

# Extension of the Fuzzy Dominance-Based Rough Set Approach Using Ordered Weighted Average Operators

Marko Palangetic<sup>a</sup> and Chris Cornelis<sup>a</sup> and Salvatore Greco<sup>b,c</sup> and Roman Słowiński<sup>d,e</sup>

<sup>a</sup>Department of Applied Mathematics, Computer Science and Statistics, Faculty of Science, Ghent University, Krijgslaan 281 S9, 9000 Ghent, Belgium, {marko.palangetic, chris.cornelis}@ugent.be

<sup>b</sup>Department of Economics and Business,

University of Catania, Corso Italia, 55, Catania 95129, Italy, salgreco@unict.it

<sup>c</sup>Portsmouth Business School, Centre of Operations Research and Logistics (CORL), University of Portsmouth, Richmond Building, Portland Street, Portsmouth PO1 3DE, United Kingdom

<sup>d</sup> Institute of Computing Science,

Poznań University of Technology, 60-965 Poznań, Poland, roman.slowinski@cs.put.poznan.pl

<sup>e</sup>Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland

## Abstract

In the article we first review some known results on fuzzy versions of the dominance-based rough set approach (DRSA) where we expand the theory considering additional properties. Also, we apply Ordinal Weighted Average (OWA) operators in fuzzy DRSA. OWA operators have shown a lot of potential in handling outliers and noisy data in decision tables when it is combined with the indiscernibility-based rough set approach (IRSA). We examine theoretical properties of the proposed combination with fuzzy DRSA.

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

## 1 Introduction

The main purpose of the rough set theory is to deal with inconsistencies in information. It is done in a way that we distinguish objects which are fully consistent with the available knowledge (lower approximation) from the objects which are possibly consistent (upper approximation). The original rough set theory proposed by Pawlak [13] defines these approximations based on an equivalence relation between objects, while the dominance based rough set approach proposed by Greco, Matarazzo and Słowiński [6] is doing this based on a dominance relation between objects. The main practical application of DRSA is the possibility to extract monotonic rules. More precisely, the approach filters data which do not satisfy the Pareto principle. Pareto principle is saying that if one object is not evaluated worse than another object on all available criteria, then it should not be assigned to a worse class than the other object. This principle is a main assumption of the monotonic classification problem, which is a special type of the ordinal regression.

This assumption is intuitive; if we have two companies where one of them has better financial parameters, then that one should have lower bankruptcy risk than the other. DRSA found large applications in machine learning and operational research [12], [19] [20]. On the other hand, fuzzy set theory is used to model gradual information, where we are grading how much some statement is true on the scale from 0 to 1. Fuzzy set theory combined with the indiscernibility based rough set approach (IRSA) found a lot of applications in machine learning, especially in domains like attribute selection [1], instance selection [10], imbalanced classification [14], multi-label classification [16] and so on. So, we would like to see if a similar hybridisation when we replace indiscernibility with a dominance relation may give improvements in ordinal classification problems.

Additionally, we will investigate the combination of the well-known aggregation operator - Ordinal Weighted Average (OWA) with fuzzy DRSA. OWA has been shown to improve IRSA in handling outliers and noisy data [15]. 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 [2]. However, we would like to see if a similar performance may be achieved with fuzzy DRSA.

The structure of the paper is as follows. The next section recalls preliminaries together with some additional properties of DRSA. In the third section, we consider different possibilities of DRSA fuzzification, while in the fourth section we present OWA integration with fuzzy DRSA. The last section is intended for conclusion and future work.

## 2 Preliminaries

### 2.1 Fuzzy logic connectives

A negator  $N : [0, 1] \rightarrow [0, 1]$  is an unary operator which is non-increasing and  $N(0) = 1, N(1) = 0$ . An involutive negator is one for which holds that  $\forall x \in [0, 1], N(N(x)) = x$ . With  $N_s$  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  $N$ -dual of  $T$  if it holds that  $S(x, y) = N(T(N(x), N(y)))$ . In this case, triplet  $T, N, S$  is called de-Morgan triplet.

An implicator  $I : [0, 1]^2 \rightarrow [0, 1]$  is a binary operator which is non-increasing in its 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 say that  $I$  is an  $S$ -implicator if there exists a negator  $N$  and a  $t$ -conorm  $S$  such that  $I(x, y) = S(N(x), y)$ . We say that  $I$  is an  $R$ -implicator if there exist 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.** [11] Residuation property holds if and only if  $T$  is left-continuous.

