

Similarity-based Classification with Dominance-based Decision Rules

Marcin Szeląg¹, Salvatore Greco^{2,3}, Roman Słowiński^{1,4}

¹ Institute of Computing Science, Poznań University of Technology,
60-965 Poznań, Poland, {mszelag, rslowinski}@cs.put.poznan.pl

² Department of Economics and Business, University of Catania,
Corso Italia, 55, 95129 Catania, Italy, salgrec@unict.it

³ Portsmouth Business School, University of Portsmouth, UK

⁴ Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw

Abstract. We consider a similarity-based classification problem where a new case (object) is classified based on its similarity to some previously classified cases. In this process of case-based reasoning (CBR), we adopt the Dominance-based Rough Set Approach (DRSA), that is able to handle monotonic relationship “the more similar is object y to object x with respect to the considered features, the closer is y to x in terms of the membership to a given decision class X ”. At the level of marginal similarity concerning single features, we consider this similarity in ordinal terms only. The marginal similarities are aggregated within induced decision rules describing monotonic relationship between comprehensive similarity of objects and their similarities with respect to single features.

Key words: Classification, Similarity, Case-based reasoning, Dominance-based Rough Set Approach, Decision rules

1 Introduction

People tend to solve new problems using the solutions of *similar* problems encountered in the past. This process is often referred to as *case-based reasoning* (CBR) [9]. As observed by Gilboa and Schmeidler [3], the basic idea of CBR can be found in the following sentence of Hume [8]: “From causes which appear similar we expect similar effects. This is the sum of all our experimental conclusions.” We can rephrase this sentence by saying: “The more similar are the causes, the more similar one expects the effects”.

We consider classification performed according to the (broadly construed) CBR paradigm, i.e., a *similarity-based classification*. In the similarity-based classification problem, there is given a finite set of training objects (case base), described by a set of *features*, a set of *marginal similarity functions* (one for each feature), and a set of predefined *decision classes*. This information is used to suggest membership of a new (unseen) object to particular decision classes.

In case-based reasoning, one needs a *similarity model* aggregating marginal similarities into comprehensive similarity. Traditionally, this model has the form

of a real-valued aggregation function (e.g., Euclidean norm) or binary relation (e.g., fuzzy relation). In this paper, we present a method based on the Dominance-based Rough Set Approach (DRSA) [4, 11, 12], using a new similarity model in terms of a set of *if-then decision rules* employing dominance relation in the space created by marginal similarity functions. The first method concerning application of DRSA to CBR was introduced in [5–7], and then extended in [14]. The method presented in this paper, first described in an unpublished PhD thesis [13], concerns revision and improvement of the approach given in [14].

The proposed rule-based similarity model makes it possible to avoid an arbitrary aggregation of marginal similarity functions. In this approach, comprehensive similarity is represented by decision rules *induced* from classification examples. These rules underline the monotonic relationship “the more similar is object y to object x with respect to the considered features, the closer is y to x in terms of the membership to a given decision class X ”. Violation of this principle causes an inconsistency in the set of objects, which is handled using DRSA. An important characteristic of the proposed approach is that induced rules employ only ordinal properties of marginal similarity functions. Thus, this approach is invariant to ordinally equivalent marginal similarity functions.

We improve over [14] by proposing a way of inducing decision rules, and by introducing a new rule-based classification scheme extending the one given in [1].

This paper is organized as follows. Section 2 describes problem setting. In Section 3, we discuss basic notions and assumptions. Section 4 defines considered similarity learning task. In Section 5, we introduce two *comprehensive closeness relations*. Section 6 defines rough approximations of the sets of objects being in either kind of comprehensive closeness relation with a reference object x . In Section 7, we describe induction of monotonic decision rules from the rough approximations. Section 8 concerns application of induced rules. In Section 9, we present an illustrative example. Section 10 concludes the paper.

2 Problem Setting

We consider the following classification problem setting. There is given a finite set of objects U (*case base*) and a finite family of pre-defined decision classes \mathcal{D} . An object $y \in U$ (a “case”) is described in terms of features $f_1, \dots, f_n \in F$. For each feature $f_i \in F$, there is given a *marginal similarity function* $\sigma_{f_i} : U \times U \rightarrow [0, 1]$, such that the value $\sigma_{f_i}(y, x)$ expresses the similarity of object $y \in U$ to object $x \in U$ with respect to (w.r.t.) feature f_i , and for all $x, y \in U$, $\sigma_{f_i}(y, x) = 1 \Leftrightarrow f_i(y) = f_i(x)$. Moreover, for each object $y \in U$ there is given an information concerning *credibility* of its membership to each of the considered classes. To admit graded credibilities, each class $X \in \mathcal{D}$ is modeled as a fuzzy set in U [15], characterized by membership function $\mu_X : U \rightarrow [0, 1]$. Thus, each object $y \in U$ can belong to different decision classes with different degrees of membership. The above input information is processed to produce a recommendation concerning a new object z , in terms of a degree of membership of z to particular classes.

