object-oriented concept lattices. *Inf. Sci.* **190**, 95–106 (2012)
12. Medina, J.: Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Comput. Math/ Appl.* **64**(6), 1992–2002 (2012)
13. Morsi, N.N., Yakout, M.M.: Axiomatics for fuzzy rough sets. *Fuzzy Sets Syst.* **100**, 327–342 (1998)
14. Nakamura, A.: Fuzzy rough sets. *Note Multiple-Valued Logic Japan* **9**, 1–8 (1988)
15. Nanda, S., Majumdar, S.: Fuzzy rough sets. *Fuzzy Sets Syst.* **45**, 157–160 (1992)
16. Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L.: Efficient mining of association rules using closed itemset lattices. *Inf. Syst.* **24**(1), 25–46 (1999)
17. Pawlak, Z.: Information systems theoretical foundations. *Inf. Syst.* **6**(3), 205–218 (1981)
18. Pawlak, Z.: Rough sets. *Int. J. Comput. Inf. Sci.* **11**, 341–356 (1982)
19. Pawlak, Z.: *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell (1992)
20. Pawlak, Z.: Rough sets and decision algorithms. In: Ziarko, W., Yao, Y. (eds.) *RSCTC 2000. LNCS (LNAI)*, vol. 2005, pp. 30–45. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45554-X_3
21. Stawicki, S., Ślęzak, D., Janusz, A., Widz, S.: Decision bireducts and decision reducts - a comparison. *Int. J. Approximate Reasoning* **84**, 75–109 (2017)
22. Wang, C.Y.: Topological structures of l-fuzzy rough sets and similarity sets of l-fuzzy relations. *Int. J. Approximate Reasoning* **83**, 160–175 (2017)
23. Yadav, D., Chowdary, C.R.: Ooimasp: origin based association rule mining with order independent mostly associated sequential patterns. *Expert Syst. Appl.* **93**, 62–71 (2018)
24. Yao, Y.: Three-way decisions with probabilistic rough sets. *Inf. Sci.* **180**(3), 341–353 (2010)
25. Zedam, L., Bouremel, H., De Baets, B.: Left- and right-compatibility of order relations and fuzzy tolerance relations. *Fuzzy Sets and Systems* **360**, 65–81 (2019), Theme: Fuzzy Relations

Anexo C

Efficiency of fuzzy rough set decision algorithms

Efficiency of fuzzy rough set decision algorithms ^{*}

Fernando Chacón-Gómez, M. Eugenia Cornejo,
Jesús Medina, Eloísa Ramírez-Poussa

Department of Mathematics, University of Cádiz, Cádiz, Spain
{fernando.chacon,mariaeugenia.cornejo,
jesus.medina,eloisa.ramirez}@uca.es

Abstract. This paper addresses the study of decision algorithms given in Rough Set Theory. Considering the fuzzy framework, we present a generalized notion of efficiency, which is one of the most important notions associated with decision algorithms. This new notion can be used to know the usefulness of a fuzzy rough set decision algorithm, as well as to compare it with other fuzzy decision algorithms.

Keywords: Fuzzy Rough Set Theory, Decision Rules, Decision Algorithm, Efficiency.

1 Introduction

In rough set theory (RST), decision tables can be described from decision rules [8, 15–17, 19]. Decision rules are used to extract information and to infer conclusions in an efficient way. The notion of decision algorithm was proposed by Pawlak in [16], to establish the properties that must be satisfied by a set of decision rules in order to suitably describe the decision table. An essential measure associated with decision algorithms is the notion of efficiency. The greater the efficiency of a decision algorithm is, the better this decision algorithm represents the considered dataset. As a consequence, the notion of efficiency is very useful to compare decision algorithms.

RST has already been considered within the fuzzy paradigm in different papers [2, 3, 6, 7, 9–14, 18]. Continuing the study started in [1], this work addresses the study of efficiency of decision algorithms, within the fuzzy paradigm. We provide a new notion of efficiency for fuzzy rough sets decision algorithms. In addition, we analyze some interesting properties of this measure, proving that this novel notion of efficiency generalizes the usually considered one in the classical framework. We also include an example to illustrate the content.

^{*} Partially supported by the the 2014-2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

2 Fuzzy rough set decision algorithms

Datasets are represented as decision tables $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ in rough set theory, where U and $\mathcal{A}$ are non-empty sets of objects and attributes respectively, $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with the attribute $a \in \mathcal{A}_d$ over U , and $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a}: U \rightarrow V_a\}$. The structure of the decision tables where an attribute is highlighted allows a direct transformation of them to a set of decision rules, since in both settings certain conditions must be satisfied to carry out a specific decision. Hence, the decision rules can describe the decision table in logical terms, providing other perspective of decision tables [16] and facilitating the extraction of information. To do the mentioned transformation, we need to define a formal language. In the fuzzy framework, the set of formulas associated with $B \subseteq A$, denoted as $For(B)$, is the same as in the classical framework, that is, it is built from attribute-value pairs (a, v) , where $a \in B$ and $v \in V_a$, by using the conjunction and disjunction logical connectives. However, the interpretation of formulas is different.

Definition 1. Let S be a decision table and $B \subseteq \mathcal{A}$. Given $\Phi \in For(B)$, with $\Phi = (a, v)$, $a \in B$ and $v \in V_a$, the mapping $\|\Phi\|_S^E: U \rightarrow [0, 1]$ is defined, for each $x \in U$, as:

$$\|\Phi\|_S^E(x) = E_a(\bar{a}(x), v)$$

where $E_a: V_a \times V_a \rightarrow [0, 1]$ is a $[0, 1]$ -fuzzy tolerance relation, that is, a fuzzy reflexive and symmetric relation. Therefore, $\|\Phi\|_S^E(x)$ represents how much the object x satisfies the formula Φ , through the relationship between the value of the attribute a in the object x and the value of the attribute a in the formula Φ .

Given $\Phi, \Psi \in For(B)$, $\|\Phi \wedge \Psi\|_S^E(x)$ and $\|\Phi \vee \Psi\|_S^E(x)$ represent how much the object x satisfies the formula $\Phi \wedge \Psi$ and formula $\Phi \vee \Psi$, respectively, both are defined as follows:

$$\begin{aligned} \|\Phi \wedge \Psi\|_S^E(x) &= \min\{\|\Phi\|_S^E(x), \|\Psi\|_S^E(x)\} \\ \|\Phi \vee \Psi\|_S^E(x) &= \max\{\|\Phi\|_S^E(x), \|\Psi\|_S^E(x)\} \end{aligned}$$

Decision rules are expressions $\Phi \rightarrow \Psi$, where $\Phi \in For(B), \Psi \in For(\{d\})$. In the following, we need to consider a family of $[0, 1]$ -fuzzy tolerance relations, known as *separable relations* [20], in order to introduce the notions of *support* and *strength* of a decision rule in the fuzzy framework.

Definition 2. Let S be a decision table, $\Phi \rightarrow \Psi$ be a decision rule in S and $E = \{E_a: V_a \times V_a \rightarrow [0, 1] \mid a \in \mathcal{A}_d\}$ be a family of $[0, 1]$ -fuzzy tolerance relations such that $E_a(\bar{a}(x), v) = 1$ if and only if $\bar{a}(x) = v$.

- *E-support* of the decision rule $\Phi \rightarrow \Psi$ is $\text{supp}_S^E(\Phi, \Psi) = \text{card}_F(\|\Phi \wedge \Psi\|_S^E)$ where $\text{card}_F(\cdot)$ represents the cardinal of a fuzzy set [5].
- *E-strength* of the decision rule $\Phi \rightarrow \Psi$ is $\sigma_S^E(\Phi, \Psi) = \frac{\text{supp}_S^E(\Phi, \Psi)}{|U|}$ where $|\cdot|$ represents the cardinal of a set.

These measures are key to study decision algorithms and their efficiency. A decision algorithm is a set of decision rules verifying certain conditions, whose formal definition in the fuzzy case is given below. First, we need to define a fuzzy tolerance relation which allows us to compare formulas.

