

Fuzzy Extensions of the Dominance-Based Rough Set Approach

Marko Palangetic^a, Chris Cornelis^a, Salvatore Greco^{b,c}, Roman Słowiński^{d,e}

^a*Department of Applied Mathematics, Computer Science and Statistics, Faculty of Science, Ghent University, Ghent, Belgium, {marko.palangetic, chris.cornelis}@ugent.be*

^b*Department of Economics and Business, University of Catania, Catania, Italy, salgreco@unict.it*

^c*Portsmouth Business School, Centre of Operations Research and Logistics (CORL), University of Portsmouth, Portsmouth, United Kingdom*

^d*Institute of Computing Science, Poznań University of Technology, Poznań, Poland, roman.slowinski@cs.put.poznan.pl*

^e*Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Abstract

In this paper, we first review existing fuzzy extensions of the dominance-based rough set approach (DRSA), and advance the theory considering additional properties. Moreover, we examine the application of Ordered Weighted Average (OWA) operators to fuzzy DRSA. OWA operators have shown a lot of potential in handling outliers and noisy data in decision tables, when they are combined with the indiscernibility-based rough set approach (IRSA). We examine theoretical properties of the proposed hybridisation of OWA operators with fuzzy DRSA. At the end, we experimentally compare the robustness of the standard fuzzy DRSA approach with the OWA one.

Keywords: Dominance-based rough set approach, Fuzzy set theory, Ordered Weighted Average.

1. Introduction

The main goal of rough set theory is to deal with inconsistencies in information. It is done in a way that one is distinguishing objects which are fully consistent with the available knowledge (lower approximation) from the objects which are only possibly consistent (upper approximation). The original proposal by Pawlak [21] defines these approximations based on an equivalence relation between objects, while the dominance-based rough set approach (DRSA) proposed by Greco, Matarazzo and Słowiński [13] uses a dominance relation between objects. The main practical application of DRSA is the possibility to induce monotonic rules from data structured using DRSA [23]. More precisely, the approach filters data which do not satisfy the Pareto principle. This principle says that if one object is not evaluated worse

than another object on all considered criteria, then it should not be assigned to a worse class than the other object. It is a main assumption of the monotonic classification problem, which is a special type of ordinal regression. This assumption is intuitive; if we have two companies where one of them has better financial parameters, then it should also have lower bankruptcy risk than the other. DRSA has found numerous applications in machine learning and operational research [1, 20, 29, 30].

On the other hand, fuzzy set theory [31] is used to model gradual information, where we are evaluating how much some statement is true on a scale from 0 to 1. While the crisp DRSA approach divides elements into consistent and inconsistent ones, the fuzzy DRSA approach allows to introduce a degree of consistency. Fuzzy set theory combined with the indiscernibility-based rough set approach (IRSA) has been applied extensively in machine learning, especially in domains like attribute selection [4], instance selection [18], imbalanced classification [22], multi-label classification [25], and so on. In this paper, we investigate how a similar hybridisation performs for ordinal classification problems, when we replace indiscernibility with a dominance relation in rough approximations.

Additionally, we examine the combination of the well-known Ordered Weighted Average (OWA) aggregation operator with fuzzy DRSA. OWA has been shown to improve IRSA in handling outliers and noisy data [5, 22, 24, 26, 27]. OWA makes approximations (and thus also machine learning algorithms that use them) more robust to small changes in the data. This goes at the expense of some desirable properties. However, for IRSA at least, it was shown that the OWA extension provides the best trade-off between theoretical properties and experimental performance among noise tolerant models [7]. In this paper, we evaluate whether a similar performance may be achieved with fuzzy DRSA.

The structure of the paper is as follows. Section 2 recalls preliminaries together with some additional properties of DRSA. In Section 3, we consider various possibilities of DRSA fuzzification, while in Section 4 we present the integration of OWA operators with fuzzy DRSA. In Section 5 we provide the experimental comparison of robustness of the standard fuzzy DRSA with the OWA-based fuzzy DRSA, while Section 6 is reserved for the conclusion and a discussion of future work.

2. Preliminaries

2.1. Fuzzy logic connectives

A negator $N : [0, 1] \rightarrow [0, 1]$ is a unary operator which is non-increasing and satisfies $N(0) = 1$, $N(1) = 0$. An involutive negator is one for which holds that $\forall x \in [0, 1]$, $N(N(x)) = x$. By N_s we denote the standard negator $N_s(x) = 1 - x$.

A t -norm $T : [0, 1]^2 \rightarrow [0, 1]$ is a binary operator which is commutative, associative, non-decreasing in both parameters and for which holds that $\forall x \in [0, 1]$, $T(x, 1) = x$.

A t -conorm $S : [0, 1]^2 \rightarrow [0, 1]$ is a binary operator which is commutative, associative, non-decreasing in both parameters and for which holds that $\forall x \in [0, 1]$, $S(x, 0) = x$. For a

given involutive negator N and a t -norm T , we say that a t -conorm S is the N -dual of T if it holds that $S(x, y) = N(T(N(x), N(y)))$. In this case, the triplet (T, N, S) is called a de-Morgan triplet.

An implicator $I : [0, 1]^2 \rightarrow [0, 1]$ is a binary operator which is non-increasing in the first coordinate, non-decreasing in the second one and for which holds that $I(1, 0) = 0, I(0, 0) = I(0, 1) = I(1, 1) = 1$.

We call I an S -implicator if there exists a negator N and a t -conorm S such that $I(x, y) = S(N(x), y)$. I is an R -implicator if there exists a t -norm T such that $I(x, y) = \sup\{\lambda \in [0, 1] : T(x, \lambda) \leq y\}$.

Let T be a t -norm, and I its R -implicator, we say that the residuation property holds if $T(x, y) \leq z \Leftrightarrow x \leq I(y, z)$. We have the following proposition.

Proposition 2.1. [19] The residuation property holds if and only if T is left-continuous.

We say that N is the negator induced by an implicator I if $N(x) = I(x, 0)$.

If T is a left-continuous t -norm, I its R -implicator, and N the negator induced by I , we say that T is an IMTL t -norm if N is involutive [10]. The abbreviation IMTL stands for Involutive Monoidal t -norm based Logic.

2.2. Dominance-based rough set approach (DRSA)

We first recall the general version of DRSA. Assume that we are given a set of objects U on which we have defined a preorder relation \succsim , i.e., a binary relation on U that satisfies

- reflexive: it holds that $u \succsim u$,
- transitive: for any three elements $u, v, w \in U$ holds $u \succsim v \wedge v \succeq w \Rightarrow u \succsim w$,

and a total preorder relation \succeq which is a preorder relation that additionally satisfies

- total: for any two objects $u, v \in U$ it holds that $u \succeq v$ or $v \succeq u$,

We say that $u, v \in U$ are indifferent with respect to \succeq , denoted by $u \equiv v$, if $u \succeq v$ and $v \succeq u$, while u is strictly preferred to v , denoted by $u \succ v$, if $u \succeq v$ and $v \not\succeq u$. It is easy to conclude that \equiv is an equivalence relation and it splits our set of objects U into equivalence classes, among which we have a strict preference. We denote those equivalence classes as $Cl_t, t \in \{1, \dots, k\}$ where for all $u, v \in U, u \succ v$ if $u \in Cl_t, v \in Cl_s$ and $t > s$. Index t of equivalence class Cl_t is also called a decision label.

The sets which will be approximated are upward and downward unions of classes defined respectively as

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s, \quad Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s, \quad t = 1, \dots, k.$$

$u \in Cl_t^{\geq}$ means that “ u belongs at least to Cl_t ”, while $u \in Cl_t^{\leq}$ means that “ u belongs at most to Cl_t ”. We recall some basic properties of downward and upward unions:

- $Cl_1^{\geq} = Cl_n^{\leq} = U$,
- $Cl_n^{\geq} = Cl_n$ and $Cl_1^{\leq} = Cl_1$,
- for $t = 2, \dots, k$ holds $Cl_{t-1}^{\geq} = U - Cl_t^{\leq}$ and $Cl_{t-1}^{\leq} = U - Cl_t^{\geq}$.

For each $u \in U$ we define the following sets:

- a set $D^+(u)$ of objects dominating u , called dominating set,
 $D^+(u) = \{v \in U : v \succsim u\}$,
- a set $D^-(u)$ of objects dominated by u , called dominated set,
 $D^-(u) = \{v \in U : u \succsim v\}$.

The lower approximation $\underline{\text{apr}}_{\succsim}(Cl_t^{\geq})$ of Cl_t^{\geq} and the upper approximation $\overline{\text{apr}}_{\succsim}(Cl_t^{\geq})$ of Cl_t^{\geq} are defined as

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &= \{u \in U : D^+(u) \subseteq Cl_t^{\geq}\} \\ \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &= \{u \in U : D^-(u) \cap Cl_t^{\geq} \neq \emptyset\}.\end{aligned}$$

Analogously, we can define the lower and upper approximation of Cl_t^{\leq} as

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &= \{u \in U : D^-(u) \subseteq Cl_t^{\leq}\} \\ \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &= \{u \in U : D^+(u) \cap Cl_t^{\leq} \neq \emptyset\}.\end{aligned}$$

Remark 2.1. (Multiple criteria context) In the majority of applications, the set of objects U is described by a decision table. A decision table is a 4-tuple $\langle U, Q \cup \{d\}, V, f \rangle$ where $U = \{u_1, \dots, u_n\}$ is a finite set of objects or alternatives, $Q = \{q_1, \dots, q_m\}$ is a finite set of condition criteria, d is a decision criterion, $V = \cup_{q \in Q \cup \{d\}} V_q$, where V_q is a domain of criterion $q \in Q \cup \{d\}$ and $f : U \times Q \cup \{d\} \rightarrow V$ is an information function such that $f(u, q) \in V_q$ for each $u \in U$ and $q \in Q$. On each criterion $q \in Q \cup \{d\}$ we induce a total preorder relation based on the information function as $u \succeq_q v \Leftrightarrow f(u, q) \geq f(v, q)$. Let $P \subseteq Q$. The dominance relation \succsim_P on universe U is defined as follows: $u \succsim_P v$ if $u \succeq_q v, \forall q \in P$ and such relation is only reflexive and transitive. Based on the dominance relation \succsim_P and a total preorder \succeq_d we may define lower and upper approximations as introduced above.

We list the main properties of the lower and upper approximations [14]:

- **(inclusion):** $\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) \subseteq Cl_t^{\geq} \subseteq \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}), \quad \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) \subseteq Cl_t^{\leq} \subseteq \overline{\text{apr}}_{\succsim}(Cl_t^{\leq})$.