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

Let  $T$  be a left continuous  $t$ -norm,  $I$  its  $R$ -implicator an  $N$  negator induced by  $I$ . If  $N$  is involutive then  $T$  is IMTL  $t$ -norm. Abbreviation IMTL stands for Involutive Monoidal  $t$ -norm based Logic.

### 2.2 Dominance based rough set approach

We first recall the classical version of DRSA. Assume that we are given a decision table represented by 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 the attribute  $q \in Q \cup \{d\}$  and  $f : U \times Q \cup \{d\} \rightarrow V$  is a total function such that  $f(u, q) \in V_q$  for each  $u \in U$  and  $q \in Q \cup \{d\}$ .  $f$  is sometimes called information function. In the applications of ordinal regression,  $V_d$  is finite and its elements are named classes. Assume

that  $|V_d| = k$ . Using  $V_d$  we can split the universe set  $U$  into  $k$  different sets, where every object  $u \in U$  belongs to one and only one class from  $V_d$ . Denote those sets of objects as  $Cl_t, t \in \{1, \dots, k\}$  where  $\cup_{t=1}^k Cl_t = U$  and  $Cl_t \cap Cl_s = \emptyset$  for  $s \neq t$ . On each criterion  $q \in Q$  we are assuming that there exists a preorder relation  $\succsim_q$ . Those relations are reflexive and transitive for all  $q \in Q$  i.e.

- it holds that  $u \succsim_q u$ ,
- for any three elements  $u, v, w \in U$  holds  $u \succsim_q v \wedge v \succsim_q w \Rightarrow u \succsim_q w$ .

We say that  $u, v \in U$  are indifferent with respect to  $q \in Q$  denoted by  $u \sim_q v$  if  $u \succsim_q v$  and  $v \succsim_q u$ , while  $u$  is strictly preferred to  $v$ , denoted by  $u \succ_q v$ , if  $u \succsim_q v$  and  $v \not\succsim_q u$ . With respect to decision criterion  $d$ , we assume that the elements in the same class  $Cl_t$  are indifferent, while for all  $u, v \in U, u \succ_d v$  if  $u \in Cl_t, v \in Cl_s$  and  $t > s$ . The sets which will be approximated are upward and downward unions of classes defined respectively as

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

$u \in Cl_t^{\geq}$  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 provide some basic properties of downward and upward unions:

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

Let  $P \subseteq Q$ . The dominance relation  $D_P$  on universe  $U$  is defined as follows:  $uD_Pv$  if  $u \succsim_q v, \forall q \in P$ . For each  $u \in U$  we define the following sets:

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

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

$$\underline{\text{apr}}_{D_P}(Cl_t^{\geq}) = \{u \in U : D_P^+(u) \subseteq Cl_t^{\geq}\}$$

$$\overline{\text{apr}}_{D_P}(Cl_t^{\geq}) = \{u \in U : D_P^-(u) \cap Cl_t^{\geq} \neq \emptyset\}.$$

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

$$\underline{\text{apr}}_{D_P}(Cl_t^{\leq}) = \{u \in U : D_P^-(u) \in Cl_t^{\leq}\}$$

$$\overline{\text{apr}}_{D_P}(Cl_t^{\leq}) = \{u \in U : D_P^+(u) \cap Cl_t^{\leq} \neq \emptyset\}.$$

We list the main properties of the lower and upper approximations [7]:

