Dominance-Based Rough Set Approach to Case-Based Reasoning

Salvatore Greco[1], Benedetto Matarazzo[1], and Roman Slowinski[2]

[1] Faculty of Economics, University of Catania,
Corso Italia, 55, 95129 – Catania, Italy
[2] Institute of Computing Science, Poznan University of Technology,
60-965 Poznan, and Institute for Systems Research,
Polish Academy of Sciences, 01-447 Warsaw, Poland

Abstract. Case-based reasoning is a paradigm in machine learning whose idea is that a new problem can be solved by noticing its similarity to a set of problems previously solved. We propose a new approach to case-based reasoning. It is based on rough set theory that is a mathematical theory for reasoning about data. More precisely, we adopt Dominance-based Rough Set Approach (DRSA) that is particularly appropriate in this context for its ability of handling monotonicity relationship between ordinal properties of data related to monotonic relationships between attribute values in the considered data set. In general terms, monotonicity concerns relationship between different aspects of a phenomenon described by data: for example, "the larger the house, the higher its price" or "the closer the house to the city centre, the higher its price". In the perspective of case-based reasoning, we propose to consider monotonicity of the type "the more similar is y to x, the more credible is that y belongs to the same set as x". We show that rough approximations and decision rules induced from these approximations can be redefined in this context and that they satisfy the same fundamental properties of classical rough set theory.

1 Introduction

Case-based reasoning (for a general introduction to case-based reasoning see e.g. [10]; for a fuzzy set approach to case-based reasoning see [3]) regards the inference of some proper conclusions related to a new situation by the analysis of similar cases from a memory of previous cases. It is based on two principles [11]:

a) similar problems have similar solutions;
b) types of encountered problems tend to recur.

Gilboa and Schmeidler [4] observed that the basic idea of case-based reasoning can be found in the following sentence of Hume [9]: "From causes which appear *similar* we expect similar effects. This is the sum of all our experimental conclusions." Rephrasing Hume, one can say that "the more similar are the causes, the more similar one expects the effects." Therefore, measuring similarity is the

V. Torra et al. (Eds.): MDAI 2006, LNAI 3885, pp. 7–18, 2006.
© Springer-Verlag Berlin Heidelberg 2006

essential point of all case-based reasoning and, particularly, of fuzzy set approach to case-based reasoning [3]. This explains the many research problems that measuring similarity generates within case-based reasoning. Problems of modelling similarity are relative to two levels:

– at level of similarity with respect to single features: how to define a meaningful similarity measure with respect to a single feature?
– at level of similarity with respect to all features: how to properly aggregate the similarity measure with respect to single features in order to obtain a comprehensive similarity measure?

Taking this into account we propose an approach to case-based reasoning which tries to be possibly "neutral" and "objective" with respect to similarity relation, in the sense that at level of similarity concerning single features, we consider only ordinal properties of similarity, and at level of aggregation, we do not impose any specific functional aggregation based on very specific axioms (see for example [4]), but we consider a set of decision rules based on the very general monotonicity property of comprehensive similarity with respect to similarity of single features. Therefore our approach to case-based reasoning is very little "invasive", comparing to the many other existing approaches.

Our approach to case-based reasoning is based on rough set theory ([12, 13]). Rough set theory relies on the idea that some knowledge (data, information) is available about elements of a set. For example, knowledge about patients suffering from a certain disease may contain information about body temperature, blood pressure, etc. All patients described by the same information are indiscernible in view of the available knowledge and form groups of similar cases. These groups are called elementary sets and can be considered as elementary building blocks of the available knowledge about patients. Elementary sets can be combined into compound concepts. Any union of elementary sets is called crisp set, while other sets are referred to as rough set. Each rough set has boundary line cases, i.e. objects which, in view of the available knowledge, cannot be classified with certainty as members of the set or of its complement. Therefore, in the rough set approach, any set is associated with a pair of crisp sets called the lower and the upper approximation. Intuitively, the lower approximation consists of all objects which certainly belong to the set and the upper approximation contains all objects which possibly belong to the set. The difference between the upper and the lower approximation constitutes the boundary region of the rough set.