- (duality)

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &= U - \overline{\text{apr}}_{\succsim}(Cl_{t-1}^{\leq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &= U - \underline{\text{apr}}_{\succsim}(Cl_{t-1}^{\leq}), \\ \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &= U - \overline{\text{apr}}_{\succsim}(Cl_{t+1}^{\geq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &= U - \underline{\text{apr}}_{\succsim}(Cl_{t+1}^{\geq}).\end{aligned}$$

- (relation monotonicity) Assume we have another dominance relation $\succsim^* \subseteq \succsim$. Then we have that

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\subseteq \underline{\text{apr}}_{\succsim^*}(Cl_t^{\geq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\supseteq \overline{\text{apr}}_{\succsim^*}(Cl_t^{\geq}), \\ \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\subseteq \underline{\text{apr}}_{\succsim^*}(Cl_t^{\leq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\supseteq \overline{\text{apr}}_{\succsim^*}(Cl_t^{\leq}).\end{aligned}$$

Additionally, we also have the properties of exact approximation, decision and object monotonicity, idempotence and interaction between lower and upper approximation:

Proposition 2.2. (exact approximation)

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) = Cl_t^{\geq} &\Leftrightarrow Cl_t^{\geq} = \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}), \\ \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) = Cl_t^{\leq} &\Leftrightarrow Cl_t^{\leq} = \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}).\end{aligned}$$

Proof. We will prove the proposition for the upward unions. Analogously, it holds for the downward unions. We have the following sequence of equivalences:

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) \supseteq Cl_t^{\geq} &\Leftrightarrow (\forall u \in Cl_t^{\geq}) (D^+(u) \subseteq Cl_t^{\geq}) \\ &\Leftrightarrow (\forall u, v \in U) (u \in Cl_t^{\geq} \wedge v \succsim u \Rightarrow v \in Cl_t^{\geq}) \\ &\Leftrightarrow (\forall u, v \in U) (v \notin Cl_t^{\geq} \wedge v \succsim u \Rightarrow u \notin Cl_t^{\geq}) \\ &\Leftrightarrow (\forall u, v \in U) (u \notin Cl_t^{\geq} \wedge u \succsim v \Rightarrow v \notin Cl_t^{\geq}) \\ &\Leftrightarrow (\forall u \in U - Cl_t^{\geq}) (D^-(u) \subseteq U - Cl_t^{\geq}) \\ &\Leftrightarrow U - Cl_t^{\geq} \subseteq \underline{\text{apr}}_{\succsim}(U - Cl_t^{\geq}) \Leftrightarrow U - Cl_t^{\geq} \subseteq U - \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) \\ &\Leftrightarrow Cl_t^{\geq} \supseteq \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}).\end{aligned}$$

In the fourth equivalence we just changed the notation; v is replaced with u and u with v , while in the seventh equivalence we used the duality property. Using this with the inclusion property, we complete the proof. \square

Proposition 2.3. (decision monotonicity) For $t \geq s$ we have that

$$\begin{aligned}\underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\subseteq \underline{\text{apr}}_{\succsim}(Cl_s^{\geq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\subseteq \overline{\text{apr}}_{\succsim}(Cl_s^{\geq}), \\ \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\supseteq \underline{\text{apr}}_{\succsim}(Cl_s^{\leq}), & \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\supseteq \overline{\text{apr}}_{\succsim}(Cl_s^{\leq}).\end{aligned}$$

Proof. Obvious from the definition. \square

Proposition 2.4. (object monotonicity) Assume that $u \succsim v$. Then we have the following implications.

$$\begin{aligned} v \in \underline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\Rightarrow u \in \underline{\text{apr}}_{\succsim}(Cl_t^{\geq}), & v \in \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}) &\Rightarrow u \in \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}), \\ u \in \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\Rightarrow v \in \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}), & u \in \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}) &\Rightarrow v \in \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}). \end{aligned}$$

Proof. If $u \succsim v$ then we have that $D^+(u) \subseteq D^+(v)$ and $D^-(u) \supseteq D^-(v)$. Putting this into the definitions of the approximations we obtain the result. \square

We provide two more propositions for which the proofs can be found in [15].

Proposition 2.5. (idempotence) It holds that

$$\begin{aligned} \underline{\text{apr}}_{\succsim}(\underline{\text{apr}}_{\succsim}(Cl_t^{\geq})) &= \underline{\text{apr}}_{\succsim}(Cl_t^{\geq}), & \overline{\text{apr}}_{\succsim}(\overline{\text{apr}}_{\succsim}(Cl_t^{\geq})) &= \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}), \\ \underline{\text{apr}}_{\succsim}(\underline{\text{apr}}_{\succsim}(Cl_t^{\leq})) &= \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}), & \overline{\text{apr}}_{\succsim}(\overline{\text{apr}}_{\succsim}(Cl_t^{\leq})) &= \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}). \end{aligned}$$

Proposition 2.6. (interaction between lower and upper approximation) It holds that

$$\begin{aligned} \overline{\text{apr}}_{\succsim}(\underline{\text{apr}}_{\succsim}(Cl_t^{\geq})) &= \underline{\text{apr}}_{\succsim}(Cl_t^{\geq}), & \underline{\text{apr}}_{\succsim}(\overline{\text{apr}}_{\succsim}(Cl_t^{\geq})) &= \overline{\text{apr}}_{\succsim}(Cl_t^{\geq}), \\ \overline{\text{apr}}_{\succsim}(\underline{\text{apr}}_{\succsim}(Cl_t^{\leq})) &= \underline{\text{apr}}_{\succsim}(Cl_t^{\leq}), & \underline{\text{apr}}_{\succsim}(\overline{\text{apr}}_{\succsim}(Cl_t^{\leq})) &= \overline{\text{apr}}_{\succsim}(Cl_t^{\leq}). \end{aligned}$$

3. Fuzzy extension of DRSA

Here, we want to relax the statement that “ u is not worse than v ” adding some sort of grading. So, we would like to measure how much the previous statement is true on a scale from 0 to 1. We can interpret this as the credibility of a statement. First, we recall the approach from Greco et al. [11], [12]. Throughout this section we assume that we are given t -norm T , negator N , t -conorm S and implicator I . We define a fuzzy dominance relation as $D : U \times U \rightarrow [0, 1]$ which is reflexive: $D(u, u) = 1$ and T -transitive: $T(D(u, v), D(v, z)) \leq D(u, z)$ for a given t -norm T . We assume now that the class sets Cl_t are fuzzy sets with degrees of membership $Cl_t(u)$ for $u \in U$. The value $Cl_t(u)$ provides the credibility that element u belongs to class Cl_t . Greco et al. [12] proposed the concept of cumulative fuzzy upward and downward unions as:

$$Cl_t^{\geq}(u) = \begin{cases} 1, & \text{if } \exists s > t : Cl_s(u) > 0 \\ Cl_t(u) & \text{otherwise} \end{cases}, \quad Cl_t^{\leq}(u) = \begin{cases} 1, & \text{if } \exists s < t : Cl_s(u) > 0 \\ Cl_t(u) & \text{otherwise} \end{cases}, \quad (1)$$

while Du et al. [8] proposed them as fuzzy unions of the classes, i.e:

$$Cl_t^{\geq}(u) = \max_{s \geq t} Cl_s(u), \quad Cl_t^{\leq}(u) = \max_{s \leq t} Cl_s(u). \quad (2)$$

In both of these cases, it holds that for all $u \in U$, $Cl_t^{\geq}(u) \leq Cl_s^{\geq}(u)$ and $Cl_t^{\leq}(u) \geq Cl_s^{\leq}(u)$ if $t \geq s$ and this is the minimal requirement we would ask for any possible definition of $Cl_t^{\geq}(u)$ and $Cl_t^{\leq}(u)$. The corresponding membership degrees to such defined fuzzy sets represent the credibility of the statement: “ u is not worse (not better) than objects from class Cl_t ”. Now the question is, how can we define fuzzy lower approximations? If we rewrite the statement of the lower approximation of the upward union, we have: “ $u \in Cl_t^{\geq}$ belongs to the lower approximation if $\forall v \in U$, for which $v \succsim u$, it also holds that $v \in Cl_t^{\geq}$ ”. In a fuzzy manner, we have to define a credibility that element $u \in Cl_t^{\geq}$ from the previous statement belongs to the lower approximation. For that purpose, we need to fuzzify logical quantifiers \forall and \exists , where \forall appears in the statement for the lower approximation as above, while \exists appears in the statement for the upper approximation. We denote those fuzzy quantifiers as qua_{\forall} and qua_{\exists} . To fuzzify those quantifiers we have two proposals. The first one is from Greco et al. [12] where fuzzy logic connectives are used, i.e. $qua_{\forall} = T$, $qua_{\exists} = S$. This option is suitable when the set of objects U is finite as it is in the case of machine learning applications. The second option is proposed by Greco et al. [11] where $qua_{\forall} = \inf$, $qua_{\exists} = \sup$. This option is suitable for both cases, when U is finite or infinite and this definition goes in line with the original fuzzy rough approximation proposed by Dubois and Prade [9]. So, for $(qua_{\forall}, qua_{\exists}) \in \{(T, S), (\inf, \sup)\}$ we have the following definitions for fuzzy lower and upper approximations:

$$\begin{aligned} \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\geq})(u) &= \text{qua}_{\forall}(I(D(v, u), Cl_t^{\geq}(v)); v \in U), \\ \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\geq})(u) &= \text{qua}_{\exists}(T(D(u, v), Cl_t^{\geq}(v)); v \in U), \\ \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\leq})(u) &= \text{qua}_{\forall}(I(D(u, v), Cl_t^{\leq}(v)); v \in U), \\ \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\leq})(u) &= \text{qua}_{\exists}(T(D(v, u), Cl_t^{\leq}(v)); v \in U). \end{aligned}$$

3.1. Properties of the Fuzzy DRSA

We have properties analogous to those listed before, presented in [12] for $(qua_{\forall}, qua_{\exists}) = (T, S)$ and in [11] for $(qua_{\forall}, qua_{\exists}) = (\inf, \sup)$:

- **(inclusion)** $\forall u \in U$:

$$\begin{aligned} \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\geq})(u) &\leq Cl_t^{\geq}(u), & \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\geq})(u) &\geq Cl_t^{\geq}(u), \\ \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\leq})(u) &\geq Cl_t^{\leq}(u), & \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\leq})(u) &\geq Cl_t^{\leq}(u). \end{aligned}$$

- **(duality)** Let T, S, N be a de-Morgan triplet, N involutive negator for which holds that $\forall t, N(Cl_t^{\geq}(u)) = Cl_{t-1}^{\leq}(u)$ and let I be the S -implicator induced by S and N . Then we have:

$$\begin{aligned} N(\underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\geq})(u)) &= \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_{t-1}^{\leq})(u), & N(\underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\leq})(u)) &= \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_{t-1}^{\geq})(u), \\ N(\overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\geq})(u)) &= \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_{t-1}^{\leq})(u), & N(\overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\leq})(u)) &= \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_{t-1}^{\geq})(u). \end{aligned}$$

- **(relation monotonicity)** For two fuzzy dominance relations D and D^* such that $D^* \subseteq D$ i.e., $\forall u, v \in U, D^*(u, v) \leq D(u, v)$, we have that

$$\begin{aligned} \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\geq})(u) &\leq \underline{\text{apr}}_{D^*}^{\text{qua}_{\forall}, I}(Cl_t^{\geq})(u), & \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\geq})(u) &\geq \overline{\text{apr}}_{D^*}^{\text{qua}_{\exists}, T}(Cl_t^{\geq})(u), \\ \underline{\text{apr}}_D^{\text{qua}_{\forall}, I}(Cl_t^{\leq})(u) &\leq \underline{\text{apr}}_{D^*}^{\text{qua}_{\forall}, I}(Cl_t^{\leq})(u), & \overline{\text{apr}}_D^{\text{qua}_{\exists}, T}(Cl_t^{\leq})(u) &\geq \overline{\text{apr}}_{D^*}^{\text{qua}_{\exists}, T}(Cl_t^{\leq})(u). \end{aligned}$$

For $(\text{qua}_{\forall}, \text{qua}_{\exists}) = (\inf, \sup)$ we have the property of exact approximation, as we show below.