- **(inclusion):**  $\underline{\text{apr}}_{D_P}(Cl_t^{\geq}) \subseteq Cl_t^{\geq} \subseteq \overline{\text{apr}}_{D_P}(Cl_t^{\geq})$ ,  
 $\underline{\text{apr}}_{D_P}(Cl_t^{\leq}) \subseteq Cl_t^{\leq} \subseteq \overline{\text{apr}}_{D_P}(Cl_t^{\leq})$ .
- **(duality)**  $\underline{\text{apr}}_{D_P}(Cl_t^{\geq}) = U - \overline{\text{apr}}_{D_P}(Cl_{t-1}^{\leq})$ ,  
 $\overline{\text{apr}}_{D_P}(Cl_t^{\geq}) = U - \underline{\text{apr}}_{D_P}(Cl_{t-1}^{\leq})$ ,  
 $\underline{\text{apr}}_{D_P}(Cl_t^{\leq}) = U - \overline{\text{apr}}_{D_P}(Cl_{t+1}^{\geq})$ ,  
 $\overline{\text{apr}}_{D_P}(Cl_t^{\leq}) = U - \underline{\text{apr}}_{D_P}(Cl_{t+1}^{\geq})$ .
- **(criteria monotonicity)** If we have two sets of condition criteria  $M \subseteq P \subseteq Q$ , we have that  
 $\underline{\text{apr}}_{D_M}(Cl_t^{\geq}) \subseteq \underline{\text{apr}}_{D_P}(Cl_t^{\geq})$ ,  
 $\overline{\text{apr}}_{D_M}(Cl_t^{\geq}) \supseteq \overline{\text{apr}}_{D_P}(Cl_t^{\geq})$ ,  
 $\underline{\text{apr}}_{D_M}(Cl_t^{\leq}) \subseteq \underline{\text{apr}}_{D_P}(Cl_t^{\leq})$ ,  
 $\overline{\text{apr}}_{D_M}(Cl_t^{\leq}) \supseteq \overline{\text{apr}}_{D_P}(Cl_t^{\leq})$ .

**Proposition 2.2.** We have the property of exact approximation:

$$\underline{\text{apr}}_{D_P}(Cl_t^{\geq}) = Cl_t^{\geq} \Leftrightarrow Cl_t^{\geq} = \overline{\text{apr}}_{D_P}(Cl_t^{\geq}),$$

$$\underline{\text{apr}}_{D_P}(Cl_t^{\leq}) = Cl_t^{\leq} \Leftrightarrow Cl_t^{\leq} = \overline{\text{apr}}_{D_P}(Cl_t^{\leq}).$$

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

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

In the fourth equivalence we just changed the notation;  $v$  is replaced with  $u$  and  $u$  with  $v$ .  $\square$

### 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 scale from 0 to 1. We can name this as a credibility of a statement. In the beginning we recall the approach from Greco et al. [4], [5]. Throughout this section we assume that we are given  $t$ -norm  $T$ , negator  $N$ ,  $t$ -conorm  $S$  and implicator  $I$ . For a particular criterion  $q \in Q$ , we define a fuzzy ordering relation as  $R_q : U \times U \rightarrow [0, 1]$ . We require this fuzzy relation to be a preorder, i.e. to be reflexive:  $R_q(u, u) = 1$  and to be  $T$ -transitive:  $T(R_q(u, v), R_q(v, z)) \leq R_q(u, z)$ . Using such defined fuzzy ordering relation, we can define a fuzzy dominance relation with respect to a set of criteria  $P$  as:

$$D_P(u, v) = T_{q \in P}(R_q(u, v)).$$

We assume now that the class sets  $Cl_t$  are fuzzy sets with degrees of membership  $Cl_t(u)$  for  $u \in U$ . Value  $Cl_t(u)$  is providing us the credibility that an element  $u$  belongs to the class  $Cl_t$ . Greco et al. proposed the concept of cumulative fuzzy upward and downward unions as:

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

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

Corresponding values of such defined fuzzy sets represent credibility of the statement: ‘ $u$  is not worse (not better) than objects from the class  $Cl_t$ ’. Now the question is, how can we define our fuzzy lower approximations. If we rewrite the statement of lower approximation of the upward union we have: ‘ $u \in Cl_t^{\geq}$  belongs to the lower approximation of  $Cl_t^{\geq}$  if  $\forall v \in U$ , for which holds that  $v D_P u$  also holds  $v \in Cl_t^{\geq}$ ’. We have to define a credibility that element  $u \in Cl_t^{\geq}$  from the previous statement belongs to the lower approximation in a fuzzy manner,. For that purpose, we need to fuzzify logical quantifiers  $\forall$  and  $\exists$ , where  $\forall$  appears in a statement for the lower approximation as above, while  $\exists$  appears in a statement for the upper approximation. We denote those fuzzy quantifiers as  $\text{qua}_{\forall}$  and  $\text{qua}_{\exists}$ . To fuzzify those quantifiers we have two proposals. The first one is from Greco et al. [5] where we take fuzzy logic connectives, i.e.  $\text{qua}_{\forall} = T$ ,  $\text{qua}_{\exists} = S$ . This option is suitable when the set of objects  $U$  is finite as it is in our case and in a case of machine learning application. The second option is proposed by Greco et al. [4] where we take that  $\text{qua}_{\forall} = \inf$ ,  $\text{qua}_{\exists} = \sup$ . 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 of Dubois and Prade [3]. So, for  $(\text{qua}_{\forall}, \text{qua}_{\exists}) \in \{(T, S), (\inf, \sup)\}$  we have

