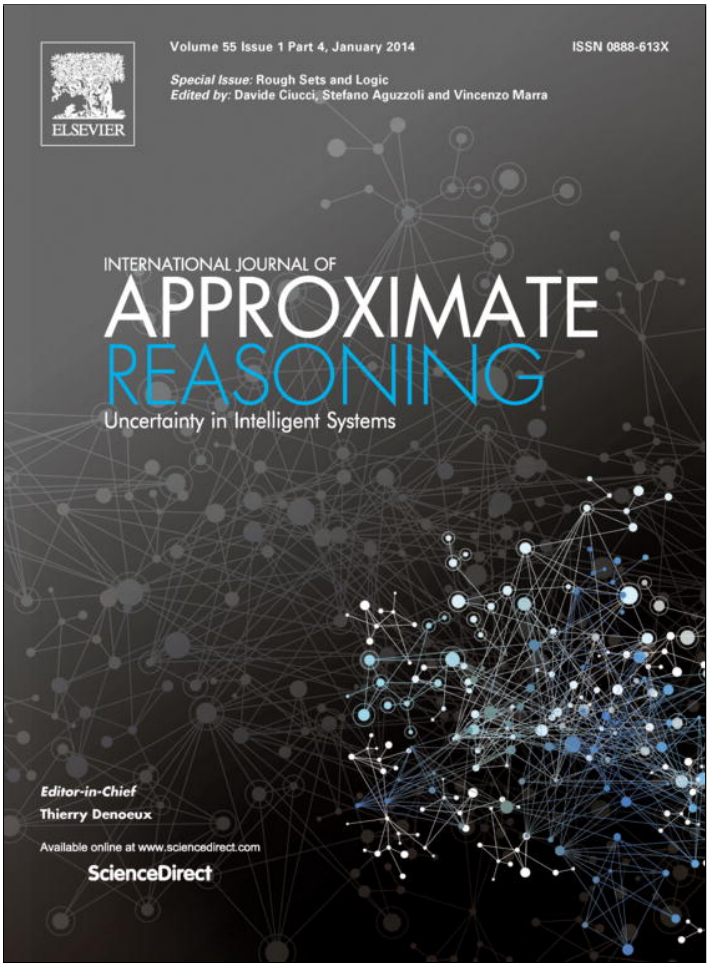
Volume 55 Issue 1 Part 4, January 2014
ISSN 0888-613X
ELSEVIER
Special Issue: Rough Sets and Logic
Edited by: Davide Ciucci, Stefano Aguzzoli and Vincenzo Marra
INTERNATIONAL JOURNAL OF
APPROXIMATE
REASONING
Uncertainty in Intelligent Systems
Editor-in-Chief
Thierry Denoeux
Available online at www.sciencedirect.com
ScienceDirect

International Journal of Approximate Reasoning 55 (2014) 412–426

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection

Chris Cornelis [a,c,1], Jesús Medina [b,2,*], Nele Verbiest [c]

[a] Department of Computer Science and Artificial Intelligence, University of Granada, Spain
[b] Department of Mathematics, University of Cádiz, Spain
[c] Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Gent, Belgium

ARTICLE INFO

Article history:
Received 18 September 2012
Received in revised form 13 July 2013
Accepted 1 September 2013
Available online 20 September 2013

Keywords:
Rough sets
Concept lattices
Fuzzy sets
Attribute selection
Data analysis
Decision reducts

ABSTRACT

This paper introduces a flexible extension of rough set theory: multi-adjoint fuzzy rough sets, in which a family of adjoint pairs are considered to compute the lower and upper approximations. This new setting increases the number of applications in which rough set theory can be used. An important feature of the presented framework is that the user may represent explicit preferences among the objects in a decision system, by associating a particular adjoint triple with any pair of objects.
Moreover, we verify mathematical properties of the model, study its relationships to multi-adjoint property-oriented concept lattices and discuss attribute selection in this framework.

1. Introduction

Pawlak proposed rough set theory [23] in the 1980s as a formal tool for modeling and processing incomplete information in information systems.

On the other hand, formal concept analysis, introduced by Wille in the decade of 1980 [28], arose as another mathematical theory for qualitative data analysis and, currently, has become an interesting research topic both with regard to its mathematical foundations [16,25] and with regard to its multiple applications [5,6].

These mathematical theories have been related in several papers [7,8,14,15,18,19]. In particular, property-oriented concept lattices [1,4,10] and object-oriented concept lattices [29] were introduced in order to extend formal concept lattices [9], with constructs from rough set theory; notably, they invoke the lower and upper approximation operators, which are often referred to in this research field as necessity and possibility operators, respectively.

More recently, multi-adjoint property-oriented and object-oriented concept lattices were studied [17,18], with the aim of introducing adjoint triples of fuzzy logic operators (in particular, a conjunctor and its two residuated implications) to define "soft" extensions of the necessity and possibility operators. This is similar to what happens in fuzzy rough set theory, where a t-norm and fuzzy implication are used in order to extend the classical rough lower and upper approximation operators.

* Corresponding author.
E-mail addresses: chriscornelis@ugr.es (C. Cornelis), jesus.medina@uca.es (J. Medina), Nele.Verbiest@UGent.be (N. Verbiest).
[1] Partially supported by the Spanish Ministry of Science and Technology under grant TIN2011-28488.
[2] Partially supported by Junta de Andalucía grant P09-FQM-5233, and by the EU (FEDER), and the Spanish Science and Education Ministry (MEC) under grants TIN2009-14562-C05-03 and TIN2012-39353-C04-04.

0888-613X/$ – see front matter
http://dx.doi.org/10.1016/j.ijar.2013.09.007

However, the multi-adjoint paradigm goes much further in the sense that it allows us to use several adjoint triples, in order to be able to express preferences among objects or properties.

In this paper, the latter characteristic of multi-adjoint property-oriented concept lattices is introduced into the framework of fuzzy rough sets, that is to say, we propose a multi-adjoint fuzzy rough set model in which the lower and upper approximation operators are constructed using several adjoint triples. This allows us to introduce explicit preferences among the objects, by associating a particular adjoint triple with any pair of objects in a decision system.

We study various properties of the model and focus in particular on attribute selection. From the perspectives of both concept lattices and rough sets, attribute selection is an important step in reducing the computational complexity. Recently, Wang and Zhang related attribute selection in property-oriented and object-oriented concept lattices [27]. Moreover, in [22], two kinds of reduction methods have been proposed and the relationship with attribute selection in rough set theory is discussed in detail.

The remainder of this paper is structured as follows. In Section 2, we recall preliminaries from rough sets, fuzzy rough sets and multi-adjoint property-oriented concept lattices. Next, in Section 3, we define multi-adjoint fuzzy rough sets and investigate their main properties. We also define a general positive region to focus on the decision attribute and on the attribute selection based on multivalued measures. In Section 4, the notions of L-valued measure, m, and fuzzy m-decision reduct are included, and specific measures are studied based on the positive region notion and on a fuzzy discernibility function which generalize the ones given in [3]. Moreover, we introduce a relation between them. Finally, in Section 5, we conclude the paper.

2. Preliminaries

2.1. Rough set theory

In the framework of rough set theory, data is represented as an *information system* $(X, \mathcal{A})$, where $X = \{x_1, \ldots, x_n\}$ and $\mathcal{A} = \{a_1, \ldots, a_m\}$ are finite, non-empty sets of objects and attributes, respectively. Each a in $\mathcal{A}$ corresponds to a mapping $\bar{a}: X \to V_a$, where V_a is the value set of a over X. For every subset B of $\mathcal{A}$, the B-indiscernibility relation[3] R_B is defined as the equivalence relation

$$R_B = \left\{(x, y) \in X \times X \;\middle|\; \text{for all } a \in B,\ \bar{a}(x) = \bar{a}(y)\right\} \tag{1}$$

Given $A \subseteq X$, its lower and upper approximation w.r.t. B are defined by

$$R_B{\downarrow}A = \left\{x \in X \;\middle|\; [x]_{R_B} \subseteq A\right\} \tag{2}$$

$$R_B{\uparrow}A = \left\{x \in X \;\middle|\; [x]_{R_B} \cap A \neq \emptyset\right\} \tag{3}$$

A *decision system* $(X, \mathcal{A} \cup \{d\})$ is a special kind of information system, in which $d \notin \mathcal{A}$ is called the decision attribute, and its equivalence classes $[x]_{R_d}$ are called decision classes. Given $B \subseteq \mathcal{A}$, the B-positive region, POS_B, and the degree of dependency of d on B, γ_B, are defined as

$$POS_B = \bigcup_{x \in X} R_B{\downarrow}[x]_{R_d} \tag{4}$$

$$\gamma_B = \frac{|POS_B|}{|X|} \tag{5}$$

$(X, \mathcal{A} \cup \{d\})$ is called *consistent* if $\gamma_{\mathcal{A}} = 1$. A subset B of $\mathcal{A}$ is called a *decision reduct* if it satisfies $POS_B = POS_{\mathcal{A}}$ and there exists no proper subset B' of B such that $POS_{B'} = POS_{\mathcal{A}}$.