Proposition 3.1. (exact approximation) Let T be a left-continuous t -norm and let I be its R -implicator. Then we have that

$$\begin{aligned} (\forall u \in U)(\underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^{\geq})(u) = Cl_t^{\geq}(u)) &\Leftrightarrow (\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^{\geq})(u) = Cl_t^{\geq}(u)), \\ (\forall u \in U)(\underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^{\leq})(u) = Cl_t^{\leq}(u)) &\Leftrightarrow (\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^{\leq})(u) = Cl_t^{\leq}(u)). \end{aligned}$$

Proof. We provide the proof of the proposition for upward unions. An analogous proof holds for downward unions. We will prove the following:

$$(\forall u \in U)(\underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^{\geq})(u) \geq Cl_t^{\geq}(u)) \Leftrightarrow (\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^{\geq})(u) \leq Cl_t^{\geq}(u)).$$

The above equivalence, together with the inclusion property, provides the desired result. We have that

$$\begin{aligned} (\forall u \in U)(\underline{\text{apr}}_{D_N}^{\text{inf}, I}(Cl_t^{\geq})(u) \geq Cl_t^{\geq}(u)) &\Leftrightarrow (\forall u \in U)(\inf_{v \in U}(I(D_P(v, u), Cl_t^{\geq}(v))) \geq Cl_t^{\geq}(u)) \\ &\Leftrightarrow (\forall u \in U)(\forall v \in U)(I(D_P(v, u), Cl_t^{\geq}(v)) \geq Cl_t^{\geq}(u)) \\ &\Leftrightarrow (\forall u \in U)(\forall v \in U)(T(D_P(v, u), Cl_t^{\geq}(u)) \leq Cl_t^{\geq}(v)) \\ &\Leftrightarrow (\forall v \in U)(\sup_{u \in U} T(D_P(v, u), Cl_t^{\geq}(u)) \leq Cl_t^{\geq}(v)) \\ &\Leftrightarrow (\forall u \in U)(\sup_{v \in U} T(D_P(u, v), Cl_t^{\geq}(v)) \leq Cl_t^{\geq}(u)) \\ &\Leftrightarrow (\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^{\geq})(u) \leq Cl_t^{\geq}(u)). \end{aligned}$$

The third equivalence holds because of the residuation property. In the fifth one, we just change the notation where u is replaced with v and v with u . \square

In Appendix A.1, we provide a counterexample that the same property does not hold if $(\text{qua}_\forall, \text{qua}_\exists) = (T, S)$. In Appendix A.1, the applied implicator is both an S -implicator and an R -implicator. Because of that, we omit using $(\text{qua}_\forall, \text{qua}_\exists) = (T, S)$ and we continue with $(\text{qua}_\forall, \text{qua}_\exists) = (\text{inf}, \text{sup})$. We want to investigate under which conditions all properties listed above are satisfied. The properties which require additional assumptions on fuzzy logic connectives are duality and exact approximation. We construct counterexamples that the exact approximation property does not hold under the assumptions of the duality property and vice versa. In Appendix A.2, we have that under the assumptions of the duality property we do not necessarily have the exact approximation property while in Appendix A.3, we may see that R -implicators cannot be used for the duality property in general. So, we conclude that R -implicators used in the exact approximation property have to be S -implicators in the duality property to have both properties together. The conclusion below unifies those observations.

Proposition 3.2. Assume that T is an IMTL t -norm, I its R -implicator, N the negator induced by I , and S the N -dual of T . Assume also that $(\text{qua}_\forall, \text{qua}_\exists) = (\text{inf}, \text{sup})$. Then the four properties listed above hold.

We can also prove the following analogous properties to the classical dominance-based rough set approach.

Proposition 3.3. (decision monotonicity) Let cumulative unions be defined as in (1) or (2). For $t \geq s$ and for all $u \in U$, we have that

$$\begin{aligned} \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^\geq)(u) &\leq \underline{\text{apr}}_D^{\text{inf}, I}(Cl_s^\geq)(u), & \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^\geq)(u) &\leq \overline{\text{apr}}_D^{\text{sup}, T}(Cl_s^\geq)(u), \\ \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^\leq)(u) &\geq \underline{\text{apr}}_D^{\text{inf}, I}(Cl_s^\leq)(u), & \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^\leq)(u) &\geq \overline{\text{apr}}_D^{\text{sup}, T}(Cl_s^\leq)(u). \end{aligned}$$

Proof. Obvious from the definition. □

Before the next proposition, we provide a simple lemma for fuzzy logic connectives.

Lemma 3.1. Let T be a left-continuous t -norm and let I be its R -implicator. If for some values $x, y, z \in [0, 1]$, it holds that $T(x, y) \leq z$, then for all $k \in [0, 1]$ it holds that $T(x, I(z, k)) \leq I(y, k)$.

Proof. Using the residuation property we have the equivalence: $I(x, y) = I(x, y) \Leftrightarrow T(x, I(x, y)) \leq y$. Using this and the associativity of T , we have:

$$T(T(x, I(T(x, y), k)), y) = T(T(x, y), I(T(x, y), k)) \leq k.$$

We apply the residuation property and we get:

$$T(T(x, I(T(x, y), k)), y) \leq k \Leftrightarrow T(x, I(T(x, y), k)) \leq I(y, k).$$

Using monotonic properties of T and I , we have that $T(x, I(T(x, y), k)) \geq T(x, I(z, k))$. Putting this into the previous expression, we obtain the conclusion. □

Proposition 3.4. (object monotonicity) Let T be a left-continuous t -norm and let I be its R -implicator. We have that

- $D(u, v) \leq I(\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(v), \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(u)),$
- $D(u, v) \leq I(\overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\geq})(v), \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\geq})(u)),$
- $D(u, v) \geq I(\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\leq})(v), \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\leq})(u)),$
- $D(u, v) \geq I(\overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\leq})(v), \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\leq})(u)).$

Proof. We will prove the proposition for the upward unions. The proof for the downward unions can be obtained by analogy. We start with the lower approximation. We fix $w \in U$. From the T -transitivity property we have that $T(D(u, v), D(w, u)) \leq D(w, v)$. Using Lemma 3.1, for $x = D(u, v), y = D(w, u), z = D(w, v), k = Cl_t^{\geq}(w)$, we have that:

$$T(D(u, v), I(D(w, v), Cl_t^{\geq}(w))) \leq I(D(w, u), Cl_t^{\geq}(w)).$$

From this we may conclude that:

$$\inf_{w_1 \in U} T(D(u, v), I(D(w_1, v), Cl_t^{\geq}(w_1))) \leq \inf_{w_2 \in U} I(D(w_2, u), Cl_t^{\geq}(w_2)).$$

Since T is increasing, we have that

$$\inf_{w_1 \in U} T(D(u, v), I(D(w_1, v), Cl_t^{\geq}(w_1))) \geq T(D(u, v), \inf_{w_1 \in U} I(D(w_1, v), Cl_t^{\geq}(w_1))).$$

Taking this into the previous expression, we obtain:

$$T(D(u, v), \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(v)) \leq \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(u).$$

Using the residuation principle we obtain the result. For the upper approximation, we may conclude:

$$\begin{aligned} & T(D(u, v), D(v, w)) \leq D(u, w) \\ \Rightarrow & T(T(D(u, v), D(v, w)), Cl_t^{\geq}(w)) \leq T(D(u, w), Cl_t^{\geq}(w)) \\ \Rightarrow & T(D(u, v), T(D(v, w), Cl_t^{\geq}(w))) \leq T(D(u, w), Cl_t^{\geq}(w)) \\ \Rightarrow & \sup_{w_1 \in U} T(D(u, v), T(D(v, w_1), Cl_t^{\geq}(w_1))) \leq \sup_{w_2 \in U} T(D(u, w_2), Cl_t^{\geq}(w_2)) \\ \Rightarrow & T(D(u, v), \sup_{w_1 \in U} T(D(v, w_1), Cl_t^{\geq}(w_1))) \leq \sup_{w_2 \in U} T(D(u, w_2), Cl_t^{\geq}(w_2)) \\ \Rightarrow & T(D(u, v), \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\geq})(v)) \leq \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\geq})(u) \end{aligned}$$

The fourth implication holds because T is left-continuous. Applying the residuation principle to the last expression, we obtain the conclusion. \square

We provide two more properties as a corollary of the object monotonicity property.

Proposition 3.5. (idempotence) Let T be a left-continuous t -norm and let I be its R -implicator. It holds that

$$\begin{aligned}\underline{\text{apr}}_D^{\text{inf},I}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})) &= \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}), & \overline{\text{apr}}_D^{\text{sup},T}(\overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\geq})) &= \overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\geq}), \\ \underline{\text{apr}}_D^{\text{inf},I}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\leq})) &= \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\leq}), & \overline{\text{apr}}_D^{\text{sup},T}(\overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\leq})) &= \overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\leq}).\end{aligned}$$

Proof. We will prove the proposition for the first statement. The other proofs stand by analogy. By the inclusion property we have that

$$\forall u, \underline{\text{apr}}_D^{\text{inf},I}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}))(u) \leq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u)$$

On the other hand, we may apply the residuation principle to the object monotonicity property. We have that:

$$\begin{aligned}\forall u \in U, D(v, u) &\leq I(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) \\ \Leftrightarrow \forall u \in U, T(D(v, u), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u)) &\leq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v) \\ \Leftrightarrow \forall u \in U, \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u) &\leq I(D(v, u), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) \\ \Leftrightarrow \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}) &\leq \inf_{v \in U} I(D(v, u), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) \\ \Leftrightarrow \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}) &\leq \underline{\text{apr}}_D^{\text{inf},I}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})),\end{aligned}$$

which proves the equality. \square

Proposition 3.6. (interaction between lower and upper approximation) Let T be a left-continuous t -norm and let I be its R -implicator. It holds that

$$\begin{aligned}\overline{\text{apr}}_D^{\text{sup},T}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})) &= \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}), & \underline{\text{apr}}_D^{\text{inf},I}(\overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\geq})) &= \overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\geq}), \\ \overline{\text{apr}}_D^{\text{sup},T}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\leq})) &= \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\leq}), & \underline{\text{apr}}_D^{\text{inf},I}(\overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\leq})) &= \overline{\text{apr}}_D^{\text{sup},T}(Cl_t^{\leq}).\end{aligned}$$

Proof. From the inclusion property we have that

$$\forall u \in U, \overline{\text{apr}}_D^{\text{sup},T}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}))(u) \geq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u)$$

On the other hand, applying the residuation principle to the object monotonicity property, we have

$$\begin{aligned}\forall u \in U, D(v, u) &\leq I(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) \\ \Leftrightarrow \forall v \in U, T(D(u, v), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) &\leq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u) \\ \Leftrightarrow \sup_{v \in U} T(D(u, v), \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(v)) &\leq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u) \\ \Leftrightarrow \overline{\text{apr}}_D^{\text{sup},T}(\underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq}))(u) &\leq \underline{\text{apr}}_D^{\text{inf},I}(Cl_t^{\geq})(u),\end{aligned}$$

which proves the equality. \square

3.2. Construction of the Fuzzy Dominance Relation

In this section, we provide an example of how to construct a fuzzy dominance relation, using previous work on fuzzy rough set theory. Although various authors have worked on proving properties of the fuzzy DRSA, none of them constructed a concrete example of a fuzzy dominance relation. We know that for a crisp dominance relation \succsim we may induce an equivalence or indiscernibility relation \sim in the following way:

$$u \sim v \Leftrightarrow u \succsim v \wedge v \succsim u. \quad (3)$$

In IRSA, rough sets are defined using the above indiscernibility relation. In fuzzy IRSA, the indiscernibility relation is replaced with a fuzzy similarity relation R which is usually assumed to be reflexive, symmetric and T -transitive for some t-norm T . In the case, when our data are represented in the form of a decision table, the most used example of such relation is a so-called triangular similarity. For a particular criterion q , two variants of this relation are defined as:

$$R_q^\lambda(u, v) = \max\left(1 - \lambda \frac{|f(u, q) - f(v, q)|}{\text{range}(q)}, 0\right), \quad R_q^\lambda(u, v) = \max\left(1 - \lambda \frac{|f(u, q) - f(v, q)|}{\sigma_q}, 0\right)$$

where $\text{range}(q)$ is the range of criterion q , σ_q is the standard deviation of q , and by $\lambda > 0$ we denote the shrinking parameter.