3 Basic Notions and Assumptions

Pairwise fuzzy information base. Given the problem setting introduced in Section 2, a *pairwise fuzzy information base* \mathbf{B} [5–7] is the 3-tuple

$$\mathbf{B} = \langle U, F, \Sigma \rangle, \quad (1)$$

where U is a finite set of objects (a case base), F is a finite set of n features, and $\Sigma = \{\sigma_{f_1}, \sigma_{f_2}, \dots, \sigma_{f_n}\}$ is a finite set of n marginal similarity functions.

Marginal similarity functions. Different marginal similarity functions can be used, depending on the value set V_{f_i} of feature $f_i \in F$. For a *numeric feature* f_i , with values on interval or ratio scale, similarity can be defined using a function, e.g., $\sigma_{f_i} = 1 - \frac{|f_i(x) - f_i(y)|}{\max_{v_i \in V_{f_i}} - \min_{v_i \in V_{f_i}}}$. For a *nominal feature* f_i , similarity can be defined using a table, like Table 1. The marginal similarity functions create an

Table 1. Exemplary definition of similarity for a nominal feature $f_i \in F$

$f_i(x) \setminus f_i(y)$	low	medium	high
low	1.0	0.6	0.3
medium	0.6	1.0	0.5
high	0.3	0.5	1.0

n -dimensional *similarity space*.

Problem decomposition. We consider the decision classes belonging to family \mathcal{D} to be mutually independent in the sense of membership function values. Then, we decompose the original multi-class problem π to a set of *single-class subproblems* π_X , where $X \in \mathcal{D}$. Thus, each subproblem concerns a single decision class $X \in \mathcal{D}$ with membership function $\mu_X : U \rightarrow [0, 1]$. In each subproblem, let

$$V_{\mu_X} = \{\mu_X(y) : y \in U\}. \quad (2)$$

Reference objects. We assume that for each subproblem π_X , there is given a set of so-called *reference objects* $U_X^R \subseteq U$. These are objects to which objects from set U are going to be compared. The reference objects may be indicated by a user, and thus, the set of reference objects should be relatively small. If such information is not available, one can use clustering to choose a suitable set of reference objects, sample U , or treat all the objects from U as the reference ones.

4 Similarity Learning

The method proposed in this paper is designed for the following learning task. Given: i) the pairwise fuzzy information base \mathbf{B} , ii) the family \mathcal{D} of decision

classes, implying subproblems π_X , $X \in \mathcal{D}$, iii) the membership functions $\mu_X : U \rightarrow [0, 1]$, $X \in \mathcal{D}$, and iv) the sets of reference objects $U_X^R \subseteq U$, $X \in \mathcal{D}$, learn, for each subproblem π_X , set of decision rules

$$R_X = \bigcup_{x \in U_X^R} R_X(x), \quad (3)$$

where $R_X(x)$ is the set of rules describing membership of an object $y \in U$ to class $X \in \mathcal{D}$ based on similarity of y to reference object $x \in U_X^R$.

5 Comprehensive Closeness of Objects

Given a decision class X being a fuzzy set in U , we define two kinds of binary *comprehensive closeness relations* on U :

$$y \succ_{\alpha, \beta}^X x \Leftrightarrow \mu_X(x) \in [\alpha, \beta] \text{ and } \mu_X(y) \in [\alpha, \beta], \quad (4)$$

$$y \lesssim_{\alpha, \beta}^X x \Leftrightarrow \mu_X(x) \in [\alpha, \beta] \text{ and } \mu_X(y) \notin (\alpha, \beta), \quad (5)$$

where $y, x \in U$ and $-\delta \leq \alpha \leq \beta \leq 1 + \delta$, where $\delta \in \mathbb{R}_+$ is any fixed positive value (a technical parameter, e.g., 0.01). When $y \succ_{\alpha, \beta}^X x$, then $\alpha \leq \mu_X(y) \leq \mu_X(x) \leq \beta$ or $\alpha \leq \mu_X(x) \leq \mu_X(y) \leq \beta$, i.e., looking from the perspective of y , $\mu_X(y)$ is on the left side of $\mu_X(x)$ but not farther than α , or $\mu_X(y)$ is on the right side of $\mu_X(x)$ but not farther than β . When $y \lesssim_{\alpha, \beta}^X x$, then $\mu_X(y)$ is on the left side of $\mu_X(x)$ but not closer than α , or $\mu_X(y)$ is on the right side of $\mu_X(x)$ but not closer than β . Thus, α and β play roles of limiting levels of membership to X .