A well-known approach to generate all reducts of a decision system is based on its discernibility matrix and function [26]. The discernibility matrix of $(X, \mathcal{A} \cup \{d\})$ is the $n \times n$ matrix O, defined by, for i and j in $\{1, \ldots, n\}$,

$$O_{ij} = \begin{cases} \emptyset & \text{if } d(x_i) = d(x_j) \\ \{a \in \mathcal{A} \mid \bar{a}(x_i) \neq \bar{a}(x_j)\} & \text{otherwise} \end{cases} \tag{6}$$

The *discernibility function* of $(X, \mathcal{A} \cup \{d\})$ is the map $f: \{0,1\}^m \to \{0,1\}$, defined by

$$f\left(a_1^*, \ldots, a_m^*\right) = \bigwedge\left\{\bigvee O_{ij}^* \;\middle|\; 1 \leqslant i < j \leqslant n \text{ and } O_{ij} \neq \emptyset\right\} \tag{7}$$

in which $O_{ij}^* = \{a^* \mid a \in O_{ij}\}$. The boolean variables $a_1^*, \ldots, a_m^*$ correspond to the attributes from $\mathcal{A}$. It can be shown that the prime implicants of f constitute exactly all decision reducts of $(X, \mathcal{A} \cup \{d\})$.

[3] When $B = \{a\}$, i.e., B is a singleton, we will write R_a instead of $R_{\{a\}}$.

2.2. Fuzzy rough set theory

In fuzzy rough set theory, a decision system $(X, \mathcal{A} \cup \{d\})$ is defined in the same way as in the crisp case. However, several other notions need to be fuzzified.

To express the approximate equality between two objects w.r.t. $a \in \mathcal{A}$, a $[0,1]$-fuzzy tolerance relation is considered [3, 13], that is, a fuzzy relation $R_a : X \times X \to [0,1]$ which is reflexive and symmetric. For any subset $B = \{a_{m_1}, \ldots, a_{m_{|B|}}\}$ of $\mathcal{A}$, the fuzzy B-indiscernibility relation is defined by, for x, y in X,

$$R_B(x, y) = \mathcal{T}\big(R_{a_{m_1}}(x, y), \ldots, R_{a_{m_{|B|}}}(x, y)\big) \tag{8}$$

in which $\mathcal{T}$ represents a t-norm.

The lower and upper approximation of a fuzzy set A in X by means of R_B is defined as in [24]: given a fuzzy implication[4] $\mathcal{I}$ and a t-norm $\mathcal{T}$, for all y in X,

$$(R_B{\downarrow}A)(y) = \inf_{x \in X} \mathcal{I}\big(R_B(x, y), A(x)\big) \tag{9}$$

$$(R_B{\uparrow}A)(y) = \sup_{x \in X} \mathcal{T}\big(R_B(x, y), A(x)\big) \tag{10}$$

Next, the fuzzy B-positive region is defined, for each y in X, as

$$POS_B(y) = \left(\bigcup_{x \in X} R_B{\downarrow}R_d x\right)(y) \tag{11}$$

where $R_d x : X \to [0,1]$ is defined as $R_d x(y) = R_d(x, y)$.

2.3. Multi-adjoint property-oriented and object-oriented concept lattices

This section only recalls the definitions and main properties of multi-adjoint property-oriented concept lattice framework, since the object-oriented one is introduced similarly. The details of both concept lattices are given in [17,18].

The basic operators in these environments are the adjoint triples [2], which are formed by three mappings: a possible non-commutative conjunctor and two residuated implications [11], that satisfy the well-known adjoint property.

Definition 1. Let $(P_1, \leqslant_1)$, $(P_2, \leqslant_2)$, $(P_3, \leqslant_3)$ be posets and $\& : P_1 \times P_2 \to P_3$, $\swarrow : P_3 \times P_2 \to P_1$, $\nwarrow : P_3 \times P_1 \to P_2$ be mappings, then $(\&, \swarrow, \nwarrow)$ is an *adjoint triple* with respect to P_1, P_2, P_3 if:

$$x \leqslant_1 z \swarrow y \quad \text{iff} \quad x \,\&\, y \leqslant_3 z \quad \text{iff} \quad y \leqslant_2 z \nwarrow x$$

where $x \in P_1$, $y \in P_2$ and $z \in P_3$.

The usual Gödel, product and Łukasiewicz t-norms together with their residuated implications are examples of adjoint triples.

Example 1. The Gödel, product and Łukasiewicz adjoint triples are defined on $[0,1]$ as:

$$\begin{aligned} &x \,\&_G\, y = \min\{x, y\} && z \nwarrow_G x = \begin{cases} 1 & \text{if } x \leqslant z \\ z & \text{otherwise} \end{cases} \\ &x \,\&_P\, y = x \cdot y && z \nwarrow_P x = \min\{1, z/x\} \\ &x \,\&_L\, y = \max\{0, x + y - 1\} && z \nwarrow_L x = \min\{1, 1 - x + z\} \end{aligned}$$

and $\swarrow^G = \nwarrow_G$, $\swarrow^P = \nwarrow_P$, $\swarrow^L = \nwarrow_L$, since $\&_P$, $\&_G$ and $\&_L$ are commutative.

In [21] more general examples of adjoint triples are given, in which $\&$ is not a t-norm and $\swarrow$ is different from $\nwarrow$.

The following easily verified lemma enumerates some properties of the adjoint triples which will be used in the next sections.

Lemma 1. *Let $(P_1, \leqslant_1)$ be a poset, $(L_2, \preceq_2)$ and $(L_3, \preceq_3)$ two complete lattices, and $(\&, \swarrow, \nwarrow)$ an adjoint triple with respect to P_1, L_2, L_3, then $\&$ preserves the supremum in the second argument and $\nwarrow$ preserves the infimum in the first argument, i.e., when I and J are two index sets and $x \in P_1$, $\{y_i \mid i \in I\} \subseteq L_2$ and $\{z_j \mid j \in J\} \subseteq L_3$, then*

[4] Note that a fuzzy implication $\mathcal{I} : [0,1] \times [0,1] \to [0,1]$ is an operator decreasing in the antecedent (in the left side) and increasing in the consequent (in the right side) and satisfying the same boundary conditions as the classical implication.

(1) $x \,\&\, (\sup\{y_i \mid i \in I\}) = \sup\{x \,\&\, y_i \mid i \in I\}$.
(2) $(\inf\{z_j \mid j \in J\}) \nwarrow x = \inf\{z_j \nwarrow x \mid j \in J\}$.

The basic structure, which allows the existence of several adjoint triples for a given triplet of a poset and two complete lattices, is the multi-adjoint property-oriented frame.

Definition 2. Given a poset $(P_1, \leqslant_1)$, two complete lattices $(L_2, \preccurlyeq_2)$ and $(L_3, \preccurlyeq_3)$ and adjoint triples with respect to P_1, L_2, L_3, $(\&_i, \swarrow^i, \nwarrow_i)$, for all $i = 1, \dots, n$, a *multi-adjoint property-oriented frame* is the tuple

$$(P_1, L_2, L_3, \&_1, \dots, \&_n)$$

The definition of context is analogous to the one given in [20].

Definition 3. Given a multi-adjoint property-oriented frame $(P_1, L_2, L_3, \&_1, \dots, \&_n)$, a *context* is a tuple (A, B, R, τ) such that A and B are non-empty sets (usually interpreted as attributes and objects, respectively), R is a P_1-fuzzy relation $R: A \times B \to P_1$ and $\tau: A \times B \to \{1, \dots, n\}$ is a mapping which associates any pair of elements in $A \times B$ with some particular adjoint triple in the frame.

Note that R is a general P_1-fuzzy relation, which in general does not need to be reflexive or symmetric when $A = B$ is assumed. From now on, we will fix a multi-adjoint property-oriented frame and context, denoted by $(P_1, L_2, L_3, \&_1, \dots, \&_n)$ and (A, B, R, τ).

Now, given[5] $g \in L_2^B$, and $f \in L_3^A$, we define the following mappings: ${}^{\uparrow_\pi}: L_2^B \to L_3^A$, ${}^{\downarrow^N}: L_3^A \to L_2^B$:

$$g^{\uparrow_\pi}(a) = \sup\{R(a,b) \,\&_{\tau(a,b)}\, g(b) \mid b \in B\} \tag{12}$$

$$f^{\downarrow^N}(b) = \inf\{f(a) \nwarrow_{\tau(a,b)} R(a,b) \mid a \in A\} \tag{13}$$

Clearly, these definitions generalize the classical possibility and necessity operators [10] and the mappings presented in Eqs. (9) and (10).

A concept, in this environment, is a pair of mappings $\langle g, f \rangle$, with $g \in L_2^B$, $f \in L_3^A$, such that $g^{\uparrow_\pi} = f$ and $f^{\downarrow^N} = g$, which will be called *multi-adjoint property-oriented formal concept*. The set of all these concepts will be denoted as $\mathcal{M}_{\pi N}$ and, together with the ordering $\preccurlyeq$, defined by $\langle g_1, f_1 \rangle \preccurlyeq \langle g_2, f_2 \rangle$ if and only if $g_1 \preccurlyeq_2 g_2$ (or equivalently $f_1 \preccurlyeq_3 f_2$), is a complete lattice [18], which is called *multi-adjoint property-oriented concept lattice*.