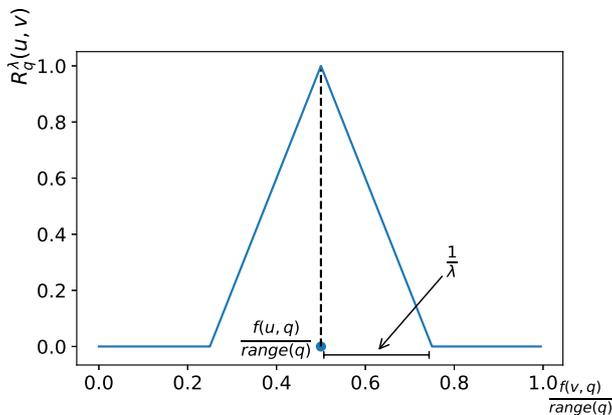


Figure 1: Examples of the similarity relation on criterion q for pair of objects (u, v) .

Figure 1 illustrates this similarity relation. The smaller λ is, the wider the triangle. To construct a similarity relation between objects taking into account subset of criteria $P \subseteq Q$, we usually use a minimum aggregation $R_P^\lambda(u, v) = \min_{q \in P} R_q^\lambda(u, v)$ or an average aggregation $R_P^\lambda(u, v) = \frac{1}{|P|} \sum_{q \in P} R_q^\lambda(u, v)$. Note that such relations are T -transitive when T is equal to the Łukasiewicz t-norm [6].

To define a fuzzy dominance relation, we want to follow a similar principle as in (3). For a fuzzy dominance relation D and a similarity relation R , the fuzzy version of (3) may be written as:

$$R(u, v) = T(D(u, v), D(v, u)). \quad (4)$$

So, we are looking for a dominance relation which satisfies (4) for the previous definitions of R , and which is reflexive and T -transitive, just like the crisp dominance relation. Assume we are given a decision table. For a particular criterion $q \in Q$, we propose two variants:

$$D_q^\lambda(u, v) = \max(\min(1 - \lambda \frac{f(v, q) - f(u, q)}{\text{range}(q)}, 1), 0) \quad (5)$$

$$D_q^\lambda(u, v) = \max(\min(1 - \lambda \frac{f(v, q) - f(u, q)}{\sigma_q}, 1), 0) \quad (6)$$

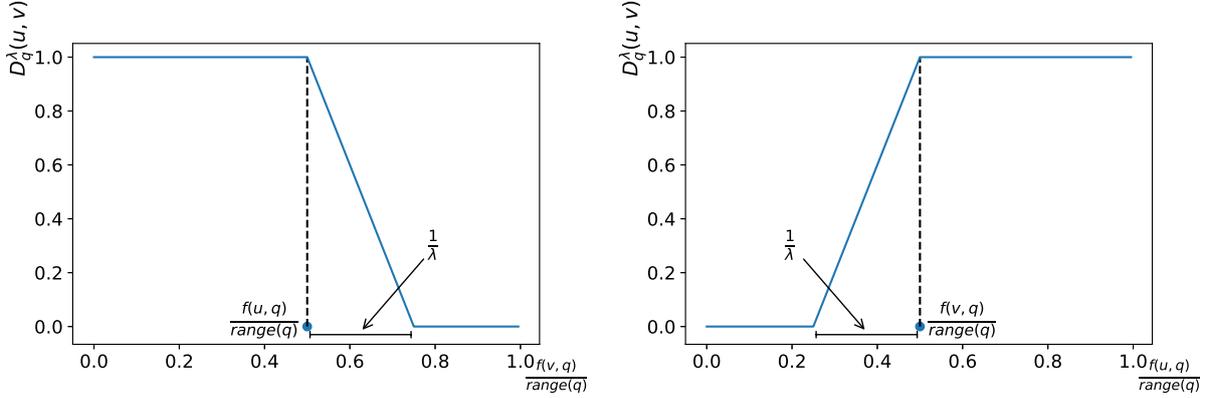


Figure 2: Examples of the dominance relation on criterion q for pair of objects (u, v) .

Examples of such defined dominance relations are given in Figure 2. The left image represents the dependence of $D_q^\lambda(u, v)$ on v when u is fixed, while the right one stands for the dependence of $D_q^\lambda(u, v)$ on u when v is fixed. It is easy to check that (4) holds in this case and that D_q^λ is reflexive. For T -transitivity we have the following result.

Proposition 3.7. Let T be the Łukasiewicz t -norm. Then the fuzzy relations defined by (5) and (6) are T -transitive, i.e, for elements $u, v, w \in U$ and a criterion q it holds that:

$$T(D_q^\lambda(u, v), D_q^\lambda(v, w)) \leq D_q^\lambda(u, w) \Leftrightarrow D_q^\lambda(u, v) + D_q^\lambda(v, w) - 1 \leq D_q^\lambda(u, w).$$

Proof. We prove the result for (5), while for (6) it is equivalent. We denote: $x = 1 - \lambda \frac{f(v, q) - f(u, q)}{\text{range}(q)}$, $y = 1 - \lambda \frac{f(w, q) - f(v, q)}{\text{range}(q)}$. Then, we have to prove that

$$\max(\min(x, 1), 0) + \max(\min(y, 1), 0) \leq \max(\min(x + y - 1, 1), 0) + 1.$$

We have three possible cases for x : $x < 0$, $0 \leq x \leq 1$, $x > 1$ and three analogous cases for y , which leads to nine possible cases of their combination. We can reduce that to six cases due to symmetry. Simple verification for each case gives us the desired result. \square

To construct a fuzzy dominance relation taking into account subset of criteria $P \subseteq Q$, we have two possibilities as before: to use a minimum or an average aggregation. If we use an average to aggregate the values, we cannot guarantee the T -transitivity in general. For a minimum aggregation we have the following result.

Proposition 3.8. Let $D_P^\lambda(u, v) = \min_{q \in P} D_q^\lambda(u, v)$. Then D_P^λ is T -transitive for any t -norm T , i.e. for elements $u, v, w \in U$ holds that

$$T(D_P^\lambda(u, v), D_P^\lambda(v, w)) \leq D_P^\lambda(u, w).$$

Proof. We have the following:

$$\begin{aligned} T(D_P^\lambda(u, v), D_P^\lambda(v, w)) &= T(\min_{q \in P} D_q^\lambda(u, v), \min_{r \in P} D_r^\lambda(w, u)) \\ &\leq \min_{q \in P} \min_{r \in P} T(D_q^\lambda(u, v), D_r^\lambda(w, u)) \\ &\leq \min_{q \in P} T(D_q^\lambda(u, v), D_q^\lambda(w, u)) \\ &\leq \min_{q \in P} D_q^\lambda(u, w) = D_P^\lambda(u, w). \end{aligned}$$

Here we used monotonicity of T and T -transitivity of D_q^λ . \square

4. Integration with OWA

In this section, we introduce the application of Ordered Weighted Average, or shortly OWA, aggregation operators in fuzzy DRSA. To evaluate if an object belongs to the lower approximation of the upward union, we check if its dominating set is contained in the upward union. In many practical approaches, we may have outliers: objects that do not follow the general distribution of the data and take some extreme values, e.g., an object with good values on all considered criteria, assigned to a worse class than many objects getting worse values on these criteria, or conversely, an object with bad values on all considered criteria assigned to a better class than many of the objects getting better values on these criteria. Because of such objects, many other objects are excluded from lower approximations of the unions of decision classes they typically belong to. Thus, the lower approximations may be poor or even empty. To avoid this, if there is some outlier, we want to reduce its significance in the calculation of the lower approximation. A lot of work has been done to handle such issues for the classical version of DRSA. Some well-known methods include Variable Precision DRSA [17] and Variable Consistency DRSA [16]. Here, we propose a new

approach suitable for fuzzy DRSA, and which is called the OWA approach. OWA operators already showed promising performance in IRSA, not only for decreasing an outlier's influence in general [7], but also in cases of imbalanced classification [22] and multi-instance learning [27]. OWA operators are used to replace the fuzzy quantifiers used for a final aggregation in the calculation of the lower and upper approximations. We recall the definition from [28]:

Definition 4.1. The OWA aggregation of set V of n real numbers with weight vector $W = (w_1, w_2, \dots, w_n)$, where $w_i \in [0, 1]$ and $\sum_{i=1}^n w_i = 1$, is given by

$$\text{OWA}_W(V) = \sum_{i=1}^n w_i v_{(i)},$$

where $v_{(i)}$ is the i -th largest element in the set V .

We have the monotonicity property of OWA:

Proposition 4.1. [28] Let V and V' be two sets of n real numbers such that for some permutation σ we have that $\forall i, V_{\sigma(i)} \geq V'_i$. If W is a vector of weights, we have that $\text{OWA}_W(V) \geq \text{OWA}_W(V')$.

It is worth of noticing that the condition from Proposition 4.1 is equivalent to saying that $\forall i, V_{(i)} \geq V'_{(i)}$.