The “special” values $-\delta$ and $1 + \delta$, where $\delta \in \mathbb{R}_+$, are considered in (5) to allow, respectively, $\mu_X(y) \notin (-\delta, \beta)$ (i.e., $\mu_X(y) \geq \beta$) and $\mu_X(y) \notin (\alpha, 1 + \delta)$ (i.e., $\mu_X(y) \leq \alpha$). This is crucial, e.g., when X is crisp – one can then consider two meaningful relations $\lesssim_{0, 1 + \delta}^X$ and $\lesssim_{-\delta, 1}^X$, composed of pairs $(y, x) \in U \times U$ such that $\mu_X(y) \leq 0$ and $\mu_X(y) \geq 1$, respectively.

Let us observe that $\succ_{\alpha, \beta}^X$ is reflexive, symmetric and transitive and thus it is an equivalence relation. Moreover, $\lesssim_{\alpha, \beta}^X$ is only transitive.

Given a class X and a reference object $x \in U_X^R$, we are interested in characterizing, in terms of similarity-based decision rules, the objects $y \in U$ being in:

- $\succ_{\alpha, \beta}^X$ relation with x , where $\alpha, \beta \in V_{\mu_X}$,
- $\lesssim_{\alpha, \beta}^X$ relation with x , where $\alpha, \beta \in V_{\mu_X} \cup \{-\delta\} \cup \{1 + \delta\}$, $\alpha < \mu_X(x) < \beta$.

Let $V_{\mu_X}^\delta = V_{\mu_X} \cup \{-\delta\} \cup \{1 + \delta\}$, where $\delta \in \mathbb{R}_+$. We define two types of sets:

$$S(\succ_{\alpha, \beta}^X, x) = \{y \in U : y \succ_{\alpha, \beta}^X x\}, \quad \text{where } \alpha, \beta \in V_{\mu_X}, \alpha \leq \mu_X(x) \leq \beta, \quad (6)$$

$$S(\lesssim_{\alpha, \beta}^X, x) = \{y \in U : y \lesssim_{\alpha, \beta}^X x\}, \quad \text{where } \alpha, \beta \in V_{\mu_X}^\delta, \alpha < \mu_X(x) < \beta. \quad (7)$$

The strict constraint $\alpha < \mu_X(x) < \beta$ in (7) prevents from considering not meaningful sets $S(\lesssim_{\alpha, \beta}^X, x)$ [13]. From this point of view, it is crucial that when $\mu_X(x) = 0$ (or $\mu_X(x) = 1$), one can take $\alpha = -\delta$ (or $\beta = 1 + \delta$, respectively).

The sets of objects defined by (6) and (7) are to be approximated using dominance cones in the similarity space created by functions $\sigma_{f_1}, \dots, \sigma_{f_n}$.

6 Rough Approximation by Dominance Relation

Let us define the dominance relation w.r.t. the similarity to an object $x \in U$, called in short *x-dominance relation*, defined over U , and denoted by D_x . For any $x, y, w \in U$, y is said to *x-dominate* w (denotation yD_xw) if for every $f_i \in F$,

$$\sigma_{f_i}(y, x) \geq \sigma_{f_i}(w, x). \quad (8)$$

Thus, object y is said to *x-dominate* object w iff for every feature $f_i \in F$, y is at least as similar to x as w is.

Given an object $y \in U$, *x-positive* and *x-negative dominance cones* of y in the similarity space are defined as follows:

$$D_x^+(y) = \{w \in U : wD_xy\}, \quad (9)$$

$$D_x^-(y) = \{w \in U : yD_xw\}. \quad (10)$$

In order to induce meaningful certain and possible decision rules concerning similarity to a reference object $x \in U_X^R$, we structure the objects $y \in U$ by calculation of lower and upper approximations of sets $S(\succ_{\alpha, \beta}^X, x)$ and $S(\preccurlyeq_{\alpha, \beta}^X, x)$.

The *lower approximations* of sets $S(\succ_{\alpha, \beta}^X, x)$ and $S(\preccurlyeq_{\alpha, \beta}^X, x)$ are defined as:

$$\underline{S(\succ_{\alpha, \beta}^X, x)} = \{y \in U : D_x^+(y) \subseteq S(\succ_{\alpha, \beta}^X, x)\}, \quad (11)$$

$$\underline{S(\preccurlyeq_{\alpha, \beta}^X, x)} = \{y \in U : D_x^-(y) \subseteq S(\preccurlyeq_{\alpha, \beta}^X, x)\}, \quad (12)$$

and the *upper approximations* of sets $S(\succ_{\alpha, \beta}^X, x)$ and $S(\preccurlyeq_{\alpha, \beta}^X, x)$ are defined as:

$$\overline{S(\succ_{\alpha, \beta}^X, x)} = \{y \in U : D_x^-(y) \cap S(\succ_{\alpha, \beta}^X, x) \neq \emptyset\}, \quad (13)$$

$$\overline{S(\preccurlyeq_{\alpha, \beta}^X, x)} = \{y \in U : D_x^+(y) \cap S(\preccurlyeq_{\alpha, \beta}^X, x) \neq \emptyset\}. \quad (14)$$