In our approach to case-based reasoning we do not consider classical rough set theory but its extension called Dominance-based Rough Set Approach (DRSA) [5, 6] that has been proposed to handle ordinal properties of data related to preferences in decision problems. The monotonicity, which is crucial for DRSA, is also meaningful for problems where preferences are not considered. Generally, monotonicity concerns relationship between different aspects of a phenomenon described by data. More specifically, it concerns mutual trends between different variables like distance and gravity in physics or inflation rate and interest rate in economics. Whenever we discover a relationship between different aspects of a phenomenon, this relationship can be represented by a monotonicity with respect to some specific measures of the considered aspects. So, in general, the monotonicity is a property translating in a formal language a primitive intuition of interaction between different concepts in our knowledge domain. As discovering is an inductive process, it is illuminating to remember the following Proposition 6.363 of Wittgenstein [17]: "*The process of induction is the process of assuming the simplest law that can be made to harmonize with our experience*". We claim that this simplest law is just monotonicity and, therefore, each data analysis method can be seen as a specific way of dealing with monotonicity.

Let us observe that monotonicity is also present in classical rough set theory. In fact, rough set philosophy employs approximation for describing relationships between concepts. For example, coming back to above example of medical diagnosis, the concept of "disease Y" can be represented in terms of such concepts as "low blood pressure and high temperature" or "muscle pain and headache". The approximation is based on a very coarse representation in the sense that, for each aspect characterizing concepts ("low blood pressure", "high temperature", "muscle pain", etc.), only its presence or its absence is considered relevant. Therefore, rough approximation within classical rough set theory involves a very primitive idea of monotonicity related to a scale with only two values: "presence" and "absence".

Monotonicity gains importance when a finer representation of the concepts is considered. A representation is finer when for each aspect characterizing concepts, not only its presence or its absence is taken into account, but also the *graduality* of its presence or absence is considered relevant. Due to graduality, the idea of monotonicity can be exploited in the whole range of its potential. Graduality is typical for fuzzy set philosophy [18] and, therefore, a joint consideration of rough sets and fuzzy sets is worthwhile. In fact, rough set and fuzzy set capture the two basic complementary features of the idea of monotonicity: rough set deals with relationships between different aspects and fuzzy sets deal with expression of different dimensions representing the considered concepts. For this reason, many approaches have been proposed to combine fuzzy sets with rough sets (see for example [2, 16]).

Greco, Matarazzo and Slowinski [7] showed how the framework of DRSA can be very naturally extended to represent any relationship of monotonicity in reasoning about data. In this context one can envisage a knowledge representation model composed of a set of decision rules with the following syntax:

"if object y presents feature f_{i1} in degree at least h_{i1}, and feature f_{i2} in degree at least h_{i2} ..., and feature f_{im} in degree at least h_{im}, then object y belongs to set X in degree at least α".

Greco, Matarazzo and Slowinski [7] proved also that the classical rough set approach [12, 13] can be seen as specific case of the general DRSA model. This is important for several reasons; in particular, this interpretation of DRSA gives an insight into fundamental properties of the classical rough set approach and permits to further generalize the rough set approach.

In this paper, we show that in the framework of DRSA a rough set approach to case-based reasoning can be developed very naturally. Here, the monotonicity concerns the relationships between similarity to some reference objects and membership to some specific sets. In this context we envisage a knowledge representation model composed of a set of decision rules with the following syntax:

"if object y is similar to object x w.r.t. feature f_{i1} in degree at least h_{i1}, and w.r.t. feature f_{i2} in degree at least h_{i2}, and ..., and w.r.t. feature f_{im} in degree at least h_{im}, then object y belongs to set X in degree at least α",
where w.r.t. means "with respect to".

These decision rules are similar to the gradual decision rules [1] being statements of the form "the more object z is X, the more it is Y" or, equivalently, but more technically,

$$\mu_X(z) \geq \alpha \Rightarrow \mu_Y(z) \geq \alpha$$

where X and Y are fuzzy sets whose membership functions are μ_Y and μ_X, and $\alpha \in [0,1]$.

Within the context of case-based reasoning gradual decision rules assume the following syntax [3]:

"the more object z is similar to a referent object x w.r.t. condition attribute s, the more z is similar to a referent object x w.r.t. decision attribute t"

or, equivalently, but more technically,

$$s(z,x) \geq \alpha \Rightarrow t(z,x) \geq \alpha$$

where s and t measure the credibility of similarity with respect to condition attribute and decision attribute, respectively.