3. Multi-adjoint fuzzy rough sets

As seen in the previous section, multi-adjoint property-oriented concept lattices generalize property-oriented concept lattices using several adjoint triples. In this framework, a context (A, B, R, τ) is considered, in which the roles of B and the relation $R: A \times B \to P$ will be related to fuzzy rough set theory.

In fuzzy rough set theory we start from a decision system $(X, \mathcal{A} \cup \{d\})$ and, given $B \subseteq \mathcal{A}$, we define a B-indiscernibility relation $R_B: X \times X \to P$. In other words, rather than relating objects with properties as in property-oriented concept lattices, in fuzzy rough set theory, R establishes a relation between the objects. As a consequence of this important difference, the attribute selection in rough set theory is not the same thing as in the concept lattice setting.

This section introduces a generalization of fuzzy rough sets considering the multi-adjoint philosophy. As a consequence, a more flexible environment is obtained: we can consider several adjoint triples, which allows us, for example, to assume preference among the objects as it was shown in [18,20].

From now on, in the rest of the paper, a decision system $(X, \mathcal{A} \cup \{d\})$ a poset $(P, \leqslant)$ and a complete lattice $(L, \preccurlyeq)$ will be fixed. We assume that the poset P has a top element $\top_P$ and the bottom and top element of the lattice L are denoted by $\bot_L$ and $\top_L$ respectively.

3.1. Definitions

We first introduce a P-fuzzy generalization of the B-indiscernibility relation[6] associated with the decision system $(X, \mathcal{A} \cup \{d\})$. For that, we need the notion of a P-fuzzy tolerance relation $R: X \times X \to P$, i.e., a P-fuzzy relation that satisfies the reflexivity property, $R(x,x) = \top_P$, for all $x \in X$, and the symmetric property $R(x,y) = R(y,x)$, for all $x, y \in X$.

[5] Given two sets L and X, we use the shorthand L^X to denote the set of mappings from X to L, that is, $L^X = \{h: X \to L\}$.
[6] Note that B is considered to be a subset of attributes in this environment, following the usual notation in rough set theory.

Definition 4. Given $B \subseteq \mathcal{A}$ and the P-fuzzy tolerance relations $R_a : X \times X \to P$, for all $a \in \mathcal{A}$, a fuzzy B-indiscernibility relation is a P-fuzzy relation $R_B : X \times X \to P$, defined, for all $x, y \in X$, as

$$R_B(x, y) = @\big(\phi_B^{x,y}(a_1), \ldots, \phi_B^{x,y}(a_m)\big) \tag{14}$$

in which $@ : P^m \to P$ is an *aggregation operator*, that is, a monotonic operator on each argument satisfying that $@(\top_P, \ldots, \top_P) = \top_P$ and, if P has a bottom element $\bot_P$, then $@(\bot_P, \ldots, \bot_P) = \bot_P$, and $\phi_B^{x,y} : \mathcal{A} \to P$ is defined for each $a \in \mathcal{A}$ as

$$\phi_B^{x,y}(a) = \begin{cases} R_a(x, y) & \text{if } a \in B \\ \top_P & \text{otherwise} \end{cases}$$

The monotonicity property is clear for this relation just like in the classical case, i.e. given $B_1 \subseteq B_2$, then $R_{B_2} \leqslant R_{B_1}$.

The algebraic structure assumed will be a multi-adjoint property-oriented frame $(P, L, L, \&_1, \ldots, \&_n)$, and, from the decision system and each B-indiscernibility relation, a context (X, X, R_B, τ) can be considered, on which the possibility and necessity operators can be defined. In particular, given $g \in L^X$, and $f \in L^X$, the mappings: ${}^{\uparrow_\pi} : L^X \to L^X$, ${}^{\downarrow^N} : L^X \to L^X$ are defined, for all $x, y \in L$ as:

$$g^{\uparrow_\pi}(x) = \sup\big\{R_B(x, y) \&_{\tau(x,y)} g(y) \bigm| y \in X\big\} \tag{15}$$

$$f^{\downarrow^N}(y) = \inf\big\{f(x) \nwarrow_{\tau(x,y)} R_B(x, y) \bigm| x \in X\big\} \tag{16}$$

where $g^{\uparrow_\pi}$ can be interpreted as the *upper approximation* of g and $f^{\downarrow^N}$ the *lower approximation* of f. From now on, in order to improve readability, we will write $\&_{xy}$, $\nwarrow_{xy}$ instead of $\&_{\tau(x,y)}$, $\nwarrow_{\tau(x,y)}$.

Definition 5. Given a fuzzy subset $h \in L^X$, the pair $(h^{\downarrow^N}, h^{\uparrow_\pi})$ is called a *multi-adjoint fuzzy rough set*.

Note that this pair is not necessarily a multi-adjoint property-oriented concept, since $h^{\downarrow^N}$ may be different from $h^{\uparrow_\pi \downarrow^N}$ or equivalently, $h^{\uparrow_\pi}$ could be different from $h^{\downarrow^N \uparrow_\pi}$, as the following example illustrates.

Example 2. Consider the information system $(X, \mathcal{A})$, where $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ is the set of attributes, $X = \{x_1, x_2, \ldots, x_7\}$ is the set of objects and $\bar{a} : X \to V_a$ are the mappings given in Table 1.

Moreover, we assume $P = L = [0, 1]$. The fuzzy tolerance relations $R_{a_i} : X \times X \to [0, 1]$ are defined as $R_{a_i}(x, y) = 1 - |R(a_i, x) - R(a_i, y)|$, for all $i \in \{1, \ldots, 4\}$ and $x, y \in X$. The aggregation operator $@ : [0, 1]^4 \to [0, 1]$ is given by

$$@(l_1, l_2, l_3, l_4) = \frac{1}{6}\big(l_1 + l_2 + 2(l_3 + l_4)\big)$$

for all $l_1, l_2, l_3, l_4 \in [0, 1]$. This operator considers more important the relation w.r.t. the third and fourth attributes than w.r.t. the first and second one.

In this example, we consider the fuzzy $\mathcal{A}$-indiscernibility relation $R_\mathcal{A} : X \times X \to [0, 1]$, defined, for all $x, y \in X$, as

$$R_\mathcal{A}(x, y) = @\big(R_{a_1}(x, y), \ldots, R_{a_4}(x, y)\big) \tag{17}$$

and given in Table 2.

Table 1
The information system $(X, \mathcal{A})$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
a_1	0.34	0.21	0.52	0.85	0.43	0.21	0.09
a_2	0.13	0.09	0.36	0.17	0.1	0.04	0.06
a_3	0.31	0.71	0.92	0.65	0.89	0.47	0.93
a_4	0.75	0.5	1	1	0.5	0.25	0.25

Table 2
The $\mathcal{A}$-indiscernibility relation $R_\mathcal{A}$.

R_B	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	0.755	0.645	0.712	0.703	0.743	0.573
x_2		1	0.667	0.693	0.902	0.828	0.818
x_3			1	0.823	0.765	0.495	0.625
x_4				1	0.672	0.562	0.512
x_5					1	0.73	0.84
x_6						1	0.823
x_7							1

The algebraic structure assumed is the multi-adjoint property-oriented frame $([0,1],[0,1],[0,1],\&_1)$ with $\&_1 = \&_G$ and the context is $(X, X, R_{\mathcal{A}}, \tau)$, where τ maps all pairs of elements in $X \times X$ to 1. Given the fuzzy subset $h: X \to [0,1]$, defined as

$$\begin{array}{llll} h(x_1) = 0.625, & h(x_2) = 0.25, & h(x_3) = 0.25, & h(x_4) = 0.375 \\ h(x_5) = 0.2, & h(x_6) = 0.4, & h(x_7) = 0.25 & \end{array}$$

the corresponding multi-adjoint fuzzy rough set is the pair $(h^{\downarrow^N}, h^{\uparrow_\pi})$, where

$$\begin{array}{llll} h^{\downarrow^N}(x_1) = 0.2, & h^{\downarrow^N}(x_2) = 0.2, & h^{\downarrow^N}(x_3) = 0.2, & h^{\downarrow^N}(x_4) = 0.2 \\ h^{\downarrow^N}(x_5) = 0.2, & h^{\downarrow^N}(x_6) = 0.2, & h^{\downarrow^N}(x_7) = 0.2 & \end{array}$$

and

$$\begin{array}{llll} h^{\uparrow_\pi}(x_1) = 0.625, & h^{\uparrow_\pi}(x_2) = 0.625, & h^{\uparrow_\pi}(x_3) = 0.625, & h^{\uparrow_\pi}(x_4) = 0.625 \\ h^{\uparrow_\pi}(x_5) = 0.625, & h^{\uparrow_\pi}(x_6) = 0.625, & h^{\uparrow_\pi}(x_7) = 0.573 & \end{array}$$

Moreover, the multi-adjoint fuzzy rough set is not a multi-adjoint property-oriented concept. For example, the concept $\langle h^{\uparrow_\pi\downarrow^N}, h^{\uparrow_\pi}\rangle$ is given by the mapping $h^{\uparrow_\pi}$ computed above and $h^{\uparrow_\pi\downarrow^N}$ which is:

$$\begin{array}{llll} h^{\uparrow_\pi\downarrow^N}(x_1) = 0.625, & h^{\uparrow_\pi\downarrow^N}(x_2) = 0.573, & h^{\uparrow_\pi\downarrow^N}(x_3) = 0.573, & h^{\uparrow_\pi\downarrow^N}(x_4) = 0.625 \\ h^{\uparrow_\pi\downarrow^N}(x_5) = 0.573, & h^{\uparrow_\pi\downarrow^N}(x_6) = 0.573, & h^{\uparrow_\pi\downarrow^N}(x_7) = 0.573 & \end{array}$$

An interesting property of the proposed framework is that we can consider several adjoint triples and, as a consequence, different preferences among the objects can be assumed. For example, we can assume, in the previous example, that the values with respect to the objects x_5, x_6, x_7 are more trustworthy than the others. Therefore, we can associate x_5, x_6, x_7 with the product adjoint triple instead of the Gödel one, because the Gödel implication results in lower values and has more influence on the infimum in the lower approximation in Eq. (16).