With respect to the three basic properties of set approximations defined for rough sets in [10], it follows from definitions (6), (7), (11), (12), (13), and (14), that lower and upper approximations defined above fulfill properties of *rough inclusion* and *monotonicity of the accuracy of approximation*. Moreover, these approximations enjoy also *complementarity* property, as shown in [14].

Using (11), (12), (13), and (14), one can define the *boundary* of set $S(\succ_{\alpha, \beta}^X, x)$ (or set $S(\preccurlyeq_{\alpha, \beta}^X, x)$), as the difference between its upper and lower approximation. It is also possible to perform further DRSA-like analysis by calculating the quality of approximation, reducts, and the core (see, e.g., [4, 11, 12]).

7 Induction of Decision Rules

Lower (or upper) approximations of considered sets $S(\succ_{\alpha, \beta}^X, x)$ and $S(\preccurlyeq_{\alpha, \beta}^X, x)$ are the basis for induction of *certain* (or *possible*) *decision rules* belonging to set $R_X(x)$, $x \in U_X^R$. We distinguish two types of rules:

- (1) *at least rules*:
 if $\sigma_{f_{i1}}(y, x) \geq h_{i1} \dots$ and $\sigma_{f_{ip}}(y, x) \geq h_{ip}$, then certainly (or possibly) $y \succsim_{\alpha, \beta}^X x$,
- (2) *at most rules*:
 if $\sigma_{f_{i1}}(y, x) \leq h_{i1} \dots$ and $\sigma_{f_{ip}}(y, x) \leq h_{ip}$, then certainly (or possibly) $y \precsim_{\alpha, \beta}^X x$,

where $\{f_{i1}, \dots, f_{ip}\} \subseteq F$, $h_{i1}, \dots, h_{ip} \in [0, 1]$, and α, β satisfy $0 \leq \alpha \leq \mu_X(x) \leq \beta \leq 1$ in case of at least rules, and $-\delta \leq \alpha < \mu_X(x) < \beta \leq 1 + \delta$ in case of at most rules, $\delta \in \mathbb{R}_+$.

Remark that according to definitions (4) and (5), the decision part of the rule of type (1) and (2) can be rewritten, respectively, as:

- (1) “then certainly (or possibly) $\mu_X(y) \in [\alpha, \beta]$ ”, i.e., the conclusion is that the membership of object y to decision class X is inside the interval $[\alpha, \beta]$,
- (2) “then certainly (or possibly) $\mu_X(y) \notin (\alpha, \beta)$ ”, i.e., the conclusion is that the membership of object y to decision class X is outside the interval (α, β) .

A certain rule of type (1) is read as: “if similarity of object y to reference object x w.r.t. feature f_{i1} is at least $h_{i1} \dots$ and similarity of y to x w.r.t. feature f_{ip} is at least h_{ip} , then certainly y belongs to class X with credibility between α and β . A possible rule of type (2) is read as: “if similarity of object y to reference object x w.r.t. feature f_{i1} is at most $h_{i1} \dots$ and similarity of y to x w.r.t. feature f_{ip} is at most h_{ip} , then possibly y belongs to class X with credibility at most α or at least β .”

Decision rules of type (1) and (2) can be induced using the VC-DomLEM algorithm [2]. On one hand, these rules reveal similarity-based patterns present in the training data. On the other hand, set $R_X = \bigcup_{x \in U_X^R} R_X(x)$ of induced certain/possible rules can be applied to *classify* new objects (new cases).

8 Application of Decision Rules

The rules from R_X can be applied to a new object z , described in terms of features $f_1, \dots, f_n \in F$, to predict its degree of membership to class X . Then, the rules covering z may give an ambiguous classification suggestion (intervals of μ_X instead of a crisp value). In order to resolve this ambiguity, we adapt and revise the rule classification scheme described in [1]. In this way, one can obtain a precise (crisp) value of membership $\mu_X(z)$. Let us consider three situations, assuming that $Cov_z \subseteq R_X$ denotes the set of rules covering object z , $Cond_\rho \subseteq U$ denotes the set of objects covered by rule ρ , $U_X^t = \{y \in U : \mu_X(y) = t\}$, and $|\cdot|$ denotes cardinality of a set.