When there is a plurality of condition attributes and decision attributes, functions s and t aggregate similarity with respect to these attributes.

Let us observe that the decision rules we propose do not need the aggregation of the similarity with respect to different features in one comprehensive similarity. This is important because it permits to avoid using aggregation operators (weighted average, min, etc.) which are always arbitrary to some extend. Moreover, the decision rules we propose permit to consider different thresholds for degrees of credibility in the premise and in the conclusion. This is not considered in the gradual decision rules, where the threshold is the same, α, in the premise and in the conclusion.

This article is organized as follows. Section 2 introduces DRSA approach to case-based similarity. Section 3 contains conclusions.

2 Rough Approximation for Case Based Reasoning

In this section, we consider rough approximation of a fuzzy set using a similarity relation in the context of case-based reasoning. The introduced rough approximation is inspired by the rough approximation of a pairwise comparison table within the Dominance-based Rough Set Approach (DRSA).

Let us consider a *pairwise fuzzy information base* being the 3-tuple

$$\boldsymbol{B} = < U, F, \sigma >,$$

where U is a finite set of *objects* (universe), $F=\{f_1,f_2,...,f_m\}$ is a finite set of *features*, and $\sigma: U \times U \times F \rightarrow [0,1]$ is a function such that $\sigma(x,y,f_h) \in [0,1]$ expresses the credibility that object x is similar to object y w.r.t. to feature f_h. The minimal requirement function σ must satisfy is that, for all $x \in U$ and for all $f_h \in F$, $\sigma(x,x,f_h)=1$. Therefore, each pair of objects $(x,y) \in U \times U$ is described by a vector

$$Des_F(x,y)=[\sigma(x,y,f_1), \ldots, \sigma(x,y,f_m)]$$

called *description* of (x,y) in terms of the evaluations of the attributes from F; it represents the available information about similarity between x and y. Obviously, similarity between x and y, $x, y \in U$, can be described in terms of any non-empty subset $E \subseteq F$ and in this case we have

$$Des_E(x,y)=[\sigma(x,y,f_h),\ f_h \in E].$$

With respect to any $E \subseteq F$ we can define the dominance relation D_E on $U \times U$ as follows: for any $x,y,w,z \in U$, (x,y) dominates (w,z) with respect to E (denotation $(x,y)D_E(w,z)$) if for any $f_h \in E$

$$\sigma(x,y,f_h) \geq \sigma(w,z,f_h).$$

Given $E \subseteq F$ and $x,y \in U$, let

$$D_E^+(y,x) = \{w \in U : (w,x)D_E(y,x)\},$$
$$D_E^-(y,x) = \{w \in U : (y,x)D_E(w,x)\}.$$

In the pair (y,x), x is considered as *reference object*, while y can be called *limit object* because it is conditioning the membership of w in $D_E^+(y,x)$ and in $D_E^-(y,x)$.

Let us also consider a fuzzy set X in U, characterized by the membership function $\mu_X : U \rightarrow [0,1]$. For each cutting level $\alpha \in [0,1]$, the following sets can be defined:

$$X^{\geq\alpha}=\{y \in U : \mu_X(y) \geq \alpha\}, \quad X^{>\alpha}=\{y \in U : \mu_X(y) > \alpha\},$$
$$X^{\leq\alpha}=\{y \in U : \mu_X(y) \leq \alpha\}, \quad X^{<\alpha}=\{y \in U : \mu_X(y) < \alpha\}.$$

For each $\alpha \in [0,1]$ and $* \in \{\geq, >\}$, we can define the E-lower approximation of $X^{*\alpha}$, $\underline{E}_\sigma(X^{*\alpha})$, and the E-upper approximation of $X^{*\alpha}$, $\overline{E}_\sigma(X^{*\alpha})$, based on similarity σ with respect to $E \subseteq F$, respectively, as:

$$\underline{E}_\sigma(X^{*\alpha}) = \{(y,x) \in U \times U : D_E^+(y,x) \subseteq X^{*\alpha}\},$$
$$\overline{E}_\sigma(X^{*\alpha}) = \{(y,x) \in U \times U : D_E^-(y,x) \cap X^{*\alpha} \neq \emptyset\}.$$