the following definitions for fuzzy lower and upper approximations:

$$\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\geq)(u) = \text{qua}_\forall(I(D_P(v, u), Cl_t^\geq(v))); v \in U,$$

$$\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\geq)(u) = \text{qua}_\exists(T(D_P(u, v), Cl_t^\geq(v))); v \in U,$$

$$\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\leq)(u) = \text{qua}_\forall(I(D_P(u, v), Cl_t^\leq(v))); v \in U,$$

$$\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\leq)(u) = \text{qua}_\exists(T(D_P(v, u), Cl_t^\leq(v))); v \in U.$$

We have properties as before listed in [5] for  $(\text{qua}_\forall, \text{qua}_\exists) = (T, S)$  and in [4] for  $(\text{qua}_\forall, \text{qua}_\exists) = (\text{inf}, \text{sup})$ :

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

$$\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\geq)(u) \leq Cl_t^\geq(u),$$

$$\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\geq)(u) \geq Cl_t^\geq(u),$$

$$\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\leq)(u) \leq Cl_t^\leq(u),$$

$$\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\leq)(u) \geq Cl_t^\leq(u).$$

- **(duality)** Let  $T, S, N$  be 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:
 
$$N(\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\geq)(u)) = \overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_{t-1}^\leq)(u),$$

$$N(\underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\leq)(u)) = \overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_{t+1}^\geq)(u),$$

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

$$N(\overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\leq)(u)) = \underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_{t+1}^\geq)(u).$$

- **(criteria monotonicity)** If we have two subsets of criteria  $M \subseteq P \subseteq Q$ , then it follows that
 
$$\underline{\text{apr}}_{D_M}^{\text{qua}_\forall, I}(Cl_t^\geq)(u) \leq \underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\geq)(u),$$

$$\overline{\text{apr}}_{D_M}^{\text{qua}_\exists, T}(Cl_t^\geq)(u) \geq \overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\geq)(u),$$

$$\underline{\text{apr}}_{D_M}^{\text{qua}_\forall, I}(Cl_t^\leq)(u) \leq \underline{\text{apr}}_{D_P}^{\text{qua}_\forall, I}(Cl_t^\leq)(u),$$

$$\overline{\text{apr}}_{D_M}^{\text{qua}_\exists, T}(Cl_t^\leq)(u) \geq \overline{\text{apr}}_{D_P}^{\text{qua}_\exists, T}(Cl_t^\leq)(u).$$

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

**Proposition 3.1.** Let  $T$  be a left-continuous  $t$ -norm and let  $I$  be its  $R$ -implicator. Then we have that

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

$$(\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^\geq)(u) = Cl_t^\geq(u)),$$

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

$$(\forall u \in U)(\overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^\leq)(u) = Cl_t^\leq(u)).$$

*Proof.* We provide the proof of the proposition for upward unions. 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 inclusion property provides the desired result. We have that

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

$$(\forall u \in U)(\inf_{v \in U}(I(D_P(v, u), Cl_t^\geq(v))) \geq Cl_t^\geq(u)) \Leftrightarrow$$

$$(\forall u \in U)(\forall v \in U)(I(D_P(v, u), Cl_t^\geq(v)) \geq Cl_t^\geq(u)) \Leftrightarrow$$

$$(\forall u \in U)(\forall v \in U)(T(D_P(v, u), Cl_t^\geq(u)) \leq Cl_t^\geq(v)) \Leftrightarrow$$

$$(\forall v \in U)(\sup_{u \in U} T(D_P(v, u), Cl_t^\geq(u)) \leq Cl_t^\geq(v)) \Leftrightarrow$$

$$(\forall u \in U)(\sup_{v \in U} T(D_P(u, v), Cl_t^\geq(v)) \leq Cl_t^\geq(u)) \Leftrightarrow$$

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

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  $u$ .  $\square$

Here the question is, will the same property hold if we use that  $(\text{qua}_\forall, \text{qua}_\exists) = (T, S)$ . We provide the counterexample. Take 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 standard negator, and we take  $S$  to be  $N$ -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