Situation (i). No rule from R_X covers object z (i.e., $Cov_z = \emptyset$), so there is no reliable suggestion concerning $\mu_X(z)$. If a concrete answer is expected, one can suggest that $\mu_X(z)$ equals to the most frequent value $\mu_X(y)$, where $y \in U$.

Situation (ii). Exactly one rule $\rho \in R_X(x) \subseteq R_X$, $x \in U_X^R$, covers object z (i.e., $|Cov_z| = 1$). Then, we calculate value $Score_X^\rho(t, z)$ for each membership value $t \in V_{\mu_X}$ covered by the decision part of this rule:

$$Score_X^\rho(t, z) = \frac{|Cond_\rho \cap U_X^t|^2}{|Cond_\rho| |U_X^t|}. \quad (15)$$

Then, $\mu_X(z)$ is calculated as $\mu_X(z) = \max_t \text{Score}_X^\rho(t, z)$. Let us observe that $\text{Score}_X^\rho(t, z) \in [0, 1]$. It can be interpreted as the degree of certainty of the suggestion that $\mu_X(z)$ equals t .

Situation (iii). Several rules from R_X cover object z (i.e., $|Cov_z| > 1$). Then, we calculate value $\text{Score}_X(t, z)$ for each $t \in V_{\mu_X}$ covered by the decision part of any of the covering rules:

$$\text{Score}_X(t, z) = \text{Score}_X^+(t, z) - \text{Score}_X^-(t, z), \quad (16)$$

where $\text{Score}_X^+(t, z)$ and $\text{Score}_X^-(t, z)$ represent the positive and negative part of $\text{Score}_X(t, z)$, respectively. $\text{Score}_X^+(t, z)$ takes into account rules $\rho_1, \dots, \rho_k \in Cov_z$ whose decision part covers t :

$$\text{Score}_X^+(t, z) = \frac{|(\text{Cond}_{\rho_1} \cap U_X^t) \cup \dots \cup (\text{Cond}_{\rho_k} \cap U_X^t)|^2}{|\text{Cond}_{\rho_1} \cup \dots \cup \text{Cond}_{\rho_k}| |U_X^t|}. \quad (17)$$

Let us observe that $\text{Score}_X^+(t, z) \in [0, 1]$. $\text{Score}_X^-(t, z)$ takes into account the rules $\rho_{k+1}, \dots, \rho_h \in Cov_z$ whose decision part does not cover t . If there is no such rule, then $\text{Score}_X^-(t, z) = 0$. Otherwise:

$$\text{Score}_X^-(t, z) = \frac{|(\text{Cond}_{\rho_{k+1}} \cap U_X^{\rho_{k+1}}) \cup \dots \cup (\text{Cond}_{\rho_h} \cap U_X^{\rho_h})|^2}{|\text{Cond}_{\rho_{k+1}} \cup \dots \cup \text{Cond}_{\rho_h}| |U_X^{\rho_{k+1}} \cup \dots \cup U_X^{\rho_h}|}, \quad (18)$$

where U_X^ρ is subset of U containing objects whose membership to class X is covered by the decision part of rule ρ . Let us observe that $\text{Score}_X^-(t, z) \in [0, 1]$.

After calculating $\text{Score}_X(t, z)$ for all considered values of t , we take $\mu_X(z) = \max_t \text{Score}_X(t, z)$. It can be interpreted as a net balance of the arguments in favor and against the suggestion “the membership of object z to class X equals t ”.

9 Illustrative Example

Let us consider set U composed of five objects described by two features: f_1 , with value set $[0, 8]$, and f_2 , with value set $[0, 1]$. Moreover, let us consider decision class X , with membership function μ_X . The five objects are presented in Fig. 1.

We assume that object x is a reference object, and that there are given two marginal similarity functions $\sigma_{f_1}, \sigma_{f_2}$ defined as:

$$\sigma_{f_i}(y, x) = 1 - \frac{|f_i(y) - f_i(x)|}{f_i^{\max} - f_i^{\min}},$$

where $i = 1, 2$, and f_i^{\max}, f_i^{\min} denote max and min value in the value set of f_i .

Functions σ_{f_1} and σ_{f_2} create a 2-dimensional similarity space. Fig. 2 shows pairs of objects (\cdot, x) in this space.

First, using (9) and (10), we calculate x -positive and x -negative dominance cones in the similarity space. Two such cones are shown in Fig. 1 and in Fig. 2.