Definition 3. *Given a decision table S , we define the $[0, 1]$ -fuzzy tolerance relation R_{Fd} : $For(\mathcal{A}_d) \times For(\mathcal{A}_d) \rightarrow [0, 1]$, for all $\Phi, \Phi' \in For(\mathcal{A}_d)$, as follows:*

$$R_{Fd}(\Phi, \Phi') = 1 - \max\{|\|\Phi\|_S^E(x) - \|\Phi'\|_S^E(x)| \mid x \in U\}$$

Given $\alpha \in [0, 1]$, the R_{Fd} - α -block is defined for each $\Phi \in For(\mathcal{A}_d)$ as:

$$[\Phi]_\alpha = \{\Phi' \in For(\mathcal{A}_d) \mid \alpha \leq R_{Fd}(\Phi, \Phi')\}$$

Definition 4. *Let S be a decision table, $\alpha > 0$, $\alpha_1, \alpha_2, \alpha_4 \in [0, 1]$, $\alpha_3 \geq 0$ and $Dec(S) = \{\Phi_i \rightarrow \Psi_i\}_{i \in I}$ be a set of decision rules of S , where the index set is $I = \{1, \dots, m\}$ and $m \geq 2$. We say that:*

1. *$Dec(S)$ is a set of $\alpha_1 \alpha_2$ -pairwise mutually exclusive (independent) decision rules, if each pair of decision rules $\Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S)$ satisfies that $\Phi = \Phi'$ or $\|\Phi \wedge \Phi'\|_S^E(x) \leq \alpha_1$ and $\Psi = \Psi'$ or $\|\Psi \wedge \Psi'\|_S^E(x) \leq \alpha_2$, for all $x \in U$.*
2. *$Dec(S)$ covers U , if $|\bigvee_{i=1}^m \Phi_i|_S^E = |\bigvee_{i=1}^m \Psi_i|_S^E = |U|$.*
3. *The decision rule $\Phi \rightarrow \Psi \in Dec(S)$ is α_3 -admissible in S if $\alpha_3 < \text{supp}_S^E(\Phi, \Psi)$.*
4. *$Dec(S)$ preserves the α -consistency of S with a degree α_4 if the next inequality holds for all $x \in U$:*

$$|POS_{\mathcal{A}}(x) - \bigvee_{\Phi \rightarrow \Psi \in Dec_\alpha^+(S)} \|\Phi\|_S^E(x)| \leq \alpha_4$$

where $\alpha \in [0, 1]$ and $Dec_\alpha^+(S) = \{\Phi \rightarrow \Psi \in Dec(S) \mid \text{if for each } \Phi' \rightarrow \Psi' \in Dec(S) \text{ such that } \Phi' \in [\Phi]_\alpha, \text{ then } \Psi' \in [\Psi]_\alpha\}$

The set of decision rules $Dec(S)$ satisfying the previous properties for the values $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ is called $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm in S .

These notions were already studied in [1], where the readers can find more information, properties and illustrative examples about this research topic. In addition, it is important to emphasize that the value α must be fixed at the beginning depending on the flexibility looked for. After that, the rest of values can be computed to obtain a decision algorithm. The values α_1, α_2 set upper bounds to the degree of satisfaction of each pair of antecedents and consequents for each object of the decision table, the value α_3 provides a lower bound of the E -support of all decision rules of $Dec(S)$ and the value α_4 measures the equivalence between the consistency provided by the fuzzy positive region and the α -consistency provided by the antecedents of decision rules of $Dec_\alpha^+(S)$.

3 Efficiency of decision algorithms

This section addresses the study of the efficiency of fuzzy rough set decision algorithms, recalling its formal definition to the fuzzy framework and presenting different properties related to this notion.

Definition 5. Let S be a decision table and $Dec(S)$ a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm. Given $0 \leq \varepsilon \leq 1$, we call ε -efficiency of the decision algorithm $Dec(S)$ to the number

$$\eta^\varepsilon(Dec(S)) = \sum_{\Phi \in For(\mathcal{A})} \eta_\Phi^\varepsilon(Dec(S))$$

where $\eta_\Phi^\varepsilon(Dec(S)) = \max\{\sigma_S^E(\Phi', \Psi') \mid \Phi \rightarrow \Psi, \Phi' \rightarrow \Psi' \in Dec(S), \Phi' \in [\Phi]_\varepsilon\}$.

This notion is indispensable to measure the usefulness of a decision algorithm and compare with other algorithms as well. The following proposition includes some properties associated with the notion of ε -efficiency.

Proposition 1. Let S be a decision table and $Dec(S)$ be a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_\alpha$ -decision algorithm and $m \geq 2$. The following properties hold:

1. If $E = \{E_a: V_a \times V_a \rightarrow \{0, 1\} \mid a \in \mathcal{A}_d\}$ is a family of boolean tolerance relations and $\varepsilon > 0$ then $\eta^\varepsilon(Dec(S)) = \eta(Dec(S))$.¹
2. The ε -efficiency is decreasing in ε .
3. If $\max\{R_{Fd}(\Phi_i, \Phi_j) \mid \Phi_i \neq \Phi_j, i, j \in \{1, 2, \dots, m\}\} < \varepsilon$ and $\Phi_i \neq \Phi_j$ when $i \neq j$, then the ε -efficiency of the decision algorithm $Dec(S)$ is given as

$$\eta^\varepsilon(Dec(S)) = \sum_{\Phi \rightarrow \Psi \in Dec(S)} \sigma_S^E(\Phi, \Psi)$$

4. Let $\Phi_j \rightarrow \Psi_j$ be a decision rule in $Dec(S)$, with $j \in \{1, 2, \dots, m\}$, such that

$$\sigma_S^E(\Phi_j, \Psi_j) = \max\{\sigma_S^E(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in Dec(S)\}$$

If $\varepsilon \leq \min\{R_{Fd}(\Phi_i, \Phi_j) \mid i \in \{1, 2, \dots, m\}\}$ then the ε -efficiency of the decision algorithm $Dec(S)$ is given as

$$\eta^\varepsilon(Dec(S)) = \sigma_S^E(\Phi_j, \Psi_j) \cdot |\{\Phi \mid \Phi \rightarrow \Psi \in Dec(S)\}|$$

Next remarks are needed to emphasize the importance of these properties.

- From the first item in Proposition 1, we can conclude that Definition 5 generalizes the notion of efficiency given in the classical case. Moreover, it is convenient to mention that the condition $\varepsilon > 0$ is required because otherwise all the antecedents would be considered as equals.

¹ The notion of efficiency of a decision algorithm in the classical case, $\eta(Dec(S))$, can be found in [16].

- With respect to the second item in Proposition 1, the comparison of the antecedents of the decision rules in $Dec(S)$ becomes more restrictive, when the value of ε increases, since the cardinal of R_{Fd} - ε -blocks decreases. Therefore, it is expected that the ε -efficiency decreases. This property is fundamental since the user must choose a balanced value of ε to obtain a good value of ε -efficiency while a good level of flexibility is considered, which is given by ε .
- From the third item in Proposition 1, if the fixed value ε allows to discern all antecedents of decision rules in $Dec(S)$, then only one decision rule is considered in the computation of each addend of the ε -efficiency. For this reason and the previous property, this ε -efficiency is the minimum possible value for the decision algorithm $Dec(S)$.
- The last item states that when the antecedent of the decision rule with the greatest value of E -strength is R_{Fd} - ε -related with the rest of antecedents, this decision rule is considered in each addend of the ε -efficiency. As a consequence, the computation of ε -efficiency boils down to the product of E -strength times the number of different antecedents in $Dec(S)$. In addition, under the hypothesis required in the fourth item, we can ensure the decision rule $\Phi_j \rightarrow \Psi_j$ is the most representative of the table. Therefore, this ε -efficiency is the maximum possible value for $Dec(S)$.

The following example continues the one given in [1] in order to illustrate Definition 5 and Proposition 1.

Example 1. Consider the decision table $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ where the set of objects is $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, the set of attributes is $\mathcal{A} = \{a_1, a_2, a_3\}$ and $\alpha = 0.75$. This decision table is represented in Table 1, together with the E -Support and E -strength of the decision rules given by each object and the relation between each pair of antecedents of decision rules. The considered decision rules can be found in [1] and they are easily obtained from the row of S .

Taking $\alpha_1 \geq 0.78$, $\alpha_2 \geq 0$, $\alpha_3 < 1.61$ and $\alpha_4 \geq 0.83$, we obtain that $Dec(S)$ is a $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)_{0.75}$ -decision algorithm. Now, we compute the ε -efficiency of $Dec(S)$ by using Definition 5, Proposition 1 and considering different values of ε .

- To begin with, we will compute the 0.75-efficiency of $Dec(S)$ by using Definition 5. From the relation R_{Fd} given in Table 1 it is easy to obtain the R_{Fd} -0.75-blocks of the antecedents:

$$\begin{aligned}
 [\Phi_1]_{0.75} &= \{\Phi_1\} \\
 [\Phi_2]_{0.75} &= \{\Phi_2, \Phi_5, \Phi_6, \Phi_7\} & [\Phi_4]_{0.75} &= \{\Phi_4\} & [\Phi_6]_{0.75} &= \{\Phi_2, \Phi_6\} \\
 [\Phi_3]_{0.75} &= \{\Phi_3\} & [\Phi_5]_{0.75} &= \{\Phi_2, \Phi_5\} & [\Phi_7]_{0.75} &= \{\Phi_2, \Phi_7\}
 \end{aligned}$$

Moreover, the decision rules with the antecedents whose R_{Fd} -0.75-blocks are not singleton and ranked in order of E -strength from the highest to the

Table 1: Decision table, E -support, E -strength and relation R_{Fd} of Example 1.