For the sake of simplicity, in the following we shall consider $\underline{E}_\sigma(X^{\geq\alpha})$ and $\overline{E}_\sigma(X^{\geq\alpha})$. Of course, analogous considerations hold for $\underline{E}_\sigma(X^{>\alpha})$ and $\overline{E}_\sigma(X^{>\alpha})$. Let us remark that the lower approximation of $X^{\geq\alpha}$ contains all the pairs $(y,x) \in U \times U$ such that any object w being similar to x at least as much as y is similar to x w.r.t. all the considered features $E \subseteq F$ also belongs to $X^{\geq\alpha}$. Thus, on the basis of the data from the fuzzy pairwise information base $\boldsymbol{B}$, if the similarity of an object w to x is not smaller than the similarity of y to x w.r.t. all the considered features $E \subseteq F$, then w belongs to $X^{\geq\alpha}$. In other words, in each pair $(y,x) \in \underline{E}_\sigma(X^{\geq\alpha})$, x is a reference object and y is a limit object which belongs "certainly" to set X with credibility at least α; the limit is understood such that all objects w that are similar to x w.r.t. considered features at least as much as y is similar to x, also belong to X with credibility at least α.

Analogously, the upper approximation of $X^{\geq\alpha}$ contains all the pairs $(y,x) \in U \times U$ such that there is at least one object w being similar to x at least as much as y is similar to x w.r.t. all the considered features $E \subseteq F$ which belongs to $X^{\geq\alpha}$. Thus, on the basis of the data from the fuzzy pairwise information base $\boldsymbol{B}$, if the similarity of an object w to x is not smaller than the similarity of y to x w.r.t. all the considered features $E \subseteq F$, then it is possible that w belongs to $X^{\geq\alpha}$. In other words, in each pair $(y,x) \in \overline{E}_\sigma(X^{\geq\alpha})$, x is a reference object and y is a limit object which belongs "possibly" to set X with credibility at least α; the limit is understood such that there is at least one object w that is similar to x w.r.t. considered features at least as much as y is similar to x and has membership in set X with credibility at least α.

For each $\alpha\in[0,1]$ and $* \in \{\leq,<\}$, we can define the E-lower approximation of $X^{*\alpha}$, $\underline{E}_\sigma(X^{*\alpha})$, and the E-upper approximation of $X^{*\alpha}$, $\overline{E}_\sigma(X^{*\alpha})$, based on similarity σ with respect to $E \subseteq F$, respectively, as:

$$\underline{E}_\sigma(X^{*\alpha}) = \{(y,x) \in U \times U : D_E^-(y,x) \subseteq X^{*\alpha}\},$$

$$\overline{E}_\sigma(X^{*\alpha}) = \{(y,x) \in U \times U : D_E^+(y,x) \cap X^{*\alpha} \neq \emptyset\}.$$

For the sake of simplicity, in the following we shall consider $\underline{E}_\sigma(X^{\leq\alpha})$ and $\overline{E}_\sigma(X^{\leq\alpha})$. Of course, analogous considerations hold for $\underline{E}_\sigma(X^{<\alpha})$ and $\overline{E}_\sigma(X^{<\alpha})$. Let us remark that the lower approximation of $X^{\leq\alpha}$ contains all the pairs $(y,x) \in U \times U$ such that any object w being similar to x at most as much as y is similar to x w.r.t. all the considered features $E \subseteq F$ also belongs to $X^{\leq\alpha}$. Thus, on the basis of the data from the fuzzy pairwise information base $\boldsymbol{B}$, if the similarity of an object w to x is not greater than the similarity of y to x with respect to all the considered features $E \subseteq F$, then w belongs to $X^{\leq\alpha}$. In other words, in each pair $(y,x) \in \underline{E}_\sigma(X^{\leq\alpha})$, x is a reference object and y is a limit object which belongs "certainly" to set X with credibility at most α; the limit is understood such that all objects w that are similar to x w.r.t. considered features at most as much as y is similar to x, also belong to X with credibility at most α.