Example 3. We consider the frame $([0,1],[0,1],[0,1],\&_1,\&_2)$, where $\&_1 = \&_G$ and $\&_2 = \&_P$ and the context $(X, X, R_{\mathcal{A}}, \tau)$, from Example 2, where $\tau(x_i, x_j) = 2$ if both x_i and x_j are in $\{x_5, x_6, x_7\}$, and $\tau(x_i, x_j) = 1$ otherwise.

In this case, the values given by $h^{\downarrow^N}$ for $\{x_1, x_2, x_3, x_4\}$, will be the same, since the Gödel implication is only considered. However, the values for x_5, x_6, x_7 can be different. E.g., $h^{\downarrow^N}(x_6)$ is given by:

$$\begin{aligned} h^{\downarrow^N}(x_6) &= \inf\bigl\{h(x_i) \nwarrow_{\tau(x_i,x_6)} R_{\mathcal{A}}(x_i,x_6) \bigm| x_i \in X\bigr\} \\ &= \inf\bigl\{\inf\bigl\{h(x_i) \nwarrow_G R_{\mathcal{A}}(x_i,x_6) \bigm| i \in \{1,2,3,4\}\bigr\}, \inf\bigl\{h(x_i) \nwarrow_P R_{\mathcal{A}}(x_i,x_6) \bigm| i \in \{5,6,7\}\bigr\}\bigr\} \\ &= \inf\{0.625, 0.25, 0.25, 0.375, 0.274, 0.4, 0.304\} = 0.25 \end{aligned}$$

which is closer to $h(x_6)$ than 0.2 (the value obtained considering only the Gödel triple). The new fuzzy subset $h^{\downarrow^N}$ is

$$\begin{array}{llll} h^{\downarrow^N}(x_1) = 0.2, & h^{\downarrow^N}(x_2) = 0.2, & h^{\downarrow^N}(x_3) = 0.2, & h^{\downarrow^N}(x_4) = 0.2 \\ h^{\downarrow^N}(x_5) = 0.2, & h^{\downarrow^N}(x_6) = 0.25, & h^{\downarrow^N}(x_7) = 0.238 & \end{array}$$

which is greater than the previous one. The fuzzy subset $h^{\uparrow_\pi}$ does not change, in this particular example.

This example shows that the use of different adjoint triples provides a more flexible language, considering cases that could not be assumed in other frameworks.

3.2. Properties

In this section we evaluate which properties of the classical lower and upper approximations remain satisfied for multi-adjoint fuzzy rough sets.

First, since the pair ${}^{\uparrow_\pi}: L^X \to L^X$, ${}^{\downarrow^N}: L^X \to L^X$ defined as in Eqs. (15) and (16) forms an isotone Galois connection [18], the monotonicity of ${}^{\uparrow_\pi}$ and ${}^{\downarrow^N}$ holds.

The following proposition shows the conditions which need to hold such that an L-fuzzy subset of X lies between its lower and upper approximations.

Proposition 1. *Given $h \in L^X$, we have that*

(1) *If $h(x) \nwarrow_{xx} \top_P \preccurlyeq h(x)$, for all $x \in X$, then the inequality $h^{\downarrow^N} \preccurlyeq h$ holds.*
(2) *If $h(x) \preccurlyeq \top_P \&_{xx} h(x)$, for all $x \in X$, then $h \preccurlyeq h^{\uparrow_\pi}$ is verified.*

Proof. The first statement follows directly using the hypothesis in the next chain of inequalities:

$$\begin{aligned} h^{\downarrow^N}(y) &= \inf\{h(x) \nwarrow_{xy} R_B(x,y) \mid x \in X\} \\ &\preccurlyeq h(y) \nwarrow_{yy} R_B(y,y) \\ &= h(y) \nwarrow_{yy} \top_P \\ &\preccurlyeq h(y) \end{aligned}$$

Analogously, the second statement is obtained:

$$\begin{aligned} h(x) &\preccurlyeq \top_P \&_{xx} h(x) \\ &= R_B(x,x) \&_{xx} h(x) \\ &\preccurlyeq \sup\{R_B(x,y) \&_{xy} h(y) \mid y \in X\} \\ &= h^{\uparrow_\pi}(x) \qquad \square \end{aligned}$$

Corollary 1. *If $v \nwarrow_{xx} \top_P \preccurlyeq v$ and $u \preccurlyeq \top_P \&_{xx} u$, for all $u, v \in L$ and $x \in X$, then $h^{\downarrow^N} \preccurlyeq h \preccurlyeq h^{\uparrow_\pi}$, for all $h \in L^X$.*

For example, the product, Gödel and Łukasiewicz t-norms, together with their residuated implication satisfy these conditions. Indeed, any t-norm and its residuated implication satisfy the conditions given in the previous corollary.

Next, we study the interaction between the approximation operators and the union/intersection of a family of fuzzy sets. We obtain similar results as in the standard rough set environment.

Proposition 2. *Given $\{h_i\}_{i\in\Lambda}$, with $h_i : X \to L$ for all i in the index set Λ, the following properties are verified:*

(1) $(\inf\{h_i \mid i \in \Lambda\})^{\downarrow^N} = \inf\{h_i^{\downarrow^N} \mid i \in \Lambda\}$.
(2) $(\inf\{h_i \mid i \in \Lambda\})^{\uparrow_\pi} \preccurlyeq \inf\{h_i^{\uparrow_\pi} \mid i \in \Lambda\}$.
(3) $\sup\{h_i^{\downarrow^N} \mid i \in \Lambda\} \preccurlyeq (\sup\{h_i \mid i \in \Lambda\})^{\downarrow^N}$.
(4) $(\sup\{h_i \mid i \in \Lambda\})^{\uparrow_\pi} = \sup\{h_i^{\uparrow_\pi} \mid i \in \Lambda\}$.

Proof. Statement (1) is obtained from the following chain of equalities.

$$\begin{aligned} \big(\inf\{h_i \mid i \in \Lambda\}\big)^{\downarrow^N}(y) &= \inf\big\{\big(\inf\{h_i \mid i \in \Lambda\}\big)(x) \nwarrow_{xy} R_B(x,y) \mid x \in X\big\} \\ &\overset{(*)}{=} \inf\big\{\big(\inf\{h_i(x) \nwarrow_{xy} R_B(x,y) \mid i \in \Lambda\}\big) \mid x \in X\big\} \\ &= \inf\big\{\big(\inf\{h_i(x) \nwarrow_{xy} R_B(x,y) \mid x \in X\}\big) \mid i \in \Lambda\big\} \\ &= \inf\big\{h_i^{\downarrow^N}(y) \mid i \in \Lambda\big\} \end{aligned}$$

where $(*)$ is given by Lemma 1(2).

Statement (2) holds directly. Since $(\inf\{h_i \mid i \in \Lambda\}) \preccurlyeq h_i$, for all $i \in \Lambda$, and, applying the monotonicity of ${}^{\uparrow_\pi}$, we obtain $(\inf\{h_i \mid i \in \Lambda\})^{\uparrow_\pi} \preccurlyeq h_i^{\uparrow_\pi}$. Therefore, applying the infimum property, we obtain

$$\big(\inf\{h_i \mid i \in \Lambda\}\big)^{\uparrow_\pi} \preccurlyeq \inf\big\{h_i^{\uparrow_\pi} \mid i \in \Lambda\big\}$$

Statements (3) and (4) follow similarly. $\square$

Additional properties follow from the isotone Galois connection definition. Specifically, we have that $h \preccurlyeq h^{\uparrow_\pi \downarrow^N}$, for all $h \in L^X$, and that $h^{\downarrow^N \uparrow_\pi} \preccurlyeq h$, for all $h \in L^X$, which proves that the composition ${}^{\uparrow_\pi \downarrow^N}$ is a closure operator and ${}^{\downarrow^N \uparrow_\pi}$ is an interior operator.

Not all properties that hold for classical rough sets and fuzzy rough sets hold for multi-adjoint fuzzy rough sets. For instance, for $h \in L^X$, the following equations do not hold:

$$\big(h^{\downarrow^N}\big)^{\uparrow^\pi} = \big(h^{\downarrow^N}\big)^{\downarrow^N} = h^{\downarrow^N} \tag{18}$$

$$\big(h^{\uparrow^\pi}\big)^{\uparrow^\pi} = \big(h^{\uparrow^\pi}\big)^{\downarrow^N} = h^{\uparrow^\pi} \tag{19}$$

Table 3
The $\mathcal{A}$-indiscernibility relation $R_{\mathcal{A}}$.