S	a_1	a_2	a_3	d	Rules	$supp_S^E$	σ_S^E
x_1	0.34	0.31	0.75	0	r_1	2.19	0.31
x_2	0.21	0.71	0.5	1	r_2	2.11	0.3
x_3	0.52	0.92	1	0	r_3	2.14	0.31
x_4	0.85	0.65	1	1	r_4	1.61	0.23
x_5	0.43	0.89	0.5	0	r_5	2.58	0.37
x_6	0.21	0.47	0.25	1	r_6	2	0.29
x_7	0.09	0.93	0.25	0	r_7	2.29	0.33

R_{Fd}	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
Φ_1	1	0.6	0.39	0.49	0.42	0.5	0.38
Φ_2	0.6	1	0.5	0.36	0.78	0.75	0.75
Φ_3	0.39	0.5	1	0.67	0.5	0.25	0.25
Φ_4	0.49	0.36	0.67	1	0.5	0.25	0.24
Φ_5	0.42	0.78	0.5	0.5	1	0.58	0.66
Φ_6	0.5	0.75	0.25	0.25	0.58	1	0.54
Φ_7	0.38	0.75	0.25	0.24	0.66	0.54	1

lowest are r_5, r_7, r_2 and r_6 . Therefore, the 0.75-efficiency of $Dec(S)$ is

$$\begin{aligned}
\eta^{0.75}(Dec(S)) &= \sigma_S^E(\Phi_1, \Psi_1) + \max\{\sigma_S^E(\Phi_2, \Psi_2), \sigma_S^E(\Phi_5, \Psi_5), \sigma_S^E(\Phi_6, \Psi_6), \sigma_S^E(\Phi_7, \Psi_7)\} \\
&\quad + \sigma_S^E(\Phi_3, \Psi_3) + \sigma_S^E(\Phi_4, \Psi_4) + \max\{\sigma_S^E(\Phi_2, \Psi_2), \sigma_S^E(\Phi_5, \Psi_5)\} \\
&\quad + \max\{\sigma_S^E(\Phi_2, \Psi_2), \sigma_S^E(\Phi_6, \Psi_6)\} + \max\{\sigma_S^E(\Phi_2, \Psi_2), \sigma_S^E(\Phi_7, \Psi_7)\} \\
&= \sigma_S^E(\Phi_1, \Psi_1) + \sigma_S^E(\Phi_5, \Psi_5) + \sigma_S^E(\Phi_3, \Psi_3) + \sigma_S^E(\Phi_4, \Psi_4) + \\
&\quad + \sigma_S^E(\Phi_5, \Psi_5) + \sigma_S^E(\Phi_2, \Psi_2) + \sigma_S^E(\Phi_7, \Psi_7) \\
&= 2.22
\end{aligned}$$

- To exemplify the third item in Proposition 1, we observe that there are no decision rules with the same antecedent (see [1]), and by using the relation R_{Fd} given in Table 1, we can take:

$$\varepsilon > \max\{R_{Fd}(\Phi_i, \Phi_j) \mid \Phi_i \neq \Phi_j, i, j \in \{1, 2, \dots, 7\}\} = 0.78$$

in order to apply Proposition 1, obtaining that the ε -efficiency of $Dec(S)$ is:

$$\eta^\varepsilon(Dec(S)) = \sum_{\Phi \rightarrow \Psi \in Dec(S)} \sigma_S^E(\Phi, \Psi) = 2.14$$

- Finally, we illustrate the fourth item in Proposition 1. From Table 1, the decision rule in $Dec(S)$ with the greatest value of E -strength is r_5 , that is, $\sigma_S^E(\Phi_5, \Psi_5) = 0.37$. On the other hand, from Table 1 again, we have that:

$$\min\{R_{Fd}(\Phi_i, \Phi_5) \mid i \in \{1, 2, \dots, 7\}\} = R_{Fd}(\Phi_1, \Phi_5) = 0.42$$

As a consequence, if $\varepsilon \leq 0.42$, by using Proposition 1, we obtain that

$$\begin{aligned}
\eta^\varepsilon(Dec(S)) &= \max\{\sigma_S^E(\Phi, \Psi) \mid \Phi \rightarrow \Psi \in Dec(S)\} |\{\Phi \mid \Phi \rightarrow \Psi \in Dec(S)\}| \\
&= 0.37 \cdot 7 = 2.59
\end{aligned}$$

□

Attending to these computations, we obtain that the minimum and maximum value of ε -efficiency are similar. Hence, the fact of carrying out a less flexible study of the ε -efficiency of the decision algorithm $Dec(S)$, that is, considering a higher value of ε , does not have a great effect on the obtained results. This aspect is quite important since it will allow to discern decision rules while maintaining the ε -efficiency of $Dec(S)$ without significant changes.

4 Conclusions and further work

In this paper, we have provided a new notion of efficiency, which can be used to measure how well a fuzzy rough set decision algorithm describes a dataset. We have proved that this notion generalizes the classical one and we have presented some interesting properties. We also have included an example to illustrate the generalized notion and its properties.

As future work, we want to continue studying more properties about decision algorithms in the fuzzy environment, as well as applying all the obtained results in real datasets. Taking into account the existing connection between the given notions of E -support and E -strength and those of support and confidence from fuzzy association rule mining, we will consider the alternative measures introduced in [4] and investigate how they could impact in our work.

References

1. F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Fuzzy rough set decision algorithms. *Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022)*, 2022. In press.
2. M. D. Cock, C. Cornelis, and E. E. Kerre. Fuzzy rough sets: The forgotten step. *Fuzzy Systems, IEEE Transactions on*, 15(1):121–130, feb. 2007.
3. C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.
4. M. De Cock, C. Cornelis, and E. Kerre. Elicitation of fuzzy association rules from positive and negative examples. *Fuzzy Sets and Systems*, 149(1):73–85, 2005. Fuzzy Sets in Knowledge Discovery.
5. 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.
6. D. DUBOIS and H. PRADE. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17(2-3):191–209, 1990.
7. T. Feng, H.-T. Fan, and J.-S. Mi. Uncertainty and reduction of variable precision multigranulation fuzzy rough sets based on three-way decisions. *International Journal of Approximate Reasoning*, 85:36–58, 2017.
8. 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.

9. J. Medina. Towards multi-adjoint property-oriented concept lattices. *Lecture Notes in Artificial Intelligence*, 6401:159–166, 2010.
10. J. Medina. Multi-adjoint property-oriented and object-oriented concept lattices. *Information Sciences*, 190:95–106, 2012.
11. J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
12. N. N. Morsi and M. M. Yakout. Axiomatics for fuzzy rough sets. *Fuzzy sets and Systems*, 100:327–342, 1998.
13. A. Nakamura. Fuzzy rough sets. *Note on Multiple-Valued Logic in Japan*, 9:1–8, 1988.
14. S. Nanda and S. Majumdar. Fuzzy rough sets. *Fuzzy sets and Systems*, 45:157–160, 1992.
15. Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992.
16. Z. Pawlak. Rough sets and decision algorithms. In W. Ziarko and Y. Yao, editors, *Rough Sets and Current Trends in Computing*, pages 30–45, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
17. S. Stawicki, D. Ślęzak, A. Janusz, and S. Widz. Decision bireducts and decision reducts - a comparison. *International Journal of Approximate Reasoning*, 84:75–109, 2017.
18. C. Y. Wang. Topological structures of l-fuzzy rough sets and similarity sets of l-fuzzy relations. *International Journal of Approximate Reasoning*, 83:160–175, 2017.
19. Y. Yao. Three-way decisions with probabilistic rough sets. *Information Sciences*, 180(3):341 – 353, 2010.
20. 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.

Anexo D

Relating decision rules and attribute implications

Relating decision rules and attribute implications

Fernando Chacón-Gómez^{1,*}, M. Eugenia Cornejo¹ and Jesús Medina¹

¹Department of Mathematics, University of Cádiz, Cádiz, Spain

Abstract

Formal Concept Analysis (FCA) and Rough Set Theory (RST) are two useful mathematical tools for extracting information, which have different philosophies. One important issue in data science is to obtain relationships among the main variables in the considered data set. Both theories, FCA and RST, independently study these relations by means of attribute implications in FCA and decision rules in RST. This paper introduces a preliminary comparison between these approaches, focused on the notions of valid attribute implication in FCA and true decision rule in RST.

Keywords

Rough Set Theory, Formal Concept Analysis, Decision Rules, Attribute implications

1. Introduction

Formal Concept Analysis (FCA) was introduced by Wille [19] in the eighties, as a mathematical tool for the extraction of knowledge from datasets, interpreted as a context composed by objects, attributes and a binary relation between them. From this relation and the derivation operators, pieces of information called concepts are obtained, which are pairs composed of a subset of objects and a subset of attributes, where each subset unequivocally determines the other one. An interesting topic in FCA is the study of attribute implications [5, 6, 7, 8, 9, 16] relating the attributes in the dataset (context) for obtaining information from the context, which has a different philosophy from decision rules in Rough Set Theory (RST).