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

$$(\forall u \in U)(T_{v \in U}(I(D_P(v, u), Cl_t^\geq(v))) = Cl_t^\geq(u)).$$

We have that

$$T_{v \in U}(I(D_P(v, a), Cl_t^\geq(v))) =$$

$$T[I(D_P(a, a), Cl_t^\geq(a)), T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v)))] =$$

$$T[Cl_t^\geq(a), T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v)))].$$

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

$$T_{v \neq a}(I(D_P(v, a), Cl_t^\geq(v))) = 1 \Rightarrow$$

$$(\forall v \neq a)(I(D_P(v, a), Cl_t^\geq(v)) = 1) \Rightarrow$$

$$(\forall v \neq a)(D_P(v, a) \leq Cl_t^\geq(v)).$$

Now take that  $D_P(b, a) = 0.9$  which satisfies the above condition. 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  $b$  holds that,  $\overline{\text{apr}}_{D_P}^{\text{sup},T}(Cl_t^{\leq})(b) \neq Cl_t^{\geq}(b)$  which is a counterexample. So if we want for the exact approximation property to hold, we need to keep using inf and sup as fuzzy quantifiers.

Now, we would like to investigate under which conditions we have that all properties are satisfied. Properties which require additional assumptions on fuzzy logic connectives are duality and exact approximation. First, we want to check if exact approximation property holds under the assumptions of the duality property and vice versa. More precisely, we will construct counterexamples that it does not hold. First let us solve the doubt if we may use  $(\text{qua}_{\forall}, \text{qua}_{\exists}) = (T, S)$  for the exact approximation property when we take  $S$ -implicator instead of  $R$ -implicator, which is the assumption for duality property. The answer is negative since provided Lukasiewicz implicator is both  $S$ -implicator and  $R$ -implicator so we have the same counterexample as before. Further on we take  $(\text{qua}_{\forall}, \text{qua}_{\exists}) = (\text{inf}, \text{sup})$ . Now, let us show that under the assumptions of duality property we do not have the exact approximation property. 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

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

which is a counterexample.

Now let us see that  $R$ -implicators cannot be used for duality property in general. 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  $(T, S) = (\text{inf}, \text{sup})$ . 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  $u$  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 side we have that  $N(\overline{\text{apr}}_{D_P}^{\text{qua}_{\exists},T}(Cl_t^{\geq})(u)) = N(\text{qua}_{\exists}(T(D_P(v, u), Cl_t^{\geq}(v)))) = \text{qua}_{\forall}(S(N(D_P(v, u)), N(Cl_t^{\geq}(v)))) = \text{qua}_{\forall}(S(N(D_P(v, u)), Cl_{t-1}^{\leq}(v))) = S(N(D_P(b, u)), Cl_{t-1}^{\leq}(b)) = N(D_P(b, u)) > 0$ . 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.

Now we are ready to form our final conclusion.

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

## 4 Integration with OWA

In this section, we introduce the application of aggregation approach called Ordinary Weighted Average or OWA in fuzzy DRSA. To investigate if an object belongs to the lower approximation of the upward union, we check if its  $P$ -dominating set is contained in the upward union. In many practical approaches, we may have outliers - the objects that do not follow a distribution of data. Because of such objects, we may have that all elements from  $P$ -dominating set are inside the upward union except maybe few. In that case those few elements cause that the object is excluded from the lower approximation, often causing lower approximations to be almost empty. We would like to avoid this. If there is some outlier, we want to reduce its significance to the calculation of the lower approximation or to just forget about it. A lot of work has been done to handle such issues for the classical version of DRSA. The well-known methods here are Variable Precision DRSA [9] and Variable Consistency DRSA [8]. Here, we are proposing a new approach suitable for fuzzy DRSA - OWA approach. OWA showed promising performance in IRSA, not only for decreasing of an outlier influence [2], but also in domains of imbalanced classification [14] and multi-instance learning [17]. OWA is used to replace the fuzzy quantifiers used for final aggregation when we are calculating lower and upper approximations. We recall the definition from [18]:

**Definition 4.1.** The OWA aggregation of a set of values  $V$  using weight vector  $W = (w_1, w_2, \dots, w_{|V|})$ ,

with  $w_i \in [0, 1]$  and  $\sum_{i=1}^{|V|} w_i = 1$ , is given by

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

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

We have monotonicity property of OWA:

**Proposition 4.1.** [18] Let  $V$  and  $V'$  be two sets of values such that for some permutation  $\sigma$  we have that  $\forall i, V_{\sigma(i)} \geq V'_i$ . If  $W$  is a vector of weights we have that  $\text{OWA}_W(V) \geq \text{OWA}_W(V')$ .