$R_{\mathcal{A}}$	x_1	x_2	x_3
x_1	1	0.9	0.9
x_2		1	0.9
x_3			1

In the next example, we show that $(h^{\downarrow^N})^{\downarrow^N} = h^{\downarrow^N}$ is not true in general. Note that, in fuzzy rough set theory, this equality holds in the specific case R is a similarity relation[7] and the implication is continuous. Therefore, we construct a counterexample with continuous implications and with a similarity relation.

Example 4. Consider the class of implications $\nwarrow_L^\alpha$ defined by $z \nwarrow_L^\alpha x = \sqrt[\alpha]{\min\{1, 1 + z^\alpha - x^\alpha\}}$ with $\alpha \geqslant 1$, for all x, z in $[0, 1]$. We consider an information system $(X, \mathcal{A})$ with three objects $X = \{x_1, x_2, x_3\}$ for which the similarity relation $R_{\mathcal{A}}$ is given in Table 3, and the fuzzy subset h:

$$h(x_1) = 0.4, \qquad h(x_2) = 0.1, \qquad h(x_3) = 0.8$$

The flexibility of the multi-adjoint fuzzy rough model allows us to use different implications for different pairs of instances. We use the following implications in our example:

$$\begin{array}{lll} \nwarrow_{x_1x_1} = \nwarrow_L^1 & \nwarrow_{x_1x_2} = \nwarrow_L^2 & \nwarrow_{x_1x_3} = \nwarrow_L^5 \\ \nwarrow_{x_2x_1} = \nwarrow_L^5 & \nwarrow_{x_2x_2} = \nwarrow_L^1 & \nwarrow_{x_2x_3} = \nwarrow_L^1 \\ \nwarrow_{x_3x_1} = \nwarrow_L^1 & \nwarrow_{x_3x_2} = \nwarrow_L^5 & \nwarrow_{x_3x_3} = \nwarrow_L^1 \end{array}$$

Then the lower approximation of h is given by:

$$h^{\downarrow^N}(x_1) = 0.4, \qquad h^{\downarrow^N}(x_2) = 0.1, \qquad h^{\downarrow^N}(x_3) = 0.2$$

while applying the lower approximation twice results in:

$$\left(h^{\downarrow^N}\right)^{\downarrow^N}(x_1) = 0.3, \qquad \left(h^{\downarrow^N}\right)^{\downarrow^N}(x_2) = 0.1, \qquad \left(h^{\downarrow^N}\right)^{\downarrow^N}(x_3) = 0.2$$

Therefore, $(h^{\downarrow^N})^{\downarrow^N} \neq h^{\downarrow^N}$.

Another property in classical rough set theory is that the lower and upper approximations are dual, that is, $co(R_B{\downarrow}A) = R_B{\uparrow}co(A)$, where co represents set complement. This property is not maintained for multi-adjoint fuzzy rough sets.

Example 5. Given the implications $\nwarrow_L^\alpha$ from Example 4, the adjoint conjunctors $\&_L^\alpha$ are given by $x \&_L^\alpha y = \max\{0, \sqrt[\alpha]{x^\alpha + y^\alpha - 1}\}$. We also consider the complement $co(h)$ of $h \in X^{[0,1]}$, defined by $(co(h))(x) = 1 - h(x)$, for all $x \in [0, 1]$. For the fuzzy subset h defined in Example 4, we have that $(co(h))^{\downarrow^N} \neq co(h^{\uparrow^\pi})$, that is

$$\begin{aligned} \big(co(h)\big)^{\downarrow^N}(x) &= \inf\big\{\big(co(h)\big)(x_i) \nwarrow_{x_i,x} R_{\mathcal{A}}(x_i, x) \;\big|\; x_i \in X\big\} \\ &\neq 1 - \sup\big\{R_{\mathcal{A}}(x, x_i) \&_{x,x_i} h(x_i) \;\big|\; x_i \in X\big\} \\ &= \big(co\big(h^{\uparrow^\pi}\big)\big)(x) \end{aligned}$$

for all $x \in X$.

This holds because, for example, $(co(h))^{\downarrow^N}(x_1) = 0.3$ and, on the other hand, $h^{\uparrow^\pi}(x_1) = 0.4$, and so $(co(h^{\uparrow^\pi}))(x_1) = 0.6$, which shows that both mappings are not equal.

In conclusion, we can assert that many properties of classical rough set theory are satisfied in the general framework of multi-adjoint fuzzy rough sets, such as monotonicity, inclusion, conjunction, disjunction, etc. Moreover, although the previous examples show that some classical rough set properties are violated for multi-adjoint fuzzy rough sets, this does not harm attribute selection as we will see in the remainder of the paper.

[7] A similarity relation is a reflexive and symmetrical fuzzy relation R that also satisfies $\min\{R(x, y), R(y, z)\} \leqslant R(x, z)$, for all $x, y, z \in X$.

3.3. Multi-adjoint fuzzy positive region

We use the necessity operator to define a generalization of the positive region to the multi-adjoint fuzzy rough set framework. Analogously to (11), the *multi-adjoint fuzzy B-positive region* can be defined, for each $y \in X$, as

$$\begin{aligned} POS_B(y) &= \left(\bigcup_{x\in X} (R_d x)^{\downarrow^N}\right)(y) \\ &= \sup_{x\in X} \inf_{z\in X} R_d(x,z) \nwarrow_{zy} R_B(z,y) \end{aligned} \tag{20}$$

where $R_d : X \times X \to L$ is an L-fuzzy relation. In case d is crisp, $R_d(x,y)$ takes two values, namely $\top_L$ if $\bar{d}(x) = \bar{d}(y)$ and $\bot_L$ else. In [3], the authors showed that when d is crisp, only the decision class that y belongs to needs to be inspected; we therefore introduce a simpler variant of the fuzzy rough positive region as, for each $y \in X$,

$$POS'_B(y) = (R_d y)^{\downarrow^N}(y) = \inf_{x\in X} R_d(y,x) \nwarrow_{xy} R_B(x,y) \tag{21}$$

Proposition 3. *For $y \in X$, if R_d is a crisp relation, then*

$$POS_B(y) = (R_d y)^{\downarrow^N}(y) = POS'_B(y)$$

Proof. Given $x \in X$, such that $x \notin R_d y$, we have that

$$\begin{aligned} \inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} &\leqslant R_d(x,y) \nwarrow_{yy} R_B(y,y) \\ &\overset{(*)}{=} \bot_L \nwarrow_{yy} \top_P \\ &= \bot_L \end{aligned}$$

where $(*)$ is provided because R_B is reflexive and $x \notin R_d y$ (and so $R_d(x,y) = \bot_L$). Therefore, we obtain

$$\sup\{\inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid x' \in X\} \mid x \notin R_d y\} = \bot_L$$

This result is used in the following chain of equalities:

$$\begin{aligned} POS_B(y) &= \sup\{\inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} \mid x \in X\} \\ &= \sup\{\sup\{\inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} \mid x \in R_d y\}, \\ &\quad \sup\{\inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} \mid x \notin R_d y\}\} \\ &\overset{(1)}{=} \sup\{\sup\{\inf\{R_d(x,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} \mid x \in R_d y\}, \bot_L\} \\ &\overset{(2)}{=} \inf\{R_d(y,z) \nwarrow_{zy} R_B(z,y) \mid z \in X\} \\ &= (R_d y)^{\downarrow^N}(y) \end{aligned}$$

where (1) is given by the previous comment and (2) follows from the equality $R_d(x,z) = R_d(y,z)$, which is true when $x \in R_d y$. □

4. Multi-adjoint fuzzy decision reducts

In this section, two approaches to obtain multi-adjoint fuzzy decision reducts are presented. One of them is based on the multi-adjoint fuzzy rough positive region, the other one on a generalization of the fuzzy discernibility function. Finally, a relation between them will be introduced.

4.1. Measures

First of all, the L-valued measure definition must be introduced. These operators will be used to evaluate subsets of $\mathcal{A}$ w.r.t. their ability to maintain discernibility relative to the decision attribute.

Definition 6. A monotonic mapping $m : \mathcal{P}(\mathcal{A}) \to L$ is an *L-valued measure associated with the decision system* $(X, \mathcal{A} \cup \{d\})$ if the condition

$$R_d(x,y) \nwarrow_{xy} R_B(x,y) = R_d(x,y) \nwarrow_{xy} R_{\mathcal{A}}(x,y), \quad \text{for all } x, y \in X \tag{22}$$

implies $m(B) = \top_L$.

Table 4
The decision system $(X, \mathcal{A} \cup \{d\})$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
a_1	0.34	0.21	0.52	0.85	0.43	0.21	0.09
a_2	0.13	0.09	0.36	0.17	0.1	0.04	0.06
a_3	0.31	0.71	0.92	0.65	0.89	0.47	0.93
a_4	0.75	0.5	1	1	0.5	0.25	0.25
d	0	1	0	1	0	1	0

Once such a measure is obtained, we can associate a notion of fuzzy decision reduct with it.

Definition 7 *(Fuzzy m-decision reduct to degree α).* Let $m : \mathcal{P}(\mathcal{A}) \to L$ be an L-valued measure associated with the decision system $(X, \mathcal{A} \cup \{d\})$, $B \subseteq \mathcal{A}$ and $\alpha \in L$, with $\alpha \neq \bot_L$. The set B is called a fuzzy m-decision superreduct to degree α if $\alpha \preccurlyeq m(B)$.