RST was proposed by Pawlak [12, 13] also in the eighties as a formal tool to analyze datasets with imprecise or incomplete information. This theory interprets a relational dataset as a decision table, which contains a collection of objects, attributes and mappings characterizing the objects by using the attributes. In RST, the notion of decision rule [10, 14, 15, 18, 20] arises as a useful tool for the management of information from relational datasets, characterizing decision tables and allowing the extraction of significant conclusions thanks to their logical terms. Different measures have been defined in order to describe decision rules and provide useful information, such as, the certainty which represents a conditional probability.

Both theories, FCA and RST, have been related in diverse papers [1, 4, 3, 11, 21]. However, to the best of our knowledge, none of them have studied the relationship among decision rules and

Published in Pablo Cordero, Ondrej Kridlo (Eds.): The 16th International Conference on Concept Lattices and Their Applications, CLA 2022, Tallinn, Estonia, June 20–22, 2022, Proceedings, pp. 69–78.

*Corresponding author.

✉ fernando.chacon@uca.es (F. Chacón-Gómez); mariaeugenia.cornejo@uca.es (M. E. Cornejo);
jesus.medina@uca.es (J. Medina)

ORCID 0000-0001-9710-7813 (F. Chacón-Gómez); 0000-0002-8230-0044 (M. E. Cornejo); 0000-0002-3931-5873 (J. Medina)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

attribute implications. This paper introduces a preliminary comparison between these notions in order to establish new bridges allowing that the developments in FCA can enrich RST, and vice versa. Moreover, an example to illustrate the obtained relation and some particularities is introduced. The main conclusions and next steps are introduced at the end of the paper.

2. Preliminaries

This section includes some preliminary definitions of Formal Concept Analysis [7, 8] and Rough Set Theory [14, 15] in order to make the paper self-contained.

2.1. Formal Concept Analysis

Datasets are represented as formal contexts in FCA. From a formal context, derivation operators are defined in order to characterize a subset of attributes by using a subset of objects and vice versa. These derivation operators are essential to determine the basic units of knowledge of the contexts, that is, the concepts.

Definition 1. A context is a tuple $(A, B, \mathcal{R})$ such that A and B are non-empty sets of attributes and objects, respectively, and $\mathcal{R}$ is a relation between them.

Definition 2. Given a context $(A, B, \mathcal{R})$, the derivation operators $(\uparrow, \downarrow)$ are the mappings $\uparrow : 2^B \rightarrow 2^A$ and $\downarrow : 2^A \rightarrow 2^B$ defined, for each $X \subseteq B$ and $Y \subseteq A$, as:

$$\begin{aligned} X^\uparrow &= \{a \in A \mid a\mathcal{R}b \text{ for each } b \in X\} \\ Y^\downarrow &= \{b \in B \mid a\mathcal{R}b \text{ for each } a \in Y\} \end{aligned}$$

The set $X^\uparrow$ is called intension of X and $Y^\downarrow$ is called extension of Y . Given $X \subseteq B$ and $Y \subseteq A$, we will say that (X, Y) is a concept if $X = Y^\downarrow$ and $Y = X^\uparrow$.

From now on, the set of concepts of a given context $(A, B, \mathcal{R})$ will be denoted as $\mathfrak{B}(A, B, \mathcal{R})$. In the following, we recall some important notions related to attribute implications and their validity.

Definition 3. Let Λ be an index set, $(A, B, \mathcal{R})$ be a context and $C, D, T \subseteq A$.

- An attribute implication is denoted as $C \Rightarrow D$. The set C is called antecedent of the attribute implication and the set D is called consequent.
- T respects an attribute implication $C \Rightarrow D$, if $C \not\subseteq T$ or $D \subseteq T$.
- $C \Rightarrow D$ is a valid attribute implication in $\{T_i \subseteq A \mid i \in \Lambda\}$, if each T_i respects the attribute implication $C \Rightarrow D$.

The notion of valid attribute implication is extended to formal contexts as it is shown below.

Definition 4. Let $(A, B, \mathcal{R})$ be a context and $C, D \subseteq A$. We say that $C \Rightarrow D$ is a valid attribute implication in $(A, B, \mathcal{R})$, if $C \Rightarrow D$ is a valid attribute implication in $\{b^\uparrow \mid b \in B, (b^{\uparrow\downarrow}, b^\uparrow) \in \mathfrak{B}(A, B, \mathcal{R})\}$.

Finally, we recall a useful property that allows us to obtain valid attribute implications in a context.

Proposition 1. *Let $(A, B, \mathcal{R})$ be a context and $C, D \subseteq A$. Then, $C \Rightarrow D$ is a valid attribute implication in $(A, B, \mathcal{R})$ if and only if $D \subseteq C^{\uparrow\uparrow}$.*

2.2. Rough Set Theory

Datasets are represented as decision tables in RST. Although we will only consider one decision attribute in decision tables, similar studies can be carried out in an analogous way when more than one decision attribute is considered.

Definition 5. *Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A decision table is a tuple $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, where V_a is the set of values associated with the attribute $a \in \mathcal{A}_d$ over U , and $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a} : U \rightarrow V_a\}$. In this case, the attributes of $\mathcal{A}$ are called condition attributes and d is called a decision attribute.*

Decision tables describe decisions in terms of conditions that must be satisfied in order to carry out the decision specified in the decision table [15]. In this way, decision tables can be seen as a logical description of approximation of decisions. Furthermore, a set of decision rules can be associated with every decision table, describing the decision table in logical terms. Hence, decision rules allow an easier interpretation of decision tables. In the following, we recall the formal language in which decision rules are based on.

Definition 6. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. The set of formulas associated with B , denoted as $\text{For}(B)$, is built from attribute-value pairs (a, v) , where $a \in B$ and $v \in V_a$, by means of the conjunction and disjunction logical connectives, $\wedge$ and $\vee$, respectively.*

For each $\Phi \in \text{For}(B)$, with $\Phi = (a, v)$, the set of objects $x \in U$ that satisfies Φ in S is defined as:

$$\|\Phi\|_S = \|(a, v)\|_S = \{x \in U \mid \bar{a}(x) = v\}$$

Given $\Phi, \Psi \in \text{For}(B)$, the set of objects that satisfies $\Phi \wedge \Psi$ in S is defined as $\|\Phi \wedge \Psi\|_S = \|\Phi\|_S \cap \|\Psi\|_S$ and the set of objects that satisfies $\Phi \vee \Psi$ in S is defined as $\|\Phi \vee \Psi\|_S = \|\Phi\|_S \cup \|\Psi\|_S$.

Once the formal language associated with a subset of attributes has been defined, we can introduce the notion of a decision rule.

Definition 7. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $B \subseteq \mathcal{A}$. A decision rule in S is an expression $\Phi \rightarrow \Psi$ where $\Phi \in \text{For}(B)$, $\Psi \in \text{For}(\{d\})$ where B and $\{d\}$ are condition and decision attributes, respectively. In addition, we say that an object $x \in U$ satisfies a decision rule $\Phi \rightarrow \Psi$ if $x \in \|\Phi \wedge \Psi\|_S$.*

Finally, we introduce the notion of *certainty* of a decision rule, which provides us with the conditional probability of the consequent given the antecedent.

Definition 8. *Let $S = (U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ be a decision table and $\Phi \rightarrow \Psi$ be a decision rule in S , with $\Phi \in \text{For}(B)$ and $\Psi \in \text{For}(\{d\})$. We call certainty of the decision rule $\Phi \rightarrow \Psi$ to the value:*

$$\text{cer}_S(\Phi, \Psi) = \frac{|\|\Phi \wedge \Psi\|_S|}{|\|\Phi\|_S|}$$

We will say that $\Phi \rightarrow \Psi$ is a true decision rule, if $cer_S(\Phi, \Psi) = 1$. If $cer_S(\Phi, \Psi) = 0$, we say that the decision rule is false. Otherwise, it will be called a not entirely true decision rule.

3. Relationship between true decision rules and valid attribute implications

This section focuses on comparing decision rules in RST and attribute implications in FCA, in order to obtain a relationship between them. Before carrying out this comparison, it is convenient to note that given a set of objects and a set of attributes, they are always related in some way in RST, thanks to the value given by the mapping associated to each attribute of the decision table, while an object and an attribute may be related or not in FCA. Moreover, we must take into account that, although FCA can consider multi-valued contexts [8], boolean relations are usually considered in classical FCA. Hence, in this preliminary study, we will consider boolean decision tables, following the translation given in [3] in order to obtain a decision table from a context and vice versa.