Different weight vectors are used depending if they will be used to replace qua_\forall or qua_\exists . For that purpose, we use *andness* and *orness* measures, where *andness* evaluates how close an aggregation vector is to qua_\forall while *orness* does the same for qua_\exists . They are defined as:

$$\text{orness}(W) = \frac{1}{n-1} \sum_{i=1}^n (w_i(n-i)), \quad \text{andness}(W) = 1 - \text{orness}(W).$$

It is easy to see that fuzzy quantifiers inf and sup are special cases of OWA. For them we have the corresponding OWA vectors $W_{\text{inf}} = (0, \dots, 0, 1)$ and $W_{\text{sup}} = (1, 0, \dots, 0)$. It is easy to check that:

$$\text{andness}(W_{\text{inf}}) = 1, \quad \text{orness}(W_{\text{sup}}) = 1.$$

Assume now that we are given two weight vectors W_L , $\text{andness}(W_L) > 0.5$ and W_U , $\text{orness}(W_U) > 0.5$ of size $|U|$. We have the new definitions of lower and upper approximations:

$$\begin{aligned} \underline{\text{apr}}_D^{W_L, I}(Cl_t^\geq)(u) &= \text{OWA}_{W_L}(\{I(D(v, u), Cl_t^\geq(v)); v \in U\}), \\ \overline{\text{apr}}_D^{W_U, T}(Cl_t^\geq)(u) &= \text{OWA}_{W_U}(\{T(D(u, v), Cl_t^\geq(v)); v \in U\}), \\ \underline{\text{apr}}_D^{W_L, I}(Cl_t^\leq)(u) &= \text{OWA}_{W_L}(\{I(D(u, v), Cl_t^\leq(v)); v \in U\}), \\ \overline{\text{apr}}_D^{W_U, T}(Cl_t^\leq)(u) &= \text{OWA}_{W_U}(\{T(D(v, u), Cl_t^\leq(v)); v \in U\}). \end{aligned}$$

Here, we have more freedom to relax the definition of lower and upper approximations such that we can decrease the significance of possible outlier. A non-trivial OWA example is the case of additive weights obtained from the normalization of vector $(1, 2, \dots, n)$. We have that:

$$W_L^{add} = \left(\frac{2}{n(n+1)}, \frac{4}{n(n+1)}, \dots, \frac{2(n-1)}{n(n+1)}, \frac{2}{n+1} \right),$$

$$W_U^{add} = \left(\frac{2}{n+1}, \frac{2(n-1)}{n(n+1)}, \dots, \frac{4}{n(n+1)}, \frac{2}{n(n+1)} \right).$$

It is easy to check that $andness(W_L^{add}) = \frac{2}{3} > 0.5$ and $orness(W_U^{add}) = \frac{2}{3} > 0.5$. As we can see, the largest weights are multiplied with the possible outliers, but we are including also the values of the other, non-outlying objects, into our calculation which is not the case in the standard fuzzy DRSA. There we take either maximum or minimum of the values. **An example how to calculate the OWA-based approximations may be seen in Appendix A.4.**

We will now check if the same properties hold as before. For every $u \in U$ we first notice the following:

$$\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(u) \geq \underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(u) \geq \underline{\text{apr}}_D^{T, I}(Cl_t^{\geq})(u),$$

$$\overline{\text{apr}}_D^{W_U, T}(Cl_t^{\leq})(u) \leq \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\leq})(u) \leq \overline{\text{apr}}_D^{S, T}(Cl_t^{\leq})(u).$$

Let us now identify some other properties.

Proposition 4.2. (duality) Let $W_L, orness(W_L) > 0.5$ be a weight vector, T, S, N a de Morgan triplet with $N = N_s$ and let I be an S -implicator. Define vector W_U as the reverse vector of W_L , i.e. $(W_U)_i = (W_L)_{n-i+1}$ for $i = 1, \dots, n$. Also assume that we have $\forall t, N(Cl_t^{\geq}(u)) = Cl_{t-1}^{\leq}(u)$. Then, for $u \in U$ it holds that:

$$N(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(u)) = \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\leq})(u), \quad N(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\leq})(u)) = \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\geq})(u).$$

Proof. We will prove just the first expression while the second one will follow by analogy. We fix $u \in U$. Without loss of generality we assume that

$$I(D(u_1, u), Cl_t^{\geq}(u_1)) \leq \dots \leq I(D(u_n, u), Cl_t^{\geq}(u_n)).$$

Using the assumptions of the proposition, we find:

$$\begin{aligned} & S(N(D(u_1, u), Cl_t^{\geq}(u_1))) \leq \dots \leq S(N(D(u_n, u), Cl_t^{\geq}(u_n))) \\ \Leftrightarrow & N(T(D(u_1, u), N(Cl_t^{\geq}(u_1)))) \leq \dots \leq N(T(D(u_n, u), N(Cl_t^{\geq}(u_n)))) \\ \Leftrightarrow & 1 - T(D(u_1, u), 1 - Cl_t^{\geq}(u_1)) \leq \dots \leq 1 - T(D(u_n, u), 1 - Cl_t^{\geq}(u_n)) \\ \Leftrightarrow & T(D(u_1, u), Cl_{t-1}^{\leq}(u_1)) \geq \dots \geq T(D(u_n, u), Cl_{t-1}^{\leq}(u_n)). \end{aligned}$$

Therefore, we have that:

$$\begin{aligned}
N(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(u)) &= 1 - \sum_{i=1}^n (W_L)_i \cdot I(D(u_i, u), Cl_t^{\geq}(u_i)) \\
&= 1 - \sum_{i=1}^n (W_U)_{n-i+1} \cdot S(1 - D(u_i, u), Cl_t^{\geq}(u_i)) \\
&= 1 - \sum_{i=1}^n (W_U)_{n-i+1} \cdot (1 - T(D(u_i, u), 1 - Cl_t^{\geq}(u_i))) \\
&= \sum_{i=1}^n (W_U)_{n-i+1} \cdot T(D(u_i, u), Cl_{t-1}^{\leq}(u_i)) = \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\leq})(u).
\end{aligned}$$

□

Proposition 4.3. (relation monotonicity) For two fuzzy dominance relations D and D^* for which it holds that $D \subseteq D^*$ i.e., $\forall u, v \in U, D^*(u, v) \leq D(u, v)$, and for any OWA aggregating vectors W_L and W_U we have that

$$\begin{aligned}
\underline{\text{apr}}_{D^*}^{W_L, I}(Cl_t^{\geq})(u) &\leq \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(u), & \overline{\text{apr}}_{D^*}^{W_U, T}(Cl_t^{\geq})(u) &\geq \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\geq})(u), \\
\underline{\text{apr}}_{D^*}^{W_L, I}(Cl_t^{\leq})(u) &\leq \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\leq})(u), & \overline{\text{apr}}_{D^*}^{W_U, T}(Cl_t^{\leq})(u) &\geq \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\leq})(u).
\end{aligned}$$

Proof. We use the monotonicity of I and T . We have that

$$I(D^*(v, u), Cl_t^{\geq}(v)) \leq I(D(v, u), Cl_t^{\geq}(v)), \quad T(D^*(v, u), Cl_t^{\geq}(v)) \geq T(D(v, u), Cl_t^{\geq}(v)).$$

Using Proposition 4.1 and the previous inequalities, we complete the proof. □

Proposition 4.4. (decision monotonicity) Let cumulative unions be defined as in (1) or (2). For $t \geq s$ and for all $u \in U$, we have that

$$\begin{aligned}
\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(u) &\leq \underline{\text{apr}}_D^{W_L, I}(Cl_s^{\geq})(u), & \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\geq})(u) &\leq \overline{\text{apr}}_D^{W_U, T}(Cl_s^{\geq})(u), \\
\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\leq})(u) &\geq \underline{\text{apr}}_D^{W_L, I}(Cl_s^{\leq})(u), & \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\leq})(u) &\geq \overline{\text{apr}}_D^{W_U, T}(Cl_s^{\leq})(u).
\end{aligned}$$

Proof. From the definitions, we have that

$$\forall u \in U, t \geq s \Rightarrow Cl_t^{\geq}(u) \leq Cl_s^{\geq}(u) \wedge Cl_t^{\leq}(u) \geq Cl_s^{\leq}(u).$$

. If we fix $u \in U$, then for all $v \in U$ it holds:

$$\begin{aligned}
I(D(v, u), Cl_t^{\geq}(v)) &\leq I(D(v, u), Cl_s^{\geq}(v)), & T(D(v, u), Cl_t^{\geq}(v)) &\geq T(D(v, u), Cl_s^{\geq}(v)), \\
I(D(v, u), Cl_t^{\leq}(v)) &\leq I(D(v, u), Cl_s^{\leq}(v)), & T(D(v, u), Cl_t^{\leq}(v)) &\geq T(D(v, u), Cl_s^{\leq}(v)).
\end{aligned}$$

Using Proposition 4.1 and the previous inequalities, we complete the proof. □

The counterexample in Appendix A.5 shows that the inclusion property does not hold in general. However, we may provide a modification such that the inclusion property holds. We may take:

$$\begin{aligned}\underline{\text{apr}}_{D_P}(Cl_t^{\geq})(u) &= \min(Cl_t^{\geq}(u), \underline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\geq})(u)), \\ \overline{\text{apr}}_{D_P}(Cl_t^{\geq})(u) &= \max(Cl_t^{\geq}(u), \overline{\text{apr}}_{D_P}^{W_U, T}(Cl_t^{\geq})(u)).\end{aligned}$$

It is obvious that in this case we will have the inclusion property, however, it is not clear whether this extension is useful in practical applications. The other properties we listed for the standard fuzzy DRSA approach do not hold in general. A counterexample for the exact approximation property is provided in Appendix A.6 while for the other properties, counterexamples are provided in Appendix A.7.

5. Experimental Evaluation

5.1. Experimental setup

In this section, we will compare the robustness of the standard fuzzy DRSA and the OWA-based fuzzy DRSA. For this purpose, we collected six datasets described in [3] which are used for ordinal classification with monotonicity constraints. A description of these datasets is given in Table 1.

name	number of objects	number of condition criteria	number of decision classes
cpu	209	6	4
era	1000	4	9
esl	488	4	9
fame	1328	10	5
lev	1000	4	5
swd	1000	10	4

Table 1: Data description

In this experiment, we will consider only the lower approximations of both upward and downward unions, since due to the duality property, the upper approximation performance will be the same. We define the positive region as the fuzzy union of the lower approximations of the upward and downward unions, i.e.,

$$POS_{D,t}(u) = \max(\underline{\text{apr}}_D(Cl_t^{\geq})(u), \underline{\text{apr}}_D(Cl_{t-1}^{\leq})(u))$$

where $\underline{\text{apr}}$ may stand for either the standard fuzzy DRSA or the OWA-based fuzzy DRSA. We define the positive region for each value $t = 2, 3, \dots, k$ and we will compare positive region

membership values of the OWA-based fuzzy DRSA and the standard fuzzy DRSA. If a fuzzy DRSA model is robust, we expect that the positive region does not change drastically when small changes in the data occur. This should be the case when both condition and decision criteria are affected by noise. Therefore, we will consider the cases when the condition and decision criteria are affected by noise separately.

- For the condition criteria, we add Gaussian noise to the data. For a given standard deviation and for each criterion-value pair, we generate a random number from the normal distribution with zero mean and the given standard deviation. We add that number to the criterion-value pair. We do this for different standard deviations, which represent the level of noise in the criteria.
- In the decision criterion, we are dealing with ordinal classes. So for a given class in our data, we have three options: it can be increased by one level, decreased by one level or stay the same. Increasing and decreasing may happen with the same, fixed probability. This probability represents the level of noise in the decision criterion. Again, we consider different values in the experiment.

In the experiments we use the fuzzy dominance relation described in Subsection 3.2, where $\lambda = 1$ and where minimum is used as aggregation operator over all criteria. We use Łukasiewicz t -norm $T(x, y) = \max(x + y - 1, 0)$, which is also IMTL logic, and its R -implicator $I(x, y) = \min(1 - x + y, 1)$. The class sets are constructed such that $Cl_t(u) = 1$ if u belongs to class t , while $Cl_t(u) = 0$ otherwise. Cumulative upward and downward unions are then constructed according to (1). The experiments were performed in PYTHON programming language together with NUMPY computational library. The seed for the random number generator in NUMPY was set to 0.