It is called a fuzzy m-decision reduct to degree α if, moreover, for all $B' \subset B$, $\alpha \not\preccurlyeq m(B')$.

4.2. Multi-adjoint fuzzy rough positive region based measures

Using the definitions of multi-adjoint fuzzy rough positive region in Eqs. (20) and (21), we can define increasing L-valued measures to implement the corresponding notion of fuzzy decision reducts.

In [12], the authors use fuzzy implications and the fuzzy definition of cardinality to define the degree of dependency, which is a concept that we want to use to define L-valued measures. Therefore, from now on we assume that $L = [0, 1]$. The cardinality of a fuzzy set $C : Y \to [0, 1]$ is defined as:

$$|C| = \sum_{y \in Y} C(y) \tag{23}$$

The most obvious way to define a $[0, 1]$-valued measure is to introduce a normalized extension of the degree of dependency, i.e., the mappings $\gamma' : \mathcal{P}(\mathcal{A}) \to [0, 1]$ and $\gamma : \mathcal{P}(\mathcal{A}) \to [0, 1]$ defined for all subsets B of $\mathcal{A}$ as

$$\gamma'_B = \mathcal{I}\left(\frac{|POS'_{\mathcal{A}}|}{|X|}, \frac{|POS'_B|}{|X|}\right)$$
$$\gamma_B = \mathcal{I}\left(\frac{|POS_{\mathcal{A}}|}{|X|}, \frac{|POS_B|}{|X|}\right)$$

where $\mathcal{I}$ is a fuzzy implication.

These measures resemble the one introduced by Jensen and Shen in [12] and are illustrated in the following example.

Example 6. We expand the information system $(X, \mathcal{A})$ from Example 2 with a decision attribute d as shown in Table 4. In the same way as $R_{\mathcal{A}}$ was obtained, the fuzzy indiscernibility relations $R_{\{a_1,a_2\}}$, R_{a_1} and R_{a_2} are computed.

The algebraic structure assumed is the multi-adjoint property-oriented frame $([0,1],[0,1],[0,1],\&_1, \&_2, \&_3)$ with $\&_1 = \&_L^1$, $\&_2 = \&_L^3$ and $\&_3 = \&_L^5$, and the context is $(X, X, R_{\mathcal{A}}, \tau)$, where τ is defined for all $x_i, x_j \in X$ as follows:

$$\tau(x_i, x_j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are even} \\ 2 & \text{if } i \text{ and } j \text{ are odd} \\ 3 & \text{otherwise} \end{cases}$$

In this setting, the cardinalities of the fuzzy positive regions with respect to $\mathcal{A}$, $\{a_1, a_2\}$, $\{a_1\}$ and $\{a_2\}$ are approximately:

$$|POS'_{\mathcal{A}}| = 6.25, \qquad |POS'_{\{a_1,a_2\}}| = 4.95, \qquad |POS'_{\{a_1\}}| = 3.83, \qquad |POS'_{\{a_2\}}| = 4.72$$

Therefore, using $\mathcal{I} = \nwarrow_L$, we approximately obtain:

$$\gamma'_{\mathcal{A}} = 1, \qquad \gamma'_{\{a_1,a_2\}} = 0.81, \qquad \gamma'_{\{a_1\}} = 0.78, \qquad \gamma'_{\{a_2\}} = 0.65$$

Consequently, we can assert for example that $B = \{a_1, a_2\}$ is a fuzzy γ'-decision reduct to degree 0.8, but not to degree 0.7 (because $\gamma'_{\{a_1\}} > 0.7$) nor to degree 0.9.

Alternatively, the measures $\delta' : \mathcal{P}(\mathcal{A}) \to [0, 1]$ and $\delta : \mathcal{P}(\mathcal{A}) \to [0, 1]$ are defined for each subset B of $\mathcal{A}$ as:

$$\delta'_B = \mathcal{I}\left(\min_{x \in X} POS'_{\mathcal{A}}(x), \min_{x \in X} POS'_B(x)\right)$$
$$\delta_B = \mathcal{I}\left(\min_{x \in X} POS_{\mathcal{A}}(x), \min_{x \in X} POS_B(x)\right)$$

These measures are inspired by the fact that in standard rough set theory, the property of being a (super)reduct is also determined by the worst object.

Example 7. Continuing Example 6, the minimal membership values to the corresponding positive regions are:

$$\min_{x\in X} POS'_{\mathcal{A}}(x) = 0.83, \qquad \min_{x\in X} POS'_{\{a_1,a_2\}}(x)' = 0.65$$
$$\min_{x\in X} POS'_{\{a_1,a_2\}}(x)' = 0.63, \qquad \min_{x\in X} POS'_{\{a_1,a_2\}}(x)' = 0.38$$

and, therefore, the values with respect to the measure δ', using $\nwarrow_L$ as the fuzzy implication, are:

$$\delta'_{\mathcal{A}} = 1, \qquad \delta'_{\{a_1,a_2\}} = 0.81, \qquad \delta'_{\{a_1\}} = 0.79, \qquad \delta'_{\{a_2\}} = 0.55$$

Thus, the subset $B = \{a_1, a_2\}$ is a fuzzy δ'-decision reduct to degree 0.8, for example.

A general operator which generalizes both measures above is the OWA operator. This operator will only be introduced considering POS'_B, the definition using POS_B can be given similarly.

Let $W = \{w_1, \dots, w_n\}$ be a list of weights such that $w_i \in [0,1]$ and $\sum_i^n w_i = 1$, and define two permutations ρ_1 and ρ_2 on $\{1, \dots, n\}$ such that for all i, j in $\{1, \dots, n\}$:

$$\rho_1(i) < \rho_1(j)$$

if and only if

$$POS'_B(x_{\rho_1(i)}) < POS'_B(x_{\rho_1(j)})$$

or

$$POS'_B(x_{\rho_1(i)}) = POS'_B(x_{\rho_1(j)}) \wedge i < j$$

The permutation ρ_2 is defined similarly with respect to $POS'_{\mathcal{A}}$. The OWA $[0,1]$-valued measure is the operator $\mathrm{MOWA}'_W : \mathcal{P}(\mathcal{A}) \to [0,1]$, defined, for each subset B of $\mathcal{A}$, as

$$\mathrm{MOWA}'_W(B) = \mathcal{I}\left(\sum_{i=1}^{n} w_i \cdot POS'_{\mathcal{A}}(x_{\rho_2(i)}), \sum_{i=1}^{n} w_i \cdot POS'_B(x_{\rho_1(i)})\right) \tag{24}$$

where $\mathcal{I}$ is a fuzzy implication.

The mappings γ' and δ' are particular cases of the $[0,1]$-valued measure MOWA'_W. The measure γ' is given considering $w_i = 1/n$, for all $i \in \{1, \dots, n\}$, and δ' considering $w_1 = 1$ and $w_i = 0$, for all $i \in \{2, \dots, n\}$.

This operator is indeed a $[0,1]$-valued measure associated with the decision system $(X, \mathcal{A} \cup \{d\})$, by the properties of the OWA operator, as shown in the next easily verified proposition.

Proposition 4. *Given two subsets B_1, B_2 of $\mathcal{A}$, such that $B_1 \subseteq B_2$, we obtain that $\mathrm{MOWA}'_W(B_1) \leqslant \mathrm{MOWA}'_W(B_2)$. Moreover, if Eq. (22) is satisfied for $B \subseteq \mathcal{A}$, then $\mathrm{MOWA}'_W(B) = \mathrm{MOWA}'_W(\mathcal{A}) = \top_L$.*

Proof. Straightforward. □

The following example defines a particular MOWA operator and obtains the measure for the subsets considered in Example 6.

Example 8. Consider $\mathrm{MOWA}'_W : \mathcal{P}(\mathcal{A}) \to [0,1]$, using $\nwarrow_L$ as the fuzzy implication, and $W = \{0.3, 0.25, 0.2, 0.15, 0.1, 0\}$, such that it mimics δ' definition.

Hence, in the setting of Example 6, the following values are obtained:

$$\mathrm{MOWA}'_W\big(\{a_1,a_2\}\big) = 0.79, \qquad \mathrm{MOWA}'_W\big(\{a_1\}\big) = 0.77, \qquad \mathrm{MOWA}'_W\big(\{a_2\}\big) = 0.56$$

Therefore, the subset $B = \{a_1, a_2\}$ is a fuzzy MOWA'_W-decision reduct to degree 0.78, for example.

4.3. Fuzzy discernibility function

In order to generalize the discernibility function (7) in a fuzzy environment, the idea in [3] will be used here. We generalize the discernibility function using the fuzzy tolerance relations R_B that represent objects' approximate equality. For each subset B of attributes, a value between the minimum and the maximum of the lattice will be obtained, indicating how well these attributes maintain the discernibility, relative to the decision attribute, among all objects. Next, we use this function to define an $[0,1]$-valued measure.