Different weight vectors are used depending if they will be used as  $\text{qua}_\vee$  or as  $\text{qua}_\exists$ . For that purpose, we define measures *andness* and *orness* where *andness* is telling us how much is our aggregation vector usable as  $\text{qua}_\vee$  while *orness* is measuring similarity with  $\text{qua}_\exists$ . They are defined as:

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

$$\text{andness}(W) = 1 - \text{orness}(W).$$

It is easy to see that fuzzy quantifiers *inf* and *sup* are special cases of OWA. For them we have the corresponding OWA vectors as  $W_{\text{inf}} = (0, \dots, 0, 1)$  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  $W_U, \text{orness}(W_U) > 0.5$ . We have the new definitions of lower and upper approximations:

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

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

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

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

It is easy to check that  $\text{andness}(W_L) > 0.5$  and  $\text{orness}(W_U) > 0.5$ . As we can see, we are giving the largest significance to the possible outlier, but we are including also the values of the other, non-outlying objects, into our calculation. This and some other examples are given in [15] and their performance in classification problems is experimentally tested on IRSA. We will now check if the same properties hold as before. For every  $u \in U$  we notice the following:

$$\begin{aligned} \underline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\geq})(u) & \geq \underline{\text{apr}}_{D_P}^{\text{inf}, I}(Cl_t^{\geq})(u) \\ & \geq \underline{\text{apr}}_{D_P}^{T, I}(Cl_t^{\geq})(u), \end{aligned}$$

$$\begin{aligned} \overline{\text{apr}}_{D_P}^{W_U, T}(Cl_t^{\leq})(u) & \leq \overline{\text{apr}}_{D_P}^{\text{sup}, T}(Cl_t^{\leq})(u) \\ & \leq \overline{\text{apr}}_{D_P}^{S, T}(Cl_t^{\leq})(u). \end{aligned}$$

Let us now identify some properties.

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

$$N(\underline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\geq})(u)) = \overline{\text{apr}}_{D_P}^{W_U, T}(Cl_{t-1}^{\leq})(u),$$

$$N(\overline{\text{apr}}_{D_P}^{W_L, I}(Cl_t^{\leq})(u)) = \underline{\text{apr}}_{D_P}^{W_U, T}(Cl_{t+1}^{\geq})(u),$$

*Proof.* We will prove just first expression while the second one will follow analogously. We fix  $u \in U$ . W.L.O.G. we assume that

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

Using assumptions of the proposition further we have:

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

We have that:

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

□

**Proposition 4.3.** For two sets of criteria  $M \subseteq P \subseteq Q$  and for any OWA aggregating vectors  $W_L$  and  $W_U$  we have that

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

*Proof.* For any  $t$ -norm  $T$  we have that  $T(x, y) \leq \min(x, y) \Leftrightarrow T(x, y) \leq x \wedge T(x, y) \leq y$ . For dominance relation we have that  $D_P(v, u) = T(D_M(v, u), D_{N-P}(v, u)) \leq D_M(v, u)$ . Because of the monotonicity of  $I$  and  $T$  we have that  $I(D_M(v, u), Cl_t^{\geq}(v)) \leq I(D_P(v, u), Cl_t^{\geq}(v))$  and  $T(D_M(v, u), Cl_t^{\geq}(v)) \geq T(D_P(v, u), Cl_t^{\geq}(v))$ . Using this and the monotonicity property for OWA we complete the proof of the proposition. □

Inclusion principle does not hold in general. We will give a counterexample. 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. If we apply general OWA approach with some vector  $W_L$  on the values obtained above, it is not necessary that we

will obtain 0 at the end. There is a proposed extension such that inclusion property holds. We may take that:

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

It is obvious that in this case we will have inclusion property, however, this extension did not find any interesting applications in practice.

The property of exact approximation also does not hold in general. We will give a counterexample. 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 side, 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. So, from here we can conclude that the equivalence does not hold in general.