Definition 9. Let U and $\mathcal{A}$ be non-empty sets of objects and attributes, respectively. A boolean decision table is a tuple $S_{\mathcal{B}} = (U, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ such that $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ with $d \notin \mathcal{A}$ and $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a} : U \rightarrow \{0, 1\}\}$.

Considering Definitions 1 and 9, we can obtain a boolean decision table from a given context as follows. Given a context $(A, B, \mathcal{R})$, we construct the boolean decision table $S_{\mathcal{B}} = (B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$, with an outstanding attribute $d \in A$, defining the set $\overline{\mathcal{A}_d}$ by means of the mappings $\bar{a} : B \rightarrow \{0, 1\}$ defined, for each $a \in \mathcal{A}_d$ and $b \in B$, as follows:

$$\bar{a}(b) = \begin{cases} 1 & \text{if } a\mathcal{R}b \\ 0 & \text{otherwise} \end{cases}$$

Reciprocally, given a boolean decision table $S_{\mathcal{B}} = (U, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ we define the context $(\mathcal{A}_d, U, \mathcal{R})$, where the relation R is defined as $a\mathcal{R}b$, if $\bar{a}(b) = 1$, for all attribute $a \in \mathcal{A}_d$ and object $b \in U$.

Once exposed the mechanisms to obtain a boolean decision table from a context and vice versa, we proceed to relate true decision rules in RST to valid attribute implications in FCA.

Proposition 2. Let $S_{\mathcal{B}} = (B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ be a boolean decision table, $(A, B, \mathcal{R})$ be a context and $C, D \subseteq A$, such that $C = \{a_1, \dots, a_n\}$ and $D = d$. Given a decision rule $\Phi \rightarrow \Psi$ in $S_{\mathcal{B}}$ satisfied by at least one object $x \in B$, where $\Phi = (a_1, 1) \wedge \dots \wedge (a_n, 1)$ and $\Psi = (d, 1)$, the following equivalence holds:

$$cer_{S_{\mathcal{B}}}(\Phi, \Psi) = 1 \text{ if and only if } C \Rightarrow D \text{ is a valid attribute implication in } (A, B, \mathcal{R})$$

The decision rules in which all the attributes (conditions and decision attributes) take the value 1 will be called *1-decision rule*.

As a consequence of the previous proposition, if a dataset is interpreted as a boolean decision table and a true 1-decision rule is obtained, a valid attribute implication is deduced. In contrast,

if the dataset is studied from the FCA perspective and a valid attribute implication is extracted, where at least one object is related with all the attributes of the implication, a true decision rule in RST is obtained. In short, regardless of the framework worked, it is possible to obtain significant conclusions in both of them simultaneously.

On the other hand, it is also possible to relate true 1-decision rules $\Phi \rightarrow \Psi = (a_1, 1) \wedge \dots \wedge (a_n, 1) \rightarrow (d, 1)$ to valid attribute implications $C \Rightarrow D$ in the case that $C^\downarrow \neq \emptyset$. This occurs since if an object x belongs to $C^\downarrow$ then it is related to all the attributes of C . Hence, x satisfies the antecedent Φ of the decision rule and, whether decision rule $\Phi \rightarrow \Psi$ is true or the attribute implication $C \Rightarrow D$ is valid it is obtained that x satisfies the consequent Ψ . Then, it is equivalent to require that the decision rule $\Phi \rightarrow \Psi$ is satisfied by at least one object than to require that this object belongs to $C^\downarrow$. As a consequence, the following result arises.

Corollary 1. Let $S_{\mathcal{B}} = (B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ be a boolean decision table, $(A, B, \mathcal{R})$ be a context and $C, D \subseteq A$, such that $C = \{a_1, \dots, a_n\}$ and $D = d$. Given an attribute implication $C \Rightarrow D$ with $C^\downarrow \neq \emptyset$, the following equivalence holds:

$$cer_{S_{\mathcal{B}}}(\Phi, \Psi) = 1 \text{ if and only if } C \Rightarrow D \text{ is a valid attribute implication in } (A, B, \mathcal{R})$$

where $\Phi \rightarrow \Psi$ is a 1-decision rule.

Notice that, by the philosophy in FCA, only attributes with value 1 can be considered in decision rules to be compared with attribute implications. Hence, the possibility of taking into account that decision rules is to consider mixed contexts, that is, positive and negative attributes, as it is considered in [2, 6, 17]. This extension will be studied in the future. Now, we illustrate Proposition 2 and Corollary 1 in the following example.

Example 1. Consider the context $(A, B, \mathcal{R})$ where the set of attributes is $A = \{a_1, a_2, a_3, a_4, a_5\}$, the set of objects $B = \{b_1, b_2, b_3, b_4, b_5, b_6, b_7\}$ and the relation $\mathcal{R}$ is given in Table 1.

Table 1

Context $(A, B, \mathcal{R})$ of Example 1.

$\mathcal{R}$	a_1	a_2	a_3	a_4	a_5
b_1		×		×	
b_2		×			×
b_3			×		
b_4	×	×	×	×	
b_5				×	
b_6		×	×		
b_7					×

As it was commented previously, we can obtain different boolean decision tables $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$, depending on the considered outstanding attribute $d \in A$. In order to illustrate Proposition 2, the set of attributes $\mathcal{A} = \{a_1, a_2, a_4, a_5\}$ and the decision attribute $d = a_3$ will be fixed. The boolean decision table $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ obtained following the mechanism previously explained is represented in Table 2.

Now, we consider all decision rules of Table 2 by using $\{a_1, a_2\}$ as condition attributes and $d = a_3$ as decision attribute. One decision rule is extracted for each $b \in B$. However, the pairs

Table 2Boolean decision table $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ of Example 1

	a_1	a_2	a_4	a_5	d
b_1	0	1	1	0	0
b_2	0	1	0	1	0
b_3	0	0	0	0	1
b_4	1	1	1	0	1
b_5	0	0	1	0	0
b_6	0	1	0	0	1
b_7	0	0	0	1	0

of objects b_1, b_2 and b_5, b_7 generate the same decision rule. Therefore, five decision rules are obtained, which are:

$$\begin{aligned}
r_1 &: (a_1, 0) \wedge (a_2, 1) \rightarrow (d, 0) \\
r_2 &: (a_1, 0) \wedge (a_2, 0) \rightarrow (d, 1) \\
r_3 &: (a_1, 1) \wedge (a_2, 1) \rightarrow (d, 1) \\
r_4 &: (a_1, 0) \wedge (a_2, 0) \rightarrow (d, 0) \\
r_5 &: (a_1, 0) \wedge (a_2, 1) \rightarrow (d, 1)
\end{aligned}$$

Notice that, both r_1 and r_4 are satisfied by two objects, and therefore, the number of obtained decision rules are lesser than the number of objects. Attending to the obtained decision rules, we can ensure that r_3 is the unique 1-decision rule. In order to illustrate the mentioned proposition, we will compute the certainty of r_3 , being $\Phi_3 = (a_1, 1) \wedge (a_2, 1)$ and $\Psi_3 = (d, 1)$. By Definition 8, its certainty is

$$cer_{S_{\mathcal{B}}}(\Phi_3, \Psi_3) = \frac{|\|\Phi_3 \wedge \Psi_3\|_{S_{\mathcal{B}}}|}{|\|\Phi_3\|_{S_{\mathcal{B}}}|} = \frac{1}{1} = 1$$

On the other hand, notice that

$$\begin{aligned}
\|\Phi_3\|_{S_{\mathcal{B}}} &= \|(a_1, 1) \wedge (a_2, 1)\|_{S_{\mathcal{B}}} = \|(a_1, 1)\|_{S_{\mathcal{B}}} \cap \|(a_2, 1)\|_{S_{\mathcal{B}}} \\
&= \{b \in B \mid \overline{a_1}(b) = 1\} \cap \{b \in B \mid \overline{a_2}(b) = 1\} \\
&= \{b \in B \mid a_1 \mathcal{R} b \text{ and } a_2 \mathcal{R} b\} = \{a_1, a_2\}^\downarrow
\end{aligned}$$

that is, an object satisfies Φ_3 in RST if it is related to all of the attributes of $\{a_1, a_2\}$ in FCA, obtaining a direct correspondence between $\|\Phi_3\|_{S_{\mathcal{B}}}$ and $\{a_1, a_2\}^\downarrow$. Therefore,

$$\begin{aligned}
\{a_1, a_2\}^{\downarrow\uparrow} &= (\{a_1, a_2\}^\downarrow)^\uparrow = (\|\Phi_3\|_{S_{\mathcal{B}}})^\uparrow \\
&= \{b_4\}^\uparrow = \{a \in A \mid a \mathcal{R} b_4\} = \{a_1, a_2, a_4, d\}
\end{aligned}$$

Since $\{d\} \subseteq \{a_1, a_2\}^{\downarrow\uparrow}$, we can conclude that $\{a_1, a_2\} \Rightarrow \{d\}$ is a valid attribute implication, according to Proposition 1.