Eq. (7) can be rewritten as

$$\begin{aligned} f\left(a_1^*,\ldots,a_m^*\right) &= \bigwedge\left\{\bigvee_{k=1}^{m} a_k^*\left[d(x_i)\neq d(x_j)\Rightarrow a_k(x_i)\neq a_k(x_j)\right]\,\middle|\, 1\leqslant i<j\leqslant n\right\} \\ &= \bigwedge\left\{\bigvee_{k=1}^{m} a_k^*\left[a_k(x_i)=a_k(x_j)\Rightarrow d(x_i)=d(x_j)\right]\,\middle|\, 1\leqslant i<j\leqslant n\right\} \\ &= \bigwedge\left\{\left[\bigwedge_{a_k^*=1}\left(a_k(x_i)=a_k(x_j)\right)\right]\Rightarrow d(x_i)=d(x_j)\,\middle|\, 1\leqslant i<j\leqslant n\right\} \end{aligned} \tag{25}$$

provided the decision system is consistent.[8] Note that each series of boolean values $a_1^*,\ldots,a_m^*$ corresponds to the subset of features B that contains those features a for which $a^*=1$. We can extend the discernibility function to a mapping from $\mathcal{P}(\mathcal{A})$ to L interpreting the infimum in Eq. (25) by an aggregation operator @, replacing the exact equalities by the respective approximate equalities (fuzzy indiscernibility relations), and writing the expression

$$\bigwedge_{a_k^*=1}\left(a_k(x_i)=a_k(x_j)\right)$$

as the value $R_B(x_i,x_j)$, where R_B is defined by Eq. (14). Hence, this general function is:

$$\begin{aligned} f_{@}(B) &= @\Big(\underbrace{c_{ij}(B)}_{1\leqslant i<j\leqslant n}\Big) \\ &= @\big(c_{12}(B),\ldots,c_{1n}(B),c_{23}(B),\ldots,c_{(n-1)n}(B)\big) \end{aligned} \tag{26}$$

with

$$c_{ij}(B)=R_d(x_i,x_j)\nwarrow_{x_i x_j} R_B(x_i,x_j) \tag{27}$$

By the monotonicity properties of the implications $\nwarrow_{x_i x_j}$, the degree to which a subset B serves to distinguish between objects x_i and x_j increases as their approximate equality $R_B(x_i,x_j)$ w.r.t. B decreases, and their approximate equality $R_d(x_i,x_j)$ w.r.t. d increases. Therefore, $f_{@}(B)$, expresses the degree to which, for all object pairs, different values in attributes of B correspond to different values of d.

The previous definition of discernibility function is a fuzzy generalization of the crisp one (25) and the fuzzy ones introduced in [3], assuming consistency.

We now consider two particular cases of $f_{@}$. First, we replace the aggregation operator @ by a t-norm $\mathcal{T}$ and define $f_{\mathcal{T}}:\mathcal{P}(\mathcal{A})\to[0,1]$ as:

$$f_{\mathcal{T}}(B)=\mathcal{T}\Big(\underbrace{c_{ij}(B)}_{1\leqslant i<j\leqslant n}\Big) \tag{28}$$

Secondly, we replace the aggregation operator @ in Eq. (26) by the average, denoted by $@_{\text{av}}$.

$$\begin{aligned} f_{@_{\text{av}}}(B) &= @_{\text{av}}\Big(\underbrace{c_{ij}(B)}_{1\leqslant i<j\leqslant n}\Big) \\ &= \frac{2\cdot\sum_{1\leqslant i<j\leqslant n} c_{ij}(B)}{n(n-1)} \end{aligned} \tag{29}$$

[8] Recall that if $(X,\mathcal{A}\cup\{d\})$ is inconsistent, there exist x_i and x_j such that, for all $a\in\mathcal{A}$, $a(x_i)=a(x_j)$, yet $d(x_i)\neq d(x_j)$. Such x_i and x_j are not considered in Eq. (7), since $O_{ij}=\emptyset$.

A fuzzy discernibility function $f_@$ can be used, together with a fuzzy implication $\mathcal{I}$, to define the following evaluation measure $\varepsilon : \mathcal{P}(\mathcal{A}) \to [0,1]$, defined, for all $B \subseteq \mathcal{A}$, as

$$\varepsilon_@(B) = \mathcal{I}\big(f_@(\mathcal{A}), f_@(B)\big) \tag{30}$$

As particular cases we have:

$$\varepsilon_{\mathcal{T}}(B) = \mathcal{I}\big(f_{\mathcal{T}}(\mathcal{A}), f_{\mathcal{T}}(B)\big)$$
$$\varepsilon_{@_{\mathrm{av}}}(B) = \mathcal{I}\big(f_{@_{\mathrm{av}}}(\mathcal{A}), f_{@_{\mathrm{av}}}(B)\big)$$

The following proposition shows that, given an aggregation operator @, fuzzy $\varepsilon_@$-decision reducts can be considered, that is, $\varepsilon_@$ is a $[0,1]$-valued measure w.r.t. the decision system.

Proposition 5. *For subsets B_1, B_2 of $\mathcal{A}$, if $B_1 \subseteq B_2$, then $\varepsilon_@(B_1) \leqslant \varepsilon_@(B_2)$. Moreover, if Eq. (22) is satisfied for $B \subseteq \mathcal{A}$, then $\varepsilon_@(B) = \varepsilon_@(\mathcal{A}) = 1$.*

Proof. Straightforward. □

To conclude this section, an example of the ε-measures is introduced.

Example 9. In the environment of Example 6, considering the average $@_{\mathrm{av}}$, and the minimum t-norm $\&_G$ we obtain:

$$\begin{array}{lll} \varepsilon_{\&_G}\big(\{a_1,a_2\}\big) = 0.90, & \varepsilon_{\&_G}\big(\{a_1\}\big) = 0.87, & \varepsilon_{\&_G}\big(\{a_2\}\big) = 0.78 \\ \varepsilon_{@_{\mathrm{av}}}\big(\{a_1,a_2\}\big) = 0.92, & \varepsilon_{@_{\mathrm{av}}}\big(\{a_1\}\big) = 0.92, & \varepsilon_{@_{\mathrm{av}}}\big(\{a_2\}\big) = 0.90 \end{array}$$

In this case, $B = \{a_1, a_2\}$ is a fuzzy $\varepsilon_{\&_G}$-decision reduct to degree 0.89, for example. On the other hand, B is a fuzzy $\varepsilon_{@_{\mathrm{av}}}$-decision superreduct to degree 0.9, for example; however, it is not a $\varepsilon_{@_{\mathrm{av}}}$-decision reduct for any value of α.

4.4. Relationships between fuzzy decision reducts

The particular cases of evaluation measures presented in the previous sections, γ, γ', δ, δ', satisfy the same properties as the corresponding ones introduced in [3]. That is $\delta'_B \leqslant \gamma'_B \leqslant \gamma_B$ and $\delta'_B \leqslant \delta_B \leqslant \gamma_B$ always hold, and $\gamma_B = \gamma'_B$ and $\delta_B = \delta'_B$ when the decision attribute is qualitative.

Moreover, a number of interesting relationships hold between the approaches based on the fuzzy positive region and those based on the fuzzy discernibility function, which are summed up by the following propositions; assuming the same aggregation operator and implication.

Proposition 6. *If $POS'_{\mathcal{A}} = X$, then $\varepsilon_{\mathcal{T}}(B) \leqslant \delta'_B$ and $\gamma'_B \leqslant \varepsilon_{@_{\mathrm{av}}}(B)$, for $B \subseteq \mathcal{A}$. Moreover, in case $\mathcal{T}$ is the minimum t-norm, $\varepsilon_{\mathcal{T}}(B) = \delta'_B$, regardless of $POS'_{\mathcal{A}} = X$.*

The next proposition shows that a crisp $\varepsilon_{@_{\mathrm{av}}}$-decision reduct is a crisp γ/γ'-decision reduct, for consistent data.

Proposition 7. *If $POS'_{\mathcal{A}} = X$ and $\varepsilon_{@_{\mathrm{av}}}(B) = 1$, then $\gamma'_B = 1$ and $\gamma_B = 1$, for any $B \subseteq \mathcal{A}$.*

The proofs of both propositions are analogous to the ones given in [3]. These results show that $\varepsilon_{\mathcal{T}}$ and δ are essentially built upon the same idea, with some variations due to the parameter choice, and also reveal the essential difference between γ and $\varepsilon_{@_{\mathrm{av}}}$: while the former looks at the lowest value of the formula $R_d(x,y) \nwarrow_{xy} R_B(x,y)$ for each y (reflecting to what extent there exists an x that has similar values for all the attributes in B, but a different decision), and averages over these values, the latter evaluates all pairwise evaluations of this formula.

In the general case, where the measures MOWA'_W and $\varepsilon_@$ are assumed, we obtain that MOWA'_W is a particular case of the measure $\varepsilon_@$ based on the fuzzy discernibility function.