## 5 Conclusion and future work

In this article we have presented the theoretical background of integration of fuzzy set theory and DRSA. We also proposed some improvements in the form of OWA which is used to filter outliers and noisy data. We saw that some properties which hold for classical fuzzy DRSA also hold in OWA version with specific assumption. The future work will have both, theoretical and practical aspects. 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 would like to see what will happen if the unions are replaced with an arbitrary set. In that case we would like to see if the similar properties which hold in IRSA, would also hold here. Also, we would like to include checking of these properties for other noise-tolerant extensions of DRSA like VP-DRSA and VC-DRSA. From the practical point of view, we want to see how fuzzy DRSA may improve

existing rule induction algorithms like DomLem and VC-DomLem. Namely, those algorithms showed huge potential in knowledge extraction from data in a form of ‘if ..., then ...’ rules, as well as in ordinal classification problems. It will be interesting to see 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] C. Cornelis, R. Jensen, G. Hurtado, D. Ślezak, Attribute selection with fuzzy decision reducts, *Information Sciences* 180 (2) (2010) 209–224.
- [2] L. D’eer, N. Verbiest, C. Cornelis, L. Godo, A comprehensive study of implicator–conjunctor-based and noise-tolerant fuzzy rough sets: definitions, properties and robustness analysis, *Fuzzy Sets and Systems* 275 (2015) 1–38.
- [3] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, *International Journal of General System* 17 (2-3) (1990) 191–209.
- [4] S. Greco, M. Inuiguchi, R. Słowiński, Dominance-based rough set approach using possibility and necessity measures, in: *International Conference on Rough Sets and Current Trends in Computing*, Springer, 2002, pp. 85–92.
- [5] S. Greco, B. Matarazzo, R. Słowiński, Fuzzy extension of the rough set approach to multicriteria and multiattribute sorting, in: *Preferences and decisions under incomplete knowledge*, Springer, 2000, pp. 131–151.
- [6] S. Greco, B. Matarazzo, R. Słowiński, Rough sets theory for multicriteria decision analysis, *European Journal of Operational Research* 129 (1) (2001) 1–47.
- [7] S. Greco, B. Matarazzo, R. Słowiński, Fuzzy set extensions of the dominance-based rough set approach, in: *Fuzzy sets and their extensions: representation, aggregation and models*, Springer, 2008, pp. 239–261.
- [8] S. Greco, B. Matarazzo, R. Słowiński, J. Stefanowski, Variable consistency model of dominance-based rough sets approach, in: *International Conference on Rough Sets and Current Trends in Computing*, Springer, 2000, pp. 170–181.
- [9] M. Inuiguchi, Y. Yoshioka, Variable-precision dominance-based rough set approach, in: *International Conference on Rough Sets and Current Trends in Computing*, Springer, 2006, pp. 203–212.
- [10] R. Jensen, C. Cornelis, Fuzzy-rough instance selection, in: *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, IEEE, 2010, pp. 1–7.
- [11] G. Klir, B. Yuan, *Fuzzy sets and fuzzy logic*, Vol. 4, Prentice hall New Jersey, 1995.
- [12] J. J. Liou, G.-H. Tzeng, A dominance-based rough set approach to customer behavior in the airline market, *Information Sciences* 180 (11) (2010) 2230–2238.
- [13] Z. Pawlak, Rough sets, *International journal of computer & information sciences* 11 (5) (1982) 341–356.
- [14] E. Ramentol, S. Vluymans, N. Verbiest, Y. Caballero, R. Bello, C. Cornelis, F. Herrera, Ifrowann: imbalanced fuzzy-rough ordered weighted average nearest neighbor classification, *IEEE Transactions on Fuzzy Systems* 23 (5) (2015) 1622–1637.
- [15] S. Vluymans, Dealing with imbalanced and weakly labelled data in machine learning using fuzzy and rough set methods, Ph.D. thesis, Springer (2018).
- [16] S. Vluymans, C. Cornelis, F. Herrera, Y. Saeys, Multi-label classification using a fuzzy rough neighborhood consensus, *Information Sciences* 433 (2018) 96–114.
- [17] S. Vluymans, D. Sánchez Tarragó, Y. Saeys, C. Cornelis, F. Herrera, Fuzzy rough classifiers for class imbalanced multi-instance data, *Pattern Recognition* 53 (2016) 36–45.
- [18] R. R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, *IEEE Transactions on systems, Man, and Cybernetics* 18 (1) (1988) 183–190.
- [19] X. Yang, J. Yang, C. Wu, D. Yu, Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, *Information Sciences* 178 (4) (2008) 1219–1234.
- [20] X. Yang, D. Yu, J. Yang, L. Wei, Dominance-based rough set approach to incomplete interval-valued information system, *Data & Knowledge Engineering* 68 (11) (2009) 1331–1347.