Now, we will emphasize the need to require that all of the attributes of the decision rule given in Proposition 2 must take the value 1 in order to obtain an equivalence between true decision rules and valid attribute implications. To show this, we consider a different boolean decision table, that is, $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ with $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ and $d = a_5$. We consider all of

the decision rules of Table 2 by using a_4 as condition attribute and $d = a_5$ as decision attribute. These decision rules are

$$\begin{aligned} r'_1 &: (a_4, 1) \rightarrow (d, 0) \\ r'_2 &: (a_4, 0) \rightarrow (d, 1) \\ r'_3 &: (a_4, 0) \rightarrow (d, 0) \end{aligned}$$

It is easy to check that r'_1 is a true decision rule. However, the attribute implication $\{a_4\} \Rightarrow \{d\}$ is not valid since

$$\{d\} \not\subseteq \{a_4\}^{\downarrow\uparrow} = \{b_1, b_4, b_5\}^\uparrow = \{a_4\}$$

This fact is due to the decision rule r'_1 is not a 1-decision rule.

Now, we will show that it is also possible to obtain valid attribute implications without requiring the existence of true decision rules. Notice that, by Corollary 1, in order to illustrate this case, no object can be in the extension of the antecedent of the attribute implication. Specifically, we will see that $\{a_4, a_5\} \Rightarrow \{a_2\}$ is a valid attribute implication but the expression $(a_4, 1) \wedge (a_5, 1) \rightarrow (a_2, 1)$ is not a certain decision rule. In fact, it is not even a decision rule. Furthermore, none of the decision rules obtained from the boolean decision table $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ with $\mathcal{A} = \{a_1, a_3, a_4, a_5\}$ and $d = a_2$, by using as condition attributes $\{a_4, a_5\}$, is true. On the one hand, we obtain

$$\{a_4, a_5\}^{\downarrow\uparrow} = \emptyset^\uparrow = A$$

As a result, $\{a_2\} \subseteq \{a_4, a_5\}^{\downarrow\uparrow} = A$. Hence, $\{a_4, a_5\} \Rightarrow \{a_2\}$ is a valid attribute implication. On the other hand, the expression:

$$(a_4, 1) \wedge (a_5, 1) \rightarrow (d, 1)$$

is not a decision rule if we take into account the condition attributes and decision attribute mentioned above. Even more, computing all the decision rules:

$$\begin{aligned} r''_1 &: (a_4, 1) \wedge (a_5, 0) \rightarrow (d, 1) \\ r''_2 &: (a_4, 0) \wedge (a_5, 1) \rightarrow (d, 1) \\ r''_3 &: (a_4, 0) \wedge (a_5, 0) \rightarrow (d, 0) \\ r''_4 &: (a_4, 1) \wedge (a_5, 0) \rightarrow (d, 0) \\ r''_5 &: (a_4, 0) \wedge (a_5, 0) \rightarrow (d, 1) \\ r''_6 &: (a_4, 0) \wedge (a_5, 1) \rightarrow (d, 0) \end{aligned} \tag{1}$$

and their certainty, collected in Table 3, we can conclude that there are no true decision rules with $\{a_4, a_5\}$ as condition attributes and a_2 as decision attribute, whereas $\{a_4, a_5\} \Rightarrow \{a_2\}$ is a valid attribute implication. The main reason of this fact is that in the derivation operator $\downarrow$ only objects related to the given subset of attributes are taken into account.

□

As a consequence of Proposition 2 and Corollary 1, we can also conclude that, if the context contains an object with all the attributes, then we have that the set of true 1-decision rules is equal to the set of valid attribute implications.

Table 3

Certainty of each decision rule of (1).

Rule	$cer_{S_{\mathcal{D}}}$
r_1''	0.67
r_2''	0.5
r_3''	0.33
r_4''	0.5
r_5''	0.5
r_6''	0.5

4. Conclusions and future work

In this paper, we have related true decision rules in RST and valid attributes implication in FCA. For this purpose, we have taken into account a way of obtaining boolean decision tables from a given context and vice versa. We have explored that a true decision rule is equivalent to a valid attribute implication, when at least one object has the attributes in the antecedent. This hypothesis is equivalent to check whether the object satisfies the 1-decision rule. Moreover, these theoretical results are illustrated with an example.

The obtained consequences show a first and encouraging step to relate both notions, decision rules and attribute implications, with the main goal of each framework takes advantage of the definitions and results developed during the last decades in the other framework. Therefore, as a future work, we will continue studying the relationships among these notions in the classical setting, as well as the well-known notion of functional dependencies in Database Management System (DBMS). There exist important notions to be taken into consideration, such as, association rules, algorithm of decision rules, efficiency, etc. Moreover, mixed contexts with negative attributes will be incorporated in the study. In addition, we are interested in the study of this relationship in the fuzzy setting.

Acknowledgments

Partially supported by the the 2014-2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

References

- [1] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Reducing concept lattices by means of a weaker notion of congruence. *Fuzzy Sets and Systems*, 418:153–169, 2021. Algebra.
- [2] E. Bartl and J. Konecny. L-concept lattices with positive and negative attributes: Modeling uncertainty and reduction of size. *Information Sciences*, 472:163 – 179, 2019.

- [3] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis. *Fuzzy Sets and Systems*, 2019.
- [4] M. J. Benítez-Caballero, J. Medina, and E. Ramírez-Poussa. Attribute reduction in rough set theory and formal concept analysis. *Lecture Notes in Computer Science*, 10314:513–525, 2017.
- [5] P. Cordero, M. Enciso, A. Mora, and M. Ojeda-Aciego. Computing left-minimal direct basis of implications. In M. Ojeda-Aciego and J. Outrata, editors, *CLA*, volume 1062 of *CEUR Workshop Proceedings*, pages 293–298. CEUR-WS.org, 2013.
- [6] D. Dubois, J. Medina, H. Prade, and E. Ramírez-Poussa. Disjunctive attribute dependencies in formal concept analysis under the epistemic view of formal contexts. *Information Sciences*, 2021.
- [7] B. Ganter, G. Stumme, and R. Wille, editors. *Formal Concept Analysis: Foundations and Applications*. Lecture Notes in Computer Science. Springer Verlag, 2005.
- [8] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.
- [9] J.-L. Guigues and V. Duquenne. Familles minimales d’implications informatives résultant d’un tableau de données binaires. *Mathématiques et Sciences Humaines*, 95:5–18, 1986.
- [10] 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.
- [11] J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
- [12] Z. Pawlak. Information systems theoretical foundations. *Information Systems*, 6(3):205 – 218, 1981.
- [13] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
- [14] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992.
- [15] Z. Pawlak. Rough sets and decision algorithms. In W. Ziarko and Y. Yao, editors, *Rough Sets and Current Trends in Computing*, pages 30–45, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
- [16] J. Rodríguez-Jiménez, P. Cordero, M. Enciso, and A. Mora. Negative attributes and implications in formal concept analysis. *Procedia Computer Science*, 31:758 – 765, 2014.
- [17] J. Rodríguez-Jiménez, P. Cordero, M. Enciso, and S. Rudolph. Concept lattices with negative information: A characterization theorem. *Information Sciences*, 369:51 – 62, 2016.
- [18] S. Stawicki, D. Ślęzak, A. Janusz, and S. Widz. Decision bireducts and decision reducts - a comparison. *International Journal of Approximate Reasoning*, 84:75–109, 2017.
- [19] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival, editor, *Ordered Sets*, pages 445–470. Reidel, 1982.
- [20] Y. Yao. Three-way decisions with probabilistic rough sets. *Information Sciences*, 180(3):341 – 353, 2010.
- [21] Y. Yao. Rough-set concept analysis: Interpreting rs-definable concepts based on ideas from formal concept analysis. *Information Sciences*, 346-347:442–462, 2016.

Bibliografía

- [1] L. G. Albuquerque, F. de Oliveira Roque, F. Valente-Neto, R. Koroiva, D. F. Buss, D. F. Baptista, L. U. Hepp, M. L. Kuhlmann, S. Sundar, A. P. Covich, and J. O. P. Pinto. Large-scale prediction of tropical stream water quality using rough sets theory. *Ecological Informatics*, 61:101226, 2021.
- [2] L. Antoni, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute classification and reduct computation in multi-adjoint concept lattices. *IEEE Transactions on Fuzzy Systems*, 29(5):1121–1132, 2021.
- [3] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Identifying non-sublattice equivalence classes induced by an attribute reduction in fca. *Mathematics*, 9(5), 2021.
- [4] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Impact of local congruences in variable selection from datasets. *Journal of Computational and Applied Mathematics*, page 113416, 2022. <https://doi.org/10.1016/j.cam.2021.113416>.
- [5] H. Bai, Y. Ge, J. Wang, D. Li, Y. Liao, and X. Zheng. A method for extracting rules from spatial data based on rough fuzzy sets. *Knowledge-Based Systems*, 57:28–40, 2014.