Theorem 1. *Given $\mathrm{MOWA}'_W : \mathcal{P}(\mathcal{A}) \to [0,1]$, defined by Eq. (24), $\varepsilon_@ : \mathcal{P}(\mathcal{A}) \to [0,1]$, defined by Eq. (30), and $\sum_{i=1}^n w_i \cdot POS'_{\mathcal{A}}(x_{\sigma(i)}) = 1$, $f_@(\mathcal{A}) = 1$ are satisfied. For each $B \subseteq \mathcal{A}$, there exists an aggregation operator @ such that, the following equality is obtained*

$$\mathrm{MOWA}'_W(B) = \varepsilon_@(B)$$

Proof. As $\sum_{i=1}^{n} w_i \cdot POS'_{\mathcal{A}}(x_{\sigma(i)}) = 1$, $f_{@}(\mathcal{A}) = 1$ and $\mathcal{I}$ is a fuzzy implication, the following chain of equalities can be written:

$$\begin{aligned}
\mathrm{MOWA}'_W(B) &= \sum_{i=1}^{n} w_i \cdot POS'_B(x_{\tau(i)}) \\
&= \sum_{i=1}^{n} w_i \cdot \inf\bigl\{R_d\bigl(x_{\tau(i)}, x'\bigr) \nwarrow_{x' x_{\tau(i)}} R_B\bigl(x', x_{\tau(i)}\bigr) \bigm| x' \in X\bigr\} \\
&\overset{(1)}{=} \sum_{i=1}^{n} w_i \cdot \min\bigl\{R_d\bigl(x_{\tau(i)}, x'\bigr) \nwarrow_{x' x_{\tau(i)}} R_B\bigl(x', x_{\tau(i)}\bigr) \bigm| x' \in X\bigr\} \\
&\overset{(2)}{=} \sum_{i=1}^{n} w_i \cdot \bigl(R_d(x_{\tau(i)}, x_{j_{\tau(i)}}) \nwarrow_{x' x_{\tau(i)}} R_B(x_{j_{\tau(i)}}, x_{\tau(i)})\bigr) \\
&= \sum_{1 \leqslant i < j \leqslant n} w'_{ij} \cdot \bigl(R_d(x_i, x_j) \nwarrow_{x_i, x_j} R_B(x_i, x_j)\bigr) \\
&\overset{(3)}{=} @_B\bigl(c_{12}(B), \ldots, c_{1n}(B), c_{23}(B), \ldots, c_{(n-1)n}(B)\bigr) \\
&= \varepsilon_{@}(B)
\end{aligned}$$

where (1) is given since the unit interval $[0, 1]$ is considered and X is finite; (2) is true if $x_{j_{\tau(i)}}$ is an element in which the minimum element is obtained, for each $\tau(i)$; and w'_{ij} is defined recursively, for each $k \in \{1, \ldots, n\}$, as follows:

If $k = 1$, then there exists $i \in \{1, \ldots, n\}$ such that $\tau(i) = 1$, and we need to consider two possible cases.

- If $\tau(i) = j_{\tau(i)} = 1$, then $R_d(x_{\tau(i)}, x') \nwarrow_{x' x_{\tau(i)}} R_B(x', x_{\tau(i)}) = 1$, for all $x' \in X$, and we define w'_{1j} as
$$w'_{1j} = \begin{cases} w_1 & \text{if } j = 2 \\ 0 & \text{otherwise} \end{cases}$$
- If $1 = \tau(i) < j_{\tau(i)}$, then
$$w'_{1j} = \begin{cases} w_1 & \text{if } j = j_{\tau(i)} \\ 0 & \text{otherwise} \end{cases}$$

Now, if we have defined until $w_{k-1,j}$, with $1 < k < n$, then we define w_{kj}.

- If $k = \tau(i) = j_{\tau(i)}$, then
$$w'_{kj} = \begin{cases} w_k & \text{if } j = k + 1 \\ 0 & \text{otherwise} \end{cases}$$
- If $k = \tau(i) < j_{\tau(i)}$, then
$$w'_{kj} = \begin{cases} w_k & \text{if } j = j_{\tau(i)} \\ 0 & \text{otherwise} \end{cases}$$
- If $k = \tau(i) > j_{\tau(i)}$, then $w_{kj} = 0$, for all $j \in \{1, \ldots, n\}$, and the value $w'_{j_{\tau(i)},\tau(i)}$ is rewritten as $w'_{j_{\tau(i)},\tau(i)} = w'_{j_{\tau(i)},\tau(i)} + w_k$.

When $k = n$, two cases must be considered:

- If $n = \tau(i) = j_{\tau(i)}$, then the value $w'_{n-1,n}$ is rewritten as $w'_{n-1,n} = w'_{n-1,n} + w_n$.
- If $n = \tau(i) > j_{\tau(i)}$, then the value $w'_{j_n,n}$ is rewritten as $w'_{j_n,n} = w'_{j_n,n} + w_n$.

Finally, (3) is obtained assuming @ as the weighted average obtained from the weights w'_{ij}, which is a special case of an aggregation operator. □

5. Conclusion

In this paper, we have introduced multi-adjoint fuzzy rough sets: an extension of the well-known implication/t-norm based fuzzy rough set model based on a family of adjoint pairs to compute the lower and upper approximations. Our model allows to represent explicit preferences among the objects in a decision system, by associating a particular adjoint triple with any pair of objects. We have pointed out the relationships and differences of our model w.r.t. property-oriented concept lattices, verified mathematical properties of the model, and discussed attribute selection in this framework.

References

[1] Y. Chen, Y. Yao, A multiview approach for intelligent data analysis based on data operators, Inf. Sci. 178 (1) (2008) 1–20.
[2] M. Cornejo, J. Medina, E. Ramírez, A comparative study of adjoint triples, Fuzzy Sets Syst. 211 (2013) 1–14.
[3] C. Cornelis, R. Jensen, G. Hurtado, D. Ślezak, Attribute selection with fuzzy decision reducts, Inf. Sci. 180 (2010) 209–224.
[4] I. Düntsch, G. Gediga, Approximation operators in qualitative data analysis, in: Theory and Applications of Relational Structures as Knowledge Instruments, 2003, pp. 214–230.
[5] A. Formica, Concept similarity in fuzzy formal concept analysis for semantic web, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 18 (2) (2010) 153–167.
[6] B. Galitsky, J.D.L. Rosa, Concept-based learning of human behavior for customer relationship management, Inf. Sci. 181 (10) (2011) 2016–2035.
[7] B. Ganter, S.O. Kuznetsov, Scale coarsening as feature selection, Lect. Notes Comput. Sci. 4933 (2008) 217–228.
[8] B. Ganter, C. Meschke, A formal concept analysis approach to rough data tables, in: Lect. Notes Comput. Sci., vol. 5908, 2009, pp. 117–126.
[9] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundation, Springer-Verlag, 1999.
[10] G. Gediga, I. Düntsch, Modal-style operators in qualitative data analysis, in: Proc. IEEE Int. Conf. on Data Mining, 2002, pp. 155–162.
[11] P. Hájek, Metamathematics of Fuzzy Logic. Trends in Logic, Kluwer Academic, 1998.
[12] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Trans. Fuzzy Syst. 15 (1) (2007) 73–89.
[13] R. Jensen, Q. Shen, New approaches to fuzzy-rough feature selection, IEEE Trans. Fuzzy Syst. 17 (4) (2009) 824–838.
[14] H. Lai, D. Zhang, Concept lattices of fuzzy contexts: Formal concept analysis vs. rough set theory, Int. J. Approx. Reason. 50 (5) (2009) 695–707.
[15] J. Li, C. Mei, Y. Lv, Incomplete decision contexts: Approximate concept construction, rule acquisition and knowledge reduction, Int. J. Approx. Reason. 54 (1) (2013) 149–165.
[16] J.-M. Ma, W.-X. Zhang, Axiomatic characterizations of dual concept lattices, Int. J. Approx. Reason. 54 (5) (2013) 690–697.
[17] J. Medina, Towards multi-adjoint property-oriented concept lattices, in: Lecture Notes in Artificial Intelligence, vol. 6401, 2010, pp. 159–166.
[18] J. Medina, Multi-adjoint property-oriented and object-oriented concept lattices, Inf. Sci. 190 (2012) 95–106.
[19] J. Medina, Relating attribute reduction in formal, object-oriented and property-oriented concept lattices, Comput. Math. Appl. 64 (6) (2012) 1992–2002.
[20] J. Medina, M. Ojeda-Aciego, J. Ruiz-Calviño, Formal concept analysis via multi-adjoint concept lattices, Fuzzy Sets Syst. 160 (2) (2009) 130–144.
[21] J. Medina, M. Ojeda-Aciego, A. Valverde, P. Vojtáš, Towards biresiduated multi-adjoint logic programming, in: Lecture Notes in Artificial Intelligence, vol. 3040, 2004, pp. 608–617.
[22] M. Liu, M. Shao, W. Zhang, C. Wu, Reduction method for concept lattices based on rough set theory and its application, Comput. Math. Appl. 53 (2007) 1390–1410.
[23] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (1982) 341–356.
[24] A.M. Radzikowska, E.E. Kerre, A comparative study of fuzzy rough sets, Fuzzy Sets Syst. 126 (2) (2002) 137–155.
[25] L. Shen, D. Zhang, The concept lattice functors, Int. J. Approx. Reason. 54 (1) (2013) 166–183.
[26] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R.S. Owiński (Ed.), Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, 1992, pp. 331–362.
[27] X. Wang, W. Zhang, Relations of attribute reduction between object and property oriented concept lattices, Knowl.-Based Syst. 21 (2008) 398–403.
[28] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: I. Rival (Ed.), Ordered Sets, Reidel, 1982, pp. 445–470.
[29] Y.Y. Yao, A comparative study of formal concept analysis and rough set theory in data analysis, in: Lecture Notes in Artificial Intelligence, vol. 3066, 2004, pp. 59–68.