5.2. Noise in the condition criteria

We first perform the experiment where we add the noise to the criteria. With \tilde{D} we denote the fuzzy dominance relation defined on data with Gaussian noise. We define Cumulative Mass Difference (CMD) as the value which is telling us how much noise affected the positive region of our data. For the standard fuzzy DRSA it is defined as:

$$CMD = \sum_{t=2}^k \frac{\sum_{u \in U} |POS_{D,t}(u) - POS_{\tilde{D},t}(u)|}{|U|},$$

while for the OWA-based fuzzy DRSA we have that

$$CMD^{OWA} = \sum_{t=2}^k \frac{\sum_{u \in U} |POS_{D,t}^{OWA}(u) - POS_{\tilde{D},t}^{OWA}(u)|}{|U|}.$$

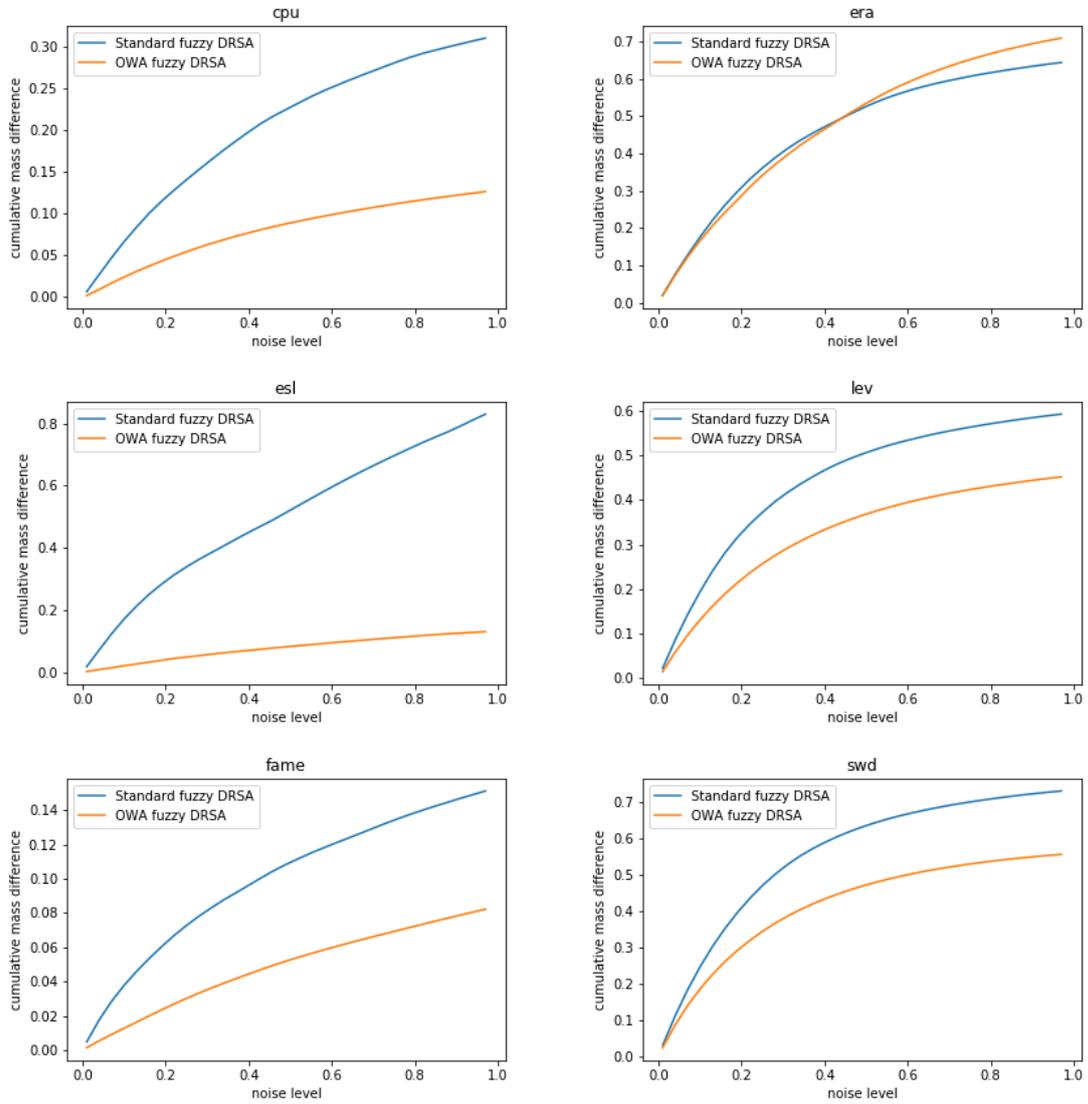


Figure 3: CMD with respect to the noise level on condition criteria

Since we want to show the robustness of OWA, we compare CMD and CMD^{OWA} as measures of robustness of the standard fuzzy DRSA and the OWA-based fuzzy DRSA, where the smaller value of CMD means more robustness. In this particular experiment we use the additive weights for the OWA operator.

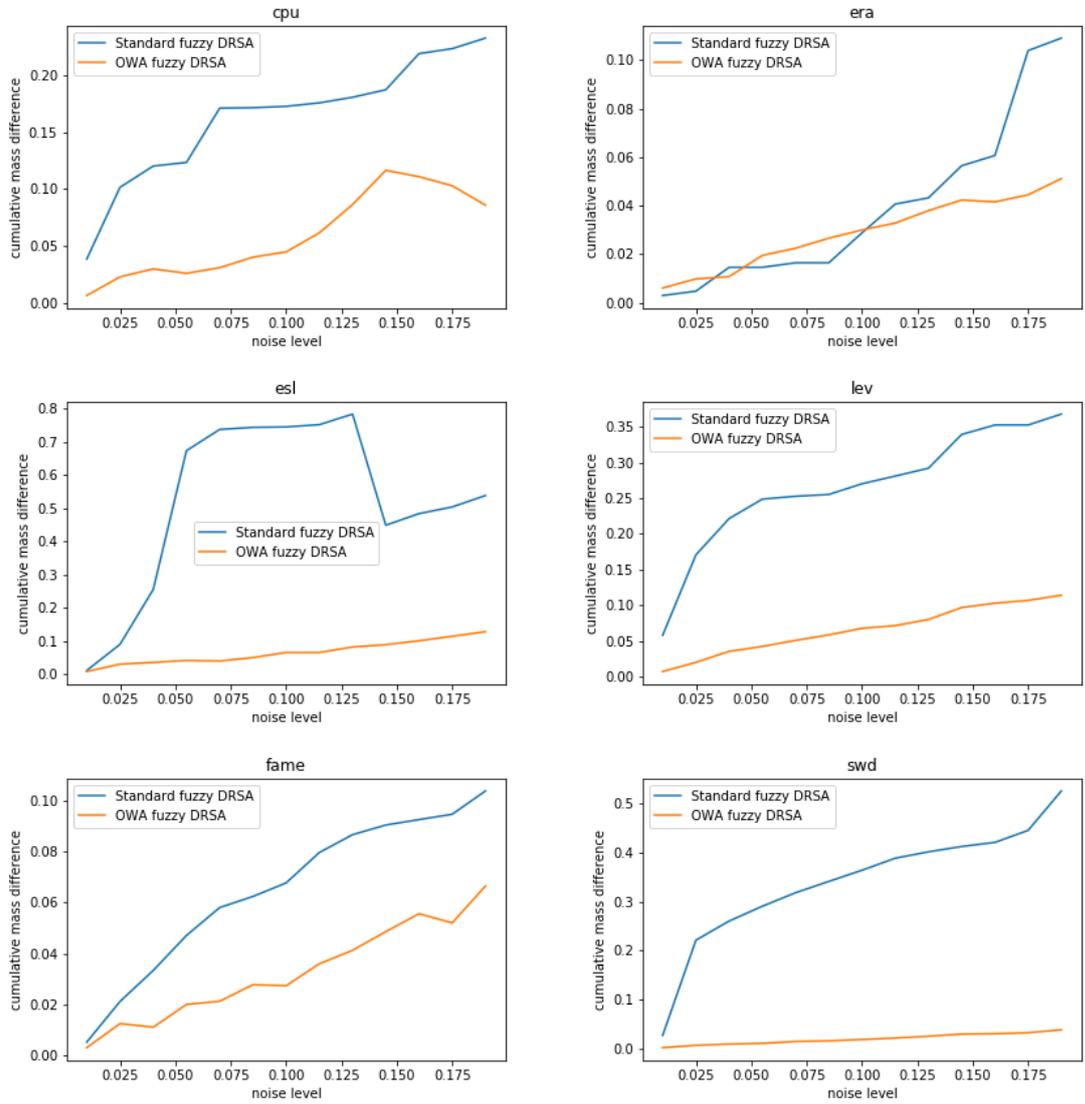


Figure 4: CMD with respect to the noise level on the decision criterion

In Figure 3, we show the dependence of CMD with respect to the noise level, i.e., with respect to the standard deviation value given to the Gaussian noise, for the different datasets described above. As we can see, 5 out of 6 images reveal that the OWA-based fuzzy DRSA is more robust than the standard fuzzy DRSA and that the robustness is increasing as the noise is being augmented. Only for the “era” dataset, both fuzzy DRSA are showing similar robustness with slightly better performance of the standard fuzzy DRSA.

5.3. Noise in the decision criterion

Next, we consider the experiment where noise is added to the decision criterion. By $Cl_{\tilde{t}}^{\geq}$ and $Cl_{\tilde{t}}^{\leq}$ we denote the respective upward and downward unions after noise is added while $POS_{D,\tilde{t}}(u)$ denotes the positive region with the noisy decision criterion. We define the CMD similarly as above, where $POS_{\tilde{D},t}(u)$ is replaced with $POS_{D,\tilde{t}}(u)$ and $POS_{\tilde{D},t}^{OWA}(u)$ is replaced with $POS_{D,\tilde{t}}^{OWA}(u)$. Again we compare CMD and CMD^{OWA} to show the robustness of OWA approach as we did before. Like before, the additive weights are used in the OWA operator. In Figure 4 we can clearly observe that the OWA-based method outperforms the standard one on 5 out of 6 datasets. As before, the "era" dataset is different where the standard DRSA is better for the small noise, while the OWA-based fuzzy DRSA outperforms for more noisy data.

5.4. Using different weights on "era" dataset

Here, we investigate why the OWA-based fuzzy DRSA failed to outperform the standard one on "era" dataset. Checking the dataset, we noticed that the positive region calculated with the standard fuzzy DRSA has many 0 membership degrees. That indicates a high presence of outliers in the "era" dataset, even without adding any artificial noise. One possibility is that the selection of the weights in this case is not appropriate.

As we stressed above, in the case of additive weights, the largest weights are multiplied with the possible outliers, which still gives some significance to the possible outliers and may affect the calculation. To avoid that, we perturb the weight vector W_L by defining a new weight vector W'_L in the following way. Let p be a percentage and let $n_1 = \lfloor pn \rfloor$. We have that:

$$(W'_L)_i = \begin{cases} (W_L)_{n-i+1} & \text{if } i = 1, \dots, n_1, \\ (W_L)_{i-n_1} & \text{if } i = n_1 + 1, \dots, n. \end{cases}$$

In this definition, we take a small percentage ($100p\%$) of values from the right side of the weight vector, flip them and add them to the left side of the weight vector. With this definition, we ensure that the small values from the end of vector W_L are now at the beginning of vector W'_L , so the possible outliers will not have as high significance as they had in W_L . We repeat the experiments for the "era" dataset, now with W'_L weights instead of W_L . We perform these experiments for $p = 0.05$ and $p = 0.1$.