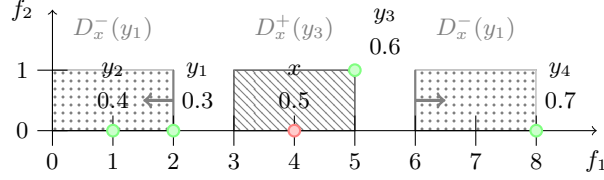


Fig. 1. Set of objects considered in the illustrative example; the number below an object id denotes the value of function μ_X for this object; the hatched area corresponds to dominance cone $D_x^+(y_3)$, and the two dotted areas (one for $f_1(y) \leq 2$, and the other for $f_1(y) \geq 6$) correspond to dominance cone $D_x^-(y_1)$

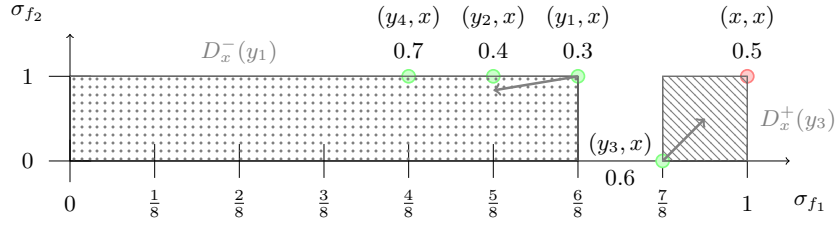


Fig. 2. Pairs of objects (\cdot, x) in the similarity space created by σ_{f_1} and σ_{f_2} ; the number below a pair of object ids denotes the value of function μ_X for the object whose id is the first in the pair; the hatched area corresponds to dominance cone $D_x^+(y_3)$, and the dotted area corresponds to dominance cone $D_x^-(y_1)$

Second, we calculate sets of objects $S(\succ_{\alpha, \beta}^X, x)$ according to (6), for $\alpha \in \{0.3, 0.4, 0.5\}$ and $\beta \in \{0.5, 0.6, 0.7\}$. Moreover, we calculate $S(\prec_{\alpha, \beta}^X, x)$ according to (7), for $\alpha \in \{-\delta, 0.3, 0.4\}$ and $\beta \in \{0.6, 0.7, 1 + \delta\}$, where $\delta \in \mathbb{R}_+$.

Third, sets $S(\succ_{\alpha, \beta}^X, x)$ and $S(\prec_{\alpha, \beta}^X, x)$ are approximated using the x -positive and x -negative dominance cones in the similarity space – see Table 2.

One can observe several inconsistencies w.r.t. the x -dominance relation in the similarity space. E.g., objects $y_2, y_4 \in S(\succ_{0.4, 0.7}^X, x)$ are inconsistent since they are x -dominated by object y_1 , and $y_1 \notin S(\succ_{0.4, 0.7}^X, x)$ (because $\mu_X(y_1) = 0.3$).

Table 3 presents minimal decision rules induced by VC-DomLEM algorithm from the non-empty lower and upper approximations shown in Table 2.

Example of application of induced decision rules. Consider a new object z such that $f_1(z) = 5.5$, $f_2(z) = 0.5$, and thus, $\sigma_{f_1}(z, x) = 6.5/8$, $\sigma_{f_2}(z, x) = 1/2$. Object z is covered by rules ρ_2 and ρ_{15} , suggesting $\mu_X(z) \in [0.3, 0.6]$ and $\mu_X(z) \notin (0.4, 0.6)$, respectively. Applying (17), (18), and (16) for each membership degree $t \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$, we get the result shown in Table 4. Consequently, one can conclude that $\mu_X(z)$ is equal to 0.3, 0.4, or 0.6.

Table 2. Approximations of sets $S(\succ_{\alpha,\beta}^X, x)$, $S(\preccurlyeq_{\alpha,\beta}^X, x)$; $\delta \in \mathbb{R}_+$; objects struck through belong to respective set but not to its lower approximation; underlined objects do not belong to respective set but belong to its upper approximation