- [6] M. J. Benítez-Caballero and J. Medina. *Characterizing Clarify and Reduce Method by Means of Left-Sided Formal Concept Analysis*, pages 79–85. Springer Nature Switzerland, Cham, 2024.
- [7] M. J. Benítez-Caballero, J. Medina, and E. Ramírez-Poussa. Attribute reduction in rough set theory and formal concept analysis. In L. Polkowski, Y. Yao, P. Artiemjew, D. Ciucci, D. Liu, D. Ślęzak, and B. Zielosko, editors, *Rough Sets*, pages 513–525, Cham, 2017. Springer International Publishing.
- [8] M. J. Benítez-Caballero, J. Medina, and E. Ramírez-Poussa. Towards a classification of rough set bireducts. In M.-J. Lesot, S. Vieira, M. Z. Reformat, J. a. P. Carvalho, A. Wilbik, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 759–770, Cham, 2020. Springer International Publishing.
- [9] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. Bireducts with tolerance relations. *Information Sciences*, 435:26 – 39, 2018.
- [10] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis. *Fuzzy Sets and Systems*, 2019.
- [11] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. A computational procedure for variable selection preserving different initial conditions. *International Journal of Computer Mathematics*, 97(1-2):387–404, 2020.
- [12] M. J. Benítez-Caballero, J. Medina-Moreno, and E. Ramírez-Poussa. *Bireducts in Formal Concept Analysis*, pages 191–198. Springer International Publishing, Cham, 2020.

- [13] B. M. Brentan, S. Carpitella, J. Izquierdo, E. Luvizotto Jr, and G. Meirelles. District metered area design through multicriteria and multi-objective optimization. *Mathematical Methods in the Applied Sciences*, 45(6):3254–3271, 2022.
- [14] F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Relating decision rules and attribute implications. *In the 16th International Conference on Concept Lattices and Their Applications (CLA 2022), CEUR Workshop Proceedings*, 2022.
- [15] F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Decision making in Fuzzy Rough Set Theory. *Mathematics*, 11(19), 2023.
- [16] F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Fuzzy Rough Set Decision Algorithms. In D. Ciucci, I. Couso, J. Medina, D. Slezak, D. Petturiti, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 19th International Conference, IPMU 2022, Milan, Italy, July 11-15, 2022, Proceedings, Part I*, volume 1601 of *Communications in Computer and Information Science*, pages 63–76. Springer, 2022.
- [17] F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. *Information Bireducts and Its Relationship with Reducts*, pages 37–44. Springer International Publishing, Cham, 2023.
- [18] F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Value reducts and bireducts: A comparative study. *Mathematical Methods in the Applied Sciences*, 46(2):1631–1650, 2023.
- [19] F. Chacón-Gómez, M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. *Efficiency of Fuzzy Rough Set Decision Algorithms*, pages 17–24. Springer Nature Switzerland, Cham, 2024.

- [20] 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.
- [21] J. Chen, J. Mi, B. Xie, and Y. Lin. A fast attribute reduction method for large formal decision contexts. *International Journal of Approximate Reasoning*, 106:1 – 17, 2019.
- [22] M. E. Cornejo, J. C. Díaz-Moreno, and J. Medina. Multi-adjoint relation equations: A decision support system for fuzzy logic. *International Journal of Intelligent Systems*, 32(8):778–800, 2017.
- [23] M. E. Cornejo, D. Lobo, and J. Medina. Bipolar fuzzy relation equations systems based on the product t-norm. *Mathematical Methods in the Applied Sciences*, 42(17):5779–5793, 2019.
- [24] M. E. Cornejo, D. Lobo, and J. Medina. Extended multi-adjoint logic programming. *Fuzzy Sets and Systems*, 388:124–145, 2020. Logic.
- [25] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. A comparative study of adjoint triples. *Fuzzy Sets and Systems*, 211:1–14, 2013.
- [26] C. Cornelis, M. De Cock, and A. M. Radzikowska. Vaguely quantified rough sets. In A. An, J. Stefanowski, S. Ramanna, C. J. Butz, W. Pedrycz, and G. Wang, editors, *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, pages 87–94, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [27] C. Cornelis, M. De Cock, and A. M. Radzikowska. *Fuzzy Rough Sets: From Theory into Practice*, pages 533–552. John Wiley & Sons, Ltd, 2008.
- [28] C. Cornelis, R. Jensen, G. Hurtado, and D. Ślęzak. Attribute selection with fuzzy decision reducts. *Information Sciences*, 180:209–224, 2010.

- [29] C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.
- [30] W. F. De Bondt and R. H. Thaler. Chapter 13 financial decision-making in markets and firms: A behavioral perspective. In *Finance*, volume 9 of *Handbooks in Operations Research and Management Science*, pages 385–410. Elsevier, 1995.
- [31] 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.
- [32] L. F. Del Cerro and E. Orłowska. Dal-a logic for data analysis. *Theoretical Computer Science*, 36:251–264, 1985.
- [33] F. Di Martino and S. Sessa. Comparison between images via bilinear fuzzy relation equations. *Journal of Ambient Intelligence and Humanized Computing*, 9(5):1517–1525, Oct 2018.
- [34] J. C. Díaz-Moreno and J. Medina. Multi-adjoint relation equations: Definition, properties and solutions using concept lattices. *Information Sciences*, 253:100–109, 2013.
- [35] J. C. Díaz-Moreno and J. Medina. Using concept lattice theory to obtain the set of solutions of multi-adjoint relation equations. *Information Sciences*, 266(0):218–225, 2014.
- [36] J. C. Díaz-Moreno, J. Medina, and M. Ojeda-Aciego. On basic conditions to generate multi-adjoint concept lattices via Galois connections. *International Journal of General Systems*, 43(2):149–161, 2014.
- [37] L. Dong, R. Wang, and D. Chen. Incremental feature selection with fuzzy rough sets for dynamic data sets. *Fuzzy Sets and Systems*, 467:108503, 2023.

- [38] D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17(2-3):191–209, 1990.
- [39] D. Dubois and H. Prade. Possibility theory, probability theory and multiple-valued logics: A clarification. *Ann. Math. Artif. Intell.*, 32:35–66, 08 2001.
- [40] D. Dubois and H. Prade. Formal concept analysis from the standpoint of possibility theory. In J. Baixeries, C. Sacarea, and M. Ojeda-Aciego, editors, *Formal Concept Analysis - 13th International Conference, ICFCA 2015, Nerja, Spain, June 23-26, 2015, Proceedings*, volume 9113 of *Lecture Notes in Computer Science*, pages 21–38. Springer, 2015.
- [41] N. Duong-Bao, J. He, L. N. Thi, K. Nguyen-Huu, and S.-W. Lee. A novel valued tolerance rough set and decision rules method for indoor positioning using wifi fingerprinting. *Sensors*, 22(15), 2022.
- [42] M. Dursun and G. Gumus. Intuitionistic fuzzy cognitive map approach for the evaluation of supply chain configuration criteria. *Mathematical Methods in the Applied Sciences*, 43(13):7788–7801, 2020.
- [43] O. G. El Barbary, A. S. Salama, and E. S. Atlam. Granular information retrieval using neighborhood systems. *Mathematical Methods in the Applied Sciences*, 41(15):5737–5753, 2018.
- [44] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.
- [45] A. Hamed, A. Sobhy, and H. Nassar. Accurate classification of covid-19 based on incomplete heterogeneous data using a knn variant algorithm. *Arabian Journal for Science and Engineering*, 46:1–12, 03 2021.
- [46] S. Harding. Supported decision making: Skills and tools for clinical practice. *The Journal for Nurse Practitioners*, 2022.

- [47] S. C. Hendrikse, S. Kluiver, J. Treur, T. F. Wilderjans, S. Dikker, and S. L. Koole. How virtual agents can learn to synchronize: an adaptive joint decision-making model of psychotherapy. *Cognitive Systems Research*, 2022.
- [48] C. M. J. Li and Y. Lv. Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction. *International Journal of Approximate Reasoning*, 54(1):149–165, 2013.
- [49] A. Janusz and D. Ślęzak. Rough set methods for attribute clustering and selection. *Applied Artificial Intelligence*, 28(3):220–242, 2014.
- [50] A. Janusz, D. Ślęzak, and H. S. Nguyen. Unsupervised similarity learning from textual data. *Fundamenta Informaticae*, Volume 119:319 – 336, 01 2012.
- [51] S. Jasanoff. The political science of risk perception. *Reliability Engineering & System Safety*, 59(1):91–99, 1998. Risk Perception Versus Risk Analysis.
- [52] R. Jensen and Q. Shen. New approaches to fuzzy-rough feature selection. *IEEE Transactions on Fuzzy Systems*, 17(4):824–838, Aug 2009.
- [53] P. Julián-Iranzo, J. Medina, and M. Ojeda-Aciego. On reductants in the framework of multi-adjoint logic programming. *Fuzzy Sets and Systems*, 317:27 – 43, 2017.
- [54] R. Kowalski. Predicate logic as programming language. volume 74, pages 569–574, 01 1974.
- [55] K. Kuczera. Application of rough set theory to identify priorities in project management. *Procedia Computer Science*, 225:4410–4415, 2023. 27th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (KES 2023).