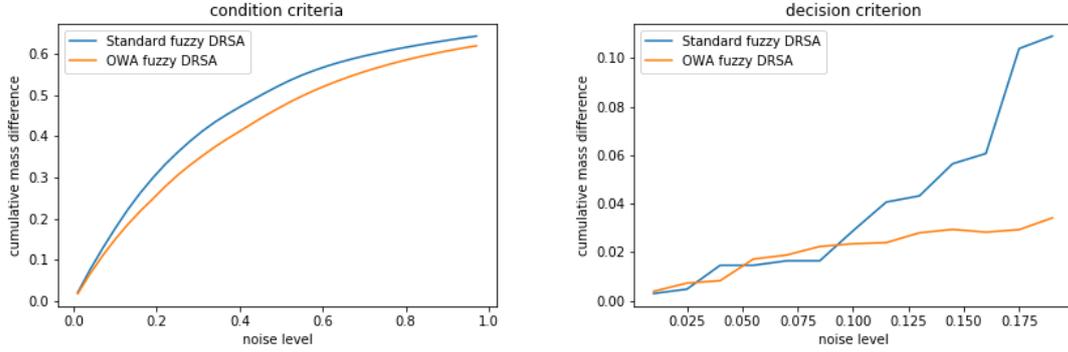


Figure 5: CMD with respect to the noise level for “era” dataset and $p = 0.05$.

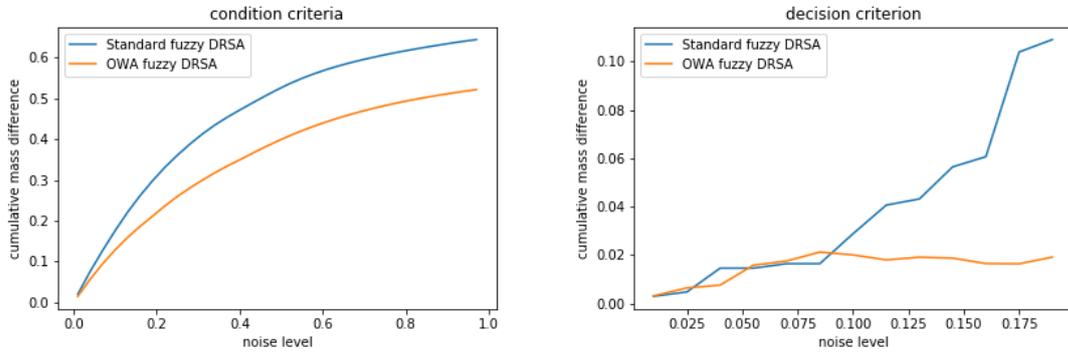


Figure 6: CMD with respect to the noise level for “era” dataset and $p = 0.1$.

Figures 5 and 6 reveal that giving smaller values to the potential outliers, improves the performance of the OWA-based fuzzy DRSA in both condition and decision case for “era” dataset. Moreover, we may see that moving a bigger portion of smaller values to the left side of the weight vector gives even better results, which again indicates the high presence of outliers in the “era” dataset.

6. Conclusion and future work

In this article we have presented the main results of the integration of fuzzy set theory and DRSA. We have also proposed some improvements using OWA operators to construct a more robust version of the fuzzy DRSA. We saw that some properties which hold for classical fuzzy DRSA also hold in the OWA version under specific assumptions. Our future work will have both theoretical and practical goals. From the theoretical point of view, we saw that the definition of upward and downward unions does not have too much influence on

the properties, it is only important that the upward unions are complements of downward unions based on a given negator. So, we will study what will happen if the unions are replaced with an arbitrary set. In that case we would like to see if similar properties to those which hold in IRSA, continue to hold here. Also, we want to verify these properties for other noise-tolerant extensions of DRSA like VP-DRSA and VC-DRSA. From the practical point of view, we want to evaluate how fuzzy DRSA may improve existing rule induction algorithms like DomLem and VC-DomLem. Rule-based inference has become an important tool in Machine Learning to interpret data or different ML models. DomLem and VC-DomLem performed well in extracting monotonic rules of the form ‘if ..., then ...’ from ordinal classification data [2]. It will be challenging to check how fuzzy set theory can improve them.

Acknowledgement

The first two authors would like to thank Odysseus project from Flanders Research Foundation (FWO) for funding their research.

References

- [1] Ali, A., Ali, M. I., Rehman, N., 2019. Soft dominance based rough sets with applications in information systems. *International Journal of Approximate Reasoning* 113, 171–195.
- [2] Błaszczyński, J., Słowiński, R., Szelaż, M., 2011. Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences* 181 (5), 987–1002.
- [3] Cano, J.-R., Gutiérrez, P. A., Krawczyk, B., Woźniak, M., García, S., 2019. Monotonic classification: An overview on algorithms, performance measures and data sets. *Neurocomputing* 341, 168–182.
- [4] Cornelis, C., Jensen, R., Hurtado, G., Ślęzak, D., 2010. Attribute selection with fuzzy decision reducts. *Information Sciences* 180 (2), 209–224.
- [5] Cornelis, C., Verbiest, N., Jensen, R., 2010. Ordered weighted average based fuzzy rough sets. In: *Proceedings of the 5th International Conference on Rough Sets and Knowledge Technology (RSKT 2010)*. Springer, pp. 78–85.
- [6] De Baets, B., Mesiar, R., 1997. Pseudo-metrics and t-equivalences. *Journal of Fuzzy Mathematics* 5, 471–481.
- [7] D’eer, L., Verbiest, N., Cornelis, C., Godo, L., 2015. A comprehensive study of implicator–conjunctive-based and noise-tolerant fuzzy rough sets: definitions, properties and robustness analysis. *Fuzzy Sets and Systems* 275, 1–38.

- [8] Du, W. S., Hu, B. Q., 2017. Dominance-based rough fuzzy set approach and its application to rule induction. *European Journal of Operational Research* 261 (2), 690–703.
- [9] Dubois, D., Prade, H., 1990. Rough fuzzy sets and fuzzy rough sets. *International Journal of General System* 17 (2-3), 191–209.
- [10] Esteva, F., Godo, L., 2001. Monoidal t-norm based logic: towards a logic for left-continuous t-norms. *Fuzzy sets and systems* 124 (3), 271–288.
- [11] Greco, S., Inuiguchi, M., Słowiński, R., 2002. Dominance-based rough set approach using possibility and necessity measures. In: *International Conference on Rough Sets and Current Trends in Computing*. Springer, pp. 85–92.
- [12] Greco, S., Matarazzo, B., Słowiński, R., 2000. Fuzzy extension of the rough set approach to multicriteria and multiattribute sorting. In: *Preferences and decisions under incomplete knowledge*. Springer, pp. 131–151.
- [13] Greco, S., Matarazzo, B., Słowiński, R., 2001. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research* 129 (1), 1–47.
- [14] Greco, S., Matarazzo, B., Słowiński, R., 2008. Fuzzy set extensions of the dominance-based rough set approach. In: *Fuzzy sets and their extensions: representation, aggregation and models*. Springer, pp. 239–261.
- [15] Greco, S., Matarazzo, B., Słowiński, R., 2012. The bipolar complemented de morgan brouwer-zadeh distributive lattice as an algebraic structure for the dominance-based rough set approach. *Fundamenta Informaticae* 115 (1), 25–56.
- [16] Greco, S., Matarazzo, B., Słowiński, R., Stefanowski, J., 2000. Variable consistency model of dominance-based rough sets approach. In: *International Conference on Rough Sets and Current Trends in Computing*. Springer, pp. 170–181.
- [17] Inuiguchi, M., Yoshioka, Y., 2006. Variable-precision dominance-based rough set approach. In: *International Conference on Rough Sets and Current Trends in Computing*. Springer, pp. 203–212.
- [18] Jensen, R., Cornelis, C., 2010. Fuzzy-rough instance selection. In: *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*. IEEE, pp. 1–7.
- [19] Klir, G., Yuan, B., 1995. *Fuzzy sets and fuzzy logic*. Vol. 4. Prentice hall New Jersey.
- [20] Liou, J. J., Tzeng, G.-H., 2010. A dominance-based rough set approach to customer behavior in the airline market. *Information Sciences* 180 (11), 2230–2238.

- [21] Pawlak, Z., 1982. Rough sets. *International Journal of Computer & Information Sciences* 11 (5), 341–356.
- [22] Ramentol, E., Vluymans, S., Verbiest, N., Caballero, Y., Bello, R., Cornelis, C., Herrera, F., 2015. Ifrowann: imbalanced fuzzy-rough ordered weighted average nearest neighbor classification. *IEEE Transactions on Fuzzy Systems* 23 (5), 1622–1637.
- [23] Słowiński, R., Greco, S., Matarazzo, B., 2015. Rough set methodology for decision aiding. In: *Springer Handbook of Computational Intelligence*. Springer, pp. 349–370.
- [24] Vluymans, S., 2018. Dealing with imbalanced and weakly labelled data in machine learning using fuzzy and rough set methods. Ph.D. thesis, Springer.
- [25] Vluymans, S., Cornelis, C., Herrera, F., Saeys, Y., 2018. Multi-label classification using a fuzzy rough neighborhood consensus. *Information Sciences* 433, 96–114.
- [26] Vluymans, S., Mac Parthalain, N., Cornelis, C., Saeys, Y., 2019. Weight selection strategies for ordered weighted average based fuzzy rough sets. *Information Sciences* 501, 155–171.
- [27] Vluymans, S., Sánchez Tarragó, D., Saeys, Y., Cornelis, C., Herrera, F., 2016. Fuzzy rough classifiers for class imbalanced multi-instance data. *Pattern Recognition* 53, 36–45.
- [28] Yager, R. R., 1988. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on systems, Man, and Cybernetics* 18 (1), 183–190.
- [29] Yang, X., Yang, J., Wu, C., Yu, D., 2008. Dominance-based rough set approach and knowledge reductions in incomplete ordered information system. *Information Sciences* 178 (4), 1219–1234.
- [30] Yang, X., Yu, D., Yang, J., Wei, L., 2009. Dominance-based rough set approach to incomplete interval-valued information system. *Data & Knowledge Engineering* 68 (11), 1331–1347.
- [31] Zadeh, L. A., 1965. Fuzzy sets. *Information and control* 8 (3), 338–353.