$S(\succ_{\alpha,\beta}^X, x)$	$\beta = 0.5$	$\beta = 0.6$	$\beta = 0.7$	$S(\preccurlyeq_{\alpha,\beta}^X, x)$	$\beta = 0.6$	$\beta = 0.7$	$\beta = 1 + \delta$
$\alpha = 0.3$	$\{y_1, y_2, x\}$	$\{y_1, y_2, y_3, x\}$	U	$\alpha = -\delta$	$\{y_3, y_4\}$	$\{y_4\}$	\emptyset
$\alpha = 0.4$	$\{\cancel{y_2}, x\}$	$\{\cancel{y_2}, y_3, x\}$	$\{\cancel{y_2}, y_3, \cancel{y_4}, x\}$	$\alpha = 0.3$	$\{\cancel{y_1}, y_3, y_4\}$	$\{\cancel{y_1}, y_4\}$	$\{\cancel{y_1}\}$
$\alpha = 0.5$	$\{x\}$	$\{y_3, x\}$	$\{y_3, \cancel{y_4}, x\}$	$\alpha = 0.4$	$\{y_1, y_2, y_3, y_4\}$	$\{y_1, y_2, y_4\}$	$\{\cancel{y_1}, \cancel{y_2}\}$
$S(\preccurlyeq_{\alpha,\beta}^X, x)$	$\beta = 0.5$	$\beta = 0.6$	$\beta = 0.7$	$S(\succ_{\alpha,\beta}^X, x)$	$\beta = 0.6$	$\beta = 0.7$	$\beta = 1 + \delta$
$\alpha = 0.3$	$\{y_1, y_2, x\}$	$\{y_1, y_2, y_3, x\}$	U	$\alpha = -\delta$	$\{y_3, y_4\}$	$\{y_4\}$	\emptyset
$\alpha = 0.4$	$\{\underline{y_1}, y_2, x\}$	$\{\underline{y_1}, y_2, y_3, x\}$	$\{\underline{y_1}, y_2, y_3, y_4, x\}$	$\alpha = 0.3$	$\{y_1, \underline{y_2}, y_3, y_4\}$	$\{y_1, \underline{y_2}, y_4\}$	$\{y_1, \underline{y_2}, \underline{y_4}\}$
$\alpha = 0.5$	$\{x\}$	$\{y_3, x\}$	$\{y_1, y_2, y_3, y_4, x\}$	$\alpha = 0.4$	$\{y_1, y_2, y_3, y_4\}$	$\{y_1, y_2, y_4\}$	$\{y_1, y_2, \underline{y_4}\}$

Table 3. Rules induced for referent x such that $f_1(x) = 4, f_2(x) = 0$; ‘Supp.’ (‘ \neg Supp.’) presents ids of objects supporting a rule (covered by a rule but not supporting it); 8 possible rules, identical (except for “possibly”) to respective certain rules, are skipped

Id	Decision rule	Supp.	\neg Supp.
ρ_1	if $\sigma_{f_1}(y, x) \geq \frac{5}{8}$ and $\sigma_{f_2}(y, x) \geq 1$, then certainly $\mu_X(y) \in [0.3, 0.5]$	$\{y_1, y_2, x\}$	
ρ_2	if $\sigma_{f_1}(y, x) \geq \frac{5}{8}$, then certainly $\mu_X(y) \in [0.3, 0.6]$	$\{y_1, y_2, y_3, x\}$	
ρ_3	if $\sigma_{f_1}(y, x) \geq 1$, then certainly $\mu_X(y) \in [0.5, 0.5]$	$\{x\}$	
ρ_4	if $\sigma_{f_1}(y, x) \geq \frac{7}{8}$, then certainly $\mu_X(y) \in [0.5, 0.6]$	$\{y_3, x\}$	
ρ_7	if $\sigma_{f_1}(y, x) \geq \frac{5}{8}$ and $\sigma_{f_2}(y, x) \geq 1$, then possibly $\mu_X(y) \in [0.4, 0.5]$	$\{y_2, x\}$	$\{y_1\}$
ρ_8	if $\sigma_{f_1}(y, x) \geq \frac{5}{8}$, then possibly $\mu_X(y) \in [0.4, 0.6]$	$\{y_2, y_3, x\}$	$\{y_1\}$
ρ_9	if $\sigma_{f_1}(y, x) \geq \frac{4}{8}$, then possibly $\mu_X(y) \in [0.4, 0.7]$	$\{y_2, y_3, y_4, x\}$	$\{y_1\}$
ρ_{12}	if $\sigma_{f_1}(y, x) \geq \frac{4}{8}$, then possibly $\mu_X(y) \in [0.5, 0.7]$	$\{y_3, y_4, x\}$	$\{y_1, y_2\}$
ρ_{13}	if $\sigma_{f_2}(y, x) \leq 0$, then certainly $\mu_X(y) \geq 0.6$	$\{y_3\}$	
ρ_{14}	if $\sigma_{f_1}(y, x) \leq \frac{4}{8}$, then certainly $\mu_X(y) \geq 0.7$	$\{y_4\}$	
ρ_{15}	if $\sigma_{f_1}(y, x) \leq \frac{7}{8}$, then certainly $\mu_X(y) \notin (0.4, 0.6)$	$\{y_1, y_2, y_3, y_4\}$	
ρ_{16}	if $\sigma_{f_1}(y, x) \leq \frac{6}{8}$, then certainly $\mu_X(y) \notin (0.4, 0.7)$	$\{y_1, y_2, y_4\}$	
ρ_{19}	if $\sigma_{f_1}(y, x) \leq \frac{7}{8}$, then possibly $\mu_X(y) \notin (0.3, 0.6)$	$\{y_1, y_3, y_4\}$	$\{y_2\}$
ρ_{20}	if $\sigma_{f_1}(y, x) \leq \frac{6}{8}$, then possibly $\mu_X(y) \notin (0.3, 0.7)$	$\{y_1, y_4\}$	$\{y_2\}$
ρ_{21}	if $\sigma_{f_1}(y, x) \leq \frac{6}{8}$, then possibly $\mu_X(y) \leq 0.3$	$\{y_1\}$	$\{y_2, y_4\}$
ρ_{24}	if $\sigma_{f_1}(y, x) \leq \frac{6}{8}$, then possibly $\mu_X(y) \leq 0.4$	$\{y_1, y_2\}$	$\{y_4\}$