- [56] G. Lang, D. Miao, and M. Cai. Three-way decision approaches to conflict analysis using decision-theoretic rough set theory. *Information Sciences*, 406-407:185–207, 2017.
- [57] Y.-J. Leng and H. Zhang. A novel evaluation method of renewable energy development based on improved rough set theory and grey cloud model. *Journal of Cleaner Production*, 428:139299, 2023.
- [58] 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.
- [59] S. Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.
- [60] N. Mac Parthalain and R. Jensen. Simultaneous feature and instance selection using fuzzy-rough bireducts. In *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2013)*, pages 1–8, July 2013.
- [61] J. Medina. Towards multi-adjoint property-oriented concept lattices. *Lecture Notes in Artificial Intelligence*, 6401:159–166, 2010.
- [62] J. Medina. Multi-adjoint property-oriented and object-oriented concept lattices. *Information Sciences*, 190:95–106, 2012.
- [63] J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
- [64] J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.
- [65] 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.

- [66] Y. Mi, Z. Wang, H. Liu, Y. Qu, G. Yu, and Y. Shi. Divide and conquer: A granular concept-cognitive computing system for dynamic classification decision making. *European Journal of Operational Research*, 2022.
- [67] A. Nakamura. Fuzzy rough sets. *Note on Multiple-Valued Logic in Japan*, 9:1–8, 1988.
- [68] H. S. Nguyen. Scalable classification method based on rough sets. In J. J. Alpigini, J. F. Peters, A. Skowron, and N. Zhong, editors, *Rough Sets and Current Trends in Computing*, pages 433–440, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [69] F. Nosheen, U. Qamar, and M. S. Raza. A parallel rule-based approach to compute rough approximations of dominance based rough set theory. *Engineering Applications of Artificial Intelligence*, 115:105285, 2022.
- [70] Z. Pawlak. Information systems theoretical foundations. *Information Systems*, 6(3):205 – 218, 1981.
- [71] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
- [72] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992.
- [73] Z. Pawlak. Rough sets and decision algorithms. In W. Ziarko and Y. Yao, editors, *Rough Sets and Current Trends in Computing*, pages 30–45, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
- [74] F. Pérez-Gámez, D. López-Rodríguez, P. Cordero, A. Mora, and M. Ojeda-Aciego. Simplifying implications with positive and negative attributes: A logic-based approach. *Mathematics*, 10(4), 2022.

- [75] M. Qi, Y. Gu, and Q. Wang. Internet financial risk management and control based on improved rough set algorithm. *Journal of Computational and Applied Mathematics*, 384:113179, 2021.
- [76] J. Qu, X. Bai, J. Gu, F. Taghizadeh-Hesary, and J. Lin. Assessment of rough set theory in relation to risks regarding hydraulic engineering investment decisions. *Mathematics*, 8(8), 2020.
- [77] A. M. Radzikowska and E. E. Kerre. A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems*, 126(2):137–155, 2002.
- [78] X. Ren, D. Li, and Y. Zhai. Research on mixed decision implications based on formal concept analysis. *International Journal of Cognitive Computing in Engineering*, 4:71–77, 2023.
- [79] J. Rodríguez-Jiménez, P. Cordero, M. Enciso, and A. Mora. Negative attributes and implications in formal concept analysis. *Procedia Computer Science*, 31:758 – 765, 2014.
- [80] J. Rodríguez-Jiménez, P. Cordero, M. Enciso, and S. Rudolph. Concept lattices with negative information: A characterization theorem. *Information Sciences*, 369:51 – 62, 2016.
- [81] J. M. Rodríguez-Jiménez, P. Cordero, M. Enciso, and A. Mora. Data mining algorithms to compute mixed concepts with negative attributes: an application to breast cancer data analysis. *Mathematical Methods in the Applied Sciences*, 39(16):4829–4845, 2016.
- [82] E. Sanchez. Resolution of composite fuzzy relation equations. *Information and Control*, 30(1):38–48, 1976.
- [83] S. Sarkar, A. Pramanik, and J. Maiti. An integrated approach using rough set theory, anfis, and z-number in occupational risk prediction. *Engineering Applications of Artificial Intelligence*, 117:105515, 2023.

- [84] S. U. Selvan, S. T. Saroglou, J. Joschinski, M. Calbi, V. Vogler, S. Barath, and Y. J. Grobman. Toward multi-species building envelopes: A critical literature review of multi-criteria decision-making for design support. *Building and Environment*, page 110006, 2023.
- [85] S. Sessa. Some results in the setting of fuzzy relation equations theory. *Fuzzy Sets and Systems*, 14(3):281–297, 1984.
- [86] Z. Shi, L. Li, C. Bo, and J. Zhan. Two novel three-way decision models based on fuzzy β -covering rough sets and prospect theory under q-rung orthopair fuzzy environments. *Expert Systems with Applications*, page 123050, 2023.
- [87] R. Słowiński and J. Stefanowski. Rough classification with valued closeness relation. In E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, and B. Burtschy, editors, *New Approaches in Classification and Data Analysis*, pages 482–489, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.
- [88] S. Stawicki and D. Ślęzak. Recent advances in decision bireducts: Complexity, heuristics and streams. In P. Lingras, M. Wolski, C. Cornelis, S. Mitra, and P. Wasilewski, editors, *Rough Sets and Knowledge Technology*, volume 8171 of *Lecture Notes in Computer Science*, pages 200–212. Springer Berlin Heidelberg, 2013.
- [89] S. Stawicki, D. Ślęzak, A. Janusz, and S. Widz. Decision bireducts and decision reducts - a comparison. *International Journal of Approximate Reasoning*, 84:75–109, 2017.
- [90] J. Stefanowski and A. Tsoukiàs. Valued tolerance and decision rules. In W. Ziarko and Y. Yao, editors, *Rough Sets and Current Trends in Computing*, pages 212–219, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

- [91] J. Stefanowski and A. Tsoukiás. Induction of decision rules and classification in the valued tolerance approach. In J. J. Alpigini, J. F. Peters, A. Skowron, and N. Zhong, editors, *Rough Sets and Current Trends in Computing*, pages 271–278, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [92] O. Svenson. Decision making and the search for fundamental psychological regularities: What can be learned from a process perspective? *Organizational Behavior and Human Decision Processes*, 65(3):252–267, 1996.
- [93] A. Theerens and C. Cornelis. Fuzzy rough sets based on fuzzy quantification. *Fuzzy Sets and Systems*, 473:108704, 2023.
- [94] G. Wang, Y. Guo, W. Zhang, S. Xie, and Q. Chen. What type of algorithm is perceived as fairer and more acceptable? A comparative analysis of rule-driven versus data-driven algorithmic decision-making in public affairs. *Government Information Quarterly*, page 101803, 2023.
- [95] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival, editor, *Ordered Sets*, pages 445–470. Reidel, 1982.
- [96] J. J. Wu, T. A. Mazzuchi, and S. Sarkani. Comparison of multi-criteria decision-making methods for online controlled experiments in a launch decision-making framework. *Information and Software Technology*, 155:107115, 2023.
- [97] Q.-E. Wu, T. Wang, X.-M. Pang, Y.-X. Huang, and J.-S. Li. Research on two different mathematical theories on control. *Journal of Computational and Applied Mathematics*, 222(2):675–689, 2008.

- [98] Z. Xiao, K. Gong, and Y. Zou. A combined forecasting approach based on fuzzy soft sets. *Journal of Computational and Applied Mathematics*, 228(1):326–333, 2009.
- [99] Y. Yao. Three-way decisions with probabilistic rough sets. *Information Sciences*, 180(3):341 – 353, 2010.
- [100] Y. Yao and Y. Zhao. Attribute reduction in decision-theoretic rough set models. *Information Sciences*, 178(17):3356–3373, 2008.
- [101] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [102] 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.
- [103] Y. Zhao, Y. Yao, and F. Luo. Data analysis based on discernibility and indiscernibility. *Information Sciences*, 177(22):4959 – 4976, 2007.
- [104] H.-J. Zimmermann. *Fuzzy Set Theory - and Its Applications*, volume 2001. 01 2001.