Appendix A. Counterexamples

Example Appendix A.1. Consider the Łukasiewicz t -norm, i.e., $T(x, y) = \max(x+y-1, 0)$. The corresponding R -implicator is $I(x, y) = \min(1 - x + y, 1)$. We induce N from I as $N(x) = 1 - x$ which is the standard negator, and we take S to be the N -dual of T , i.e.,

$S(x, y) = \min(x + y, 1)$. Let us assume that there are two objects a and b such that $Cl_t^\geq(a) = Cl_t^\geq(b) = 0.9$. Assume now that the exact approximation property holds, i.e.,

$$(\forall u \in U)(\underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^\geq)(u) = Cl_t^\geq(u)) \Leftrightarrow (\forall u \in U)(T_{v \in U}(I(D_P(v, u), Cl_t^\geq(v))) = Cl_t^\geq(u)).$$

We have that

$$\begin{aligned} T_{v \in U}(I(D_P(v, a), Cl_t^\geq(v))) &= T[I(D_P(a, a), Cl_t^\geq(a)), T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v)))] \\ &= T[Cl_t^\geq(a), T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v)))] \end{aligned}$$

The last expression is equal to $Cl_t^\geq(a)$ if

$$\begin{aligned} T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v))) = 1 &\Rightarrow (\forall v \neq a)(I(D_P(v, a), Cl_t^\geq(v)) = 1) \\ &\Rightarrow (\forall v \neq a)(D_P(v, a) \leq Cl_t^\geq(v)). \end{aligned}$$

Now assume $D_P(b, a) = 0.9$ which satisfies the condition $D_P(b, a) \leq Cl_t^\geq(b)$. We have that $T(D_P(b, a), Cl_t^\geq(a)) = 0.8$. Then we will have that

$$\begin{aligned} \overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^\leq)(b) &= S_{v \in U}(T(D_P(b, v), Cl_t^\geq(v))) \\ &\geq S[T(D_P(b, b), Cl_t^\geq(b)), T(D_P(b, a), Cl_t^\geq(a))] \\ &= S[Cl_t^\geq(b), T(D_P(b, a), Cl_t^\geq(a))] \\ &= S(0.9, 0.8) = 1 > 0.9 = Cl_t^\geq(b). \end{aligned}$$

So, we got that for a particular b it holds that, $\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^\leq)(b) \neq Cl_t^\geq(b)$ which is a counterexample to the exact approximation property.

Example Appendix A.2. We take de-Morgan triplet: $T(x, y) = \min(x, y)$, $N(x) = 1 - x$ and $S(x, y) = \max(x, y)$ with I as S -implicator, i.e., $I(x, y) = \max(1 - x, y)$. Assume that

$$\begin{aligned} &(\forall u \in U)(\underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^\geq)(u) \geq Cl_t^\geq(u)) \\ &\Leftrightarrow (\forall u \in U)(\inf_{v \in U}(I(D_P(v, u), Cl_t^\geq(v))) \geq Cl_t^\geq(u)) \\ &\Leftrightarrow (\forall u \in U)(\forall v \in U)(I(D_P(v, u), Cl_t^\geq(v)) \geq Cl_t^\geq(u)). \end{aligned}$$

Take two objects a and b such that $Cl_t^\geq(a) = 0.4$, $Cl_t^\geq(b) = 0.3$ and $D_P(b, a) = 0.5$. Then, we have that $I(D_P(b, a), Cl_t^\geq(b)) = I(0.5, 0.3) = 0.5 > 0.4 = Cl_t^\geq(a)$, so the condition above is satisfied. On the other side, we have that $T(D_P(b, a), Cl_t^\geq(a)) = T(0.5, 0.4) = 0.4 > 0.3 = Cl_t^\geq(b)$. Then, we have that

$$\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^\leq)(b) = \sup_{v \in U}(T(D_P(b, v), Cl_t^\geq(v))) \geq T(D_P(b, a), Cl_t^\geq(a)) > Cl_t^\geq(b),$$

which is a counterexample.

Example Appendix A.3. Take de-Morgan triplet $T(x, y) = \min(x, y)$, $S(x, y) = \max(x, y)$, $N = N_s$. Let I be R -implicator of T , i.e. $I(x, y) = 1$ if $x \leq y$ and $I(x, y) = y$ otherwise. It is obvious that in this case $(T, S) = (\inf, \sup)$. Assume that for some t and for unique object b we have that $Cl_{t-1}^{\leq}(b) = 0$ and $Cl_{t-1}^{\leq}(v) = 1$ for every $v \neq b$. Assume that for some u it holds that $D_P(b, u) < 1$. Then, we will have that $\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_{t-1}^{\leq})(u) = 0$ since the values of $I(D_P(u, v), Cl_{t-1}^{\leq}(v))$ are all ones with the one 0 value. On the other hand, we have that

$$\begin{aligned} N(\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^{\geq})(u)) &= N(\text{qua}_\exists(T(D_P(v, u), Cl_t^{\geq}(v)))) \\ &= \text{qua}_\forall(S(N(D_P(v, u)), N(Cl_t^{\geq}(v)))) \\ &= \text{qua}_\forall(S(N(D_P(v, u)), Cl_{t-1}^{\leq}(v))) \\ &= S(N(D_P(b, u)), Cl_{t-1}^{\leq}(b)) = N(D_P(b, u)) > 0. \end{aligned}$$

So, we got that for some u , $N(\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^{\geq})(u)) > \underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_{t-1}^{\leq})(u)$, which is a counterexample.

Example Appendix A.4. We provide an example how to calculate OWA-based fuzzy DRSA approximations, and we compare them with the standard fuzzy DRSA approximations. Let us consider the decision table shown in Table A.2.

obj.	cond1	cond2	decision
a	0.75	0.75	1
b	0.7	0.5	1
c	0.5	0.6	0
d	0.5	0.5	0

Table A.2: Example of decision table

We evaluate fuzzy dominance relations among objects of the decision table. We take a fuzzy dominance relation defined as in Subsection 3.2, where $\lambda = 1$ and the aggregation operator is minimum. We get the matrix shown in Table A.3.

	a	b	c	d
a	1	1	1	1
b	0.75	1	0.9	1
c	0.75	0.8	1	1
d	0.75	0.8	0.9	1

Table A.3: Pairwise evaluations of the fuzzy dominance relation

In the matrix, the value in cell (a, b) , for example, represents the value of $D(a, b)$. For the cumulative upward union, we take the decision vector $(1, 1, 0, 0)$ and its fuzzy set representation, where vector $(1, 1, 0, 0)$ corresponds to fuzzy membership values of objects (a, b, c, d) in fuzzy set Cl_t^{\geq} . For the cumulative downward union, we take the complementary set, i.e., set Cl_{t-1}^{\leq} with fuzzy membership values $(0, 0, 1, 1)$. We use additive weight vectors $W_L = (0.4, 0.3, 0.2, 0.1)$ and $W_U = (0.1, 0.2, 0.3, 0.4)$ for the OWA operators. To calculate the lower and upper approximations, we take Łukasiewicz t -norm $T(x, y) = \max(x + y - 1, 0)$, which is also IMTL logic, and its R -implicator $I(x, y) = \min(1 - x + y, 1)$. We provide the steps for calculation of $\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(a)$ and $\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a)$. For $v \in \{a, b, c, d\}$ we calculate the values $I(D(v, a), Cl_t^{\geq}(v))$. For $v \in \{a, b, c, d\}$ these values are $\{1, 1, 0.25, 0.25\}$. The standard lower approximation is then calculated by taking the smallest value, i.e., $\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(a) = 0.25$. For the OWA operation, we sort the values in ascending order: $\{0.25, 0.25, 1, 1\}$, and apply the weights W_L . We have that

$$\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a) = 0.4 \cdot 0.25 + 0.3 \cdot 0.25 + 0.2 \cdot 1 + 0.1 \cdot 1 = 0.375.$$

For the other objects we obtain: $\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(b) = 0.2$, $\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(c) = 0$, $\underline{\text{apr}}_D^{\text{inf}, I}(Cl_t^{\geq})(d) = 0$, $\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(b) = 0.33$, $\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(c) = 0.2$ and $\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(d) = 0.167$. For the other approximations we get:

$$\begin{aligned} \overline{\text{apr}}_D^{\text{sup}, T}(Cl_t^{\geq}) &= (1, 1, 0.8, 0.8), & \overline{\text{apr}}_D^{W_U, T}(Cl_t^{\geq}) &= (0.833, 0.75, 0.65, 0.65), \\ \underline{\text{apr}}_D^{\text{inf}, I}(Cl_{t-1}^{\leq}) &= (0, 0, 0.2, 0.2), & \underline{\text{apr}}_D^{W_L, I}(Cl_{t-1}^{\leq}) &= (0.167, 0.25, 0.35, 0.35), \\ \overline{\text{apr}}_D^{\text{sup}, T}(Cl_{t-1}^{\leq}) &= (0.75, 0.8, 1, 1), & \overline{\text{apr}}_D^{W_U, T}(Cl_{t-1}^{\leq}) &= (0.625, 0.667, 0.8, 0.833). \end{aligned}$$

Example Appendix A.5. Assume that Cl_t is crisp for every t , which means $Cl_t(u) = 1$ if $u \in Cl_t$ and $Cl_t(u) = 0$ otherwise. Let us compute $I(D_P(v, u), Cl_t^{\geq}(v))$. If we assume that I is an S -implicator, we have that:

- if $v \in Cl_t^{\geq}$, then $I(D_P(v, u), Cl_t^{\geq}(v)) = 1$,
- if $v \in Cl_{t-1}^{\leq}$, then $I(D_P(v, u), Cl_t^{\geq}(v)) = 1 - D_P(v, u)$.

So the values used for OWA aggregation are either $1 - D_P(v, u)$ or 1. Assume that $u \notin Cl_t^{\geq} \Rightarrow Cl_t^{\geq}(u) = 0$. Then, the lower approximation should be 0, but we can always construct weight vector for the OWA approach to obtain value different from 0 at the end.

Example Appendix A.6. Assume as above that Cl_t sets are crisp and assume that both W_L and W_U do not have zero weights. Then the evaluations of the implicators will be as above. For the t -norms we have that:

- if $v \in Cl_t^{\geq}$, then $T(D(u, v), Cl_t^{\geq}(v)) = D_P(u, v)$,

- if $v \in Cl_{t-1}^{\leq}$, then $T(D_P(u, v), Cl_t^{\geq}(v)) = 0$.

Let $u \in Cl_t^{\geq}$. Then we have that $\underline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\geq})(u) = Cl_t^{\geq}(u) = 1$ if and only if $\forall v \in Cl_{t-1}^{\leq}$, $1 - D_P(v, u) = 1 \Rightarrow D_P(v, u) = 0$. This holds since $\underline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\geq})(u)$ is a convex combination of the elements less or equal than 1 and it can be equal to 1 only if all elements are 1. So, it is possible to satisfy the condition. On the other hand, it is impossible to satisfy $\overline{\text{apr}}_{D_P}^{W_U, T}(Cl_t^{\geq})(u) = Cl_t^{\geq}(u) = 1$ since we have the convex combination where we have zero elements. Thus, we can conclude that the equivalence does not hold in general.

Example Appendix A.7. In this counter example, we use the same data as in Example Appendix A.4 as well as the same dominance relation, cumulative upward union and the OWA weights. Because of that, the pairwise comparison table will be the same as in Appendix A.4. To show the counterexamples for some properties, we take the left-continuous t -norm

$T(x, y) = \sqrt{\max(x^2 + y^2 - 1, 0)}$ and its R -implicator $I(x, y) = \sqrt{\min(1 - x^2 + y^2, 1)}$. We check objects a and d and we have that

$$\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a) = 0.763, \quad \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(d) = 0.3$$

We have that $I(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a), \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(d)) = 0.7126$, so

$$D(d, a) > I(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a), \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(d))$$

which is a counterexample for decision monotonicity property. Further we have that

$$\underline{\text{apr}}_D^{W_L, I}(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq}))(a) = 0.7751 \neq \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a)$$

which is a counterexample for the idempotence property while it also holds that

$$\overline{\text{apr}}_D^{W_U, T}(\underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq}))(a) = 0.6374 \neq \underline{\text{apr}}_D^{W_L, I}(Cl_t^{\geq})(a)$$

which is a counterexample for the relation between the lower and upper approximation property.