Table 4. Scores of a new object z resulting from application of induced decision rules

$t \in V_{\mu_X}$	0.3	0.4	0.5	0.6	0.7
$Score_X^+(t, z)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{4}$
$Score_X^-(t, z)$	0	0	1	0	1
$Score_X(t, z)$	$\frac{1}{5}$	$\frac{1}{5}$	$-\frac{3}{4}$	$\frac{1}{5}$	$-\frac{3}{4}$

10 Conclusions

We presented a method of similarity-based classification using the Dominance-based Rough Set Approach. This method exploits only ordinal properties of

marginal similarity functions and membership functions of decision classes. It avoids arbitrary aggregation of marginal similarities into one comprehensive similarity. Instead, it uses a rule-based similarity model employing rules describing monotonic relationship between comprehensive similarity of objects and their similarities with respect to single features. Thus, our case-based reasoning approach is as much “neutral” and “objective” as possible. Moreover, our method provides more insight when determining membership of a new object z to class X – one can see the rules matching z and the objects supporting these rules.

Acknowledgments. The first author wishes to acknowledge financial support from the Faculty of Computing at Poznań University of Technology, grant 09/91/DSMK/0609.

References

1. Błaszczyński, J., Greco, S., Słowiński, R.: Multi-criteria classification – a new scheme for application of dominance-based decision rules. *European Journal of Operational Research* 181(3), 1030–1044 (2007)
2. Błaszczyński, J., Słowiński, R., Szeląg, M.: Sequential covering rule induction algorithm for variable consistency rough set approaches. *INS* 181, 987–1002 (2011)
3. Gilboa, I., Schmeidler, D.: *A Theory of Case-Based Decisions*. No. 9780521003117 in Cambridge Books, Cambridge University Press (February 2001)
4. Greco, S., Matarazzo, B., Słowiński, R.: Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research* 129(1), 1–47 (2001)
5. Greco, S., Matarazzo, B., Słowiński, R.: Dominance-based rough set approach to case-based reasoning. In: Torra, V., et al. (eds.) *Modelling Decisions for Artificial Intelligence*. LNAI, vol. 3885, pp. 7–18. Springer, Berlin (2006)
6. Greco, S., Matarazzo, B., Słowiński, R.: Case-based reasoning using gradual rules induced from dominance-based rough approximations. In: Wang, G., et al. (eds.) *RSKT*. LNAI, vol. 5009, pp. 268–275. Springer, Berlin (2008)
7. Greco, S., Matarazzo, B., Słowiński, R.: Granular computing for reasoning about ordered data: the dominance-based rough set approach. In: Pedrycz, W., et al. (eds.) *Handbook of Granular Computing*, chap. 15, pp. 347–373. Wiley (2008)
8. Hume, D.: *An Enquiry Concerning Human Understanding*. Clarendon Press (1748)
9. Kolodner, J.: *Case-Based Reasoning*. Morgan Kaufmann, San Mateo (1993)
10. Pawlak, Z.: *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer, Dordrecht, Netherlands (1991)
11. Słowiński, R., Greco, S., Matarazzo, B.: Rough sets in decision making. In: Meyers, R. (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 7753–7786. Springer, New York (2009)
12. Słowiński, R., Greco, S., Matarazzo, B.: Rough set based decision support. In: Burke, E.K., Kendall, G. (eds.) *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, pp. 557–609. Springer (2014)
13. Szeląg, M.: *Application of the Dominance-based Rough Set Approach to Ranking and Similarity-based Classification Problems*. Ph.D. thesis, Poznań University Of Technology (2015), <http://www.cs.put.poznan.pl/mszelag/Research/MSzPhD.pdf>
14. Szeląg, M., Greco, S., Błaszczyński, J., Słowiński, R.: Case-based reasoning using dominance-based decision rules. In: Yao, J., et al. (eds.) *RSKT*. LNCS, vol. 6954, pp. 404–413. Springer, Heidelberg (2011)
15. Zadeh, L.: Fuzzy sets. *Information and Control* 8(3), 338–353 (1965)