Analogously, the upper approximation of $X^{\leq\alpha}$ contains all the pairs $(y,x) \in U \times U$ such that there is at least one object w being similar to x at most as much as y is similar to x with respect to all the considered features $E \subseteq F$

which belongs to $X^{\leq\alpha}$. Thus, on the basis of the data from the fuzzy pairwise information base $\boldsymbol{B}$, if the similarity of an object w to x is not greater than the similarity of y to x w.r.t. all the considered features $E \subseteq F$, then it is possible that that w belongs to $X^{\leq\alpha}$. In other words, in each pair $(y, x) \in \underline{E}_\sigma(X^{\leq\alpha})$, x is a reference object and y is a limit object which belongs "possibly" to set X with credibility at most α; the limit is understood such that there is at least one object w that is similar to x w.r.t. considered features at most as much as y is similar to x and has membership in set X with credibility at most α.

Let us remark that we can rewrite the rough approximations $\underline{E}_\sigma(X^{\geq\alpha})$, $\overline{E}_\sigma(X^{\geq\alpha})$, $\underline{E}_\sigma(X^{\leq\alpha})$ and $\overline{E}_\sigma(X^{\leq\alpha})$ as follows:

$$\underline{E}_\sigma(X^{\geq\alpha})=\{(y, x) \in U \times U\colon \forall w \in U,\ (w,x)D_E(y,x) \Rightarrow w \in X^{\geq\alpha}\},$$

$$\overline{E}_\sigma(X^{\geq\alpha}) = \{(y, x) \in U \times U\colon \exists w \in U \text{ such that } (w, x)D_E(y, x) \text{ and } w \in X^{\geq\alpha}\},$$

$$\underline{E}_\sigma(X^{\leq\alpha})=\{(y, x) \in U \times U\colon \forall w \in U,\ (y,x)D_E(w,x) \Rightarrow w \in X^{\leq\alpha}\},$$

$$\overline{E}_\sigma(X^{\leq\alpha}) = \{(y, x) \in U \times U : \exists w \in U \text{ such that } (y, x)D_E(w, x) \text{ and } w \in X^{\leq\alpha}\}.$$

This formulation of the rough approximation is concordant with the syntax of the decision rules induced by means of DRSA in a pairwise fuzzy information base. More precisely:

- $\underline{E}_\sigma(X^{\geq\alpha})$ is concordant with decision rules of the type: "if object w is similar to object x w.r.t. feature f_{i1} in degree at least h_{i1} and w.r.t. feature f_{i2} in degree at least h_{i2} and ... and w.r.t. feature f_{im} in degree at least h_{im}, then object w belongs to set X in degree at least α",
- $\overline{E}_\sigma(X^{\geq\alpha})$ is concordant with decision rules of the type: "if object w is similar to object x w.r.t. feature f_{i1} in degree at least h_{i1} and w.r.t. feature f_{i2} in degree at least h_{i2} and ... and w.r.t. feature f_{im} in degree at least h_{im}, then object w could belong to set X in degree at least α",
- $\underline{E}_\sigma(X^{\leq\alpha})$ is concordant with decision rules of the type: "if object w is similar to object x w.r.t. feature f_{i1} in degree at most h_{i1} and w.r.t. feature f_{i2} in degree at most h_{i2} and ... and w.r.t. feature f_{im} in degree at most h_{im}, then object w belongs to set X in degree at most α",
- $\overline{E}_\sigma(X^{\leq\alpha})$ is concordant with decision rules of the type: "if object w is similar to object x w.r.t. feature f_{i1} in degree at most h_{i1} and w.r.t. feature f_{i2} in degree at most h_{i2} and ... and w.r.t. feature f_{im} in degree at least h_{im}, then object w could belong to set X in degree at most α",

where $\{i1, \ldots, im\} = E$ and $h_{i1}, \ldots, h_{im} \in [0, 1]$.

The above definitions of rough approximations and the syntax of decision rules are based on ordinal properties of similarity relations only. In fact, no algebraic operations, such as sum or product, involving cardinal properties of function σ measuring credibility of similarity relations is considered. This is an important characteristic of our approach in comparison with alternative approaches to case-based reasoning.

Let us remark that in the above approximations, even if for two fuzzy sets X and Y we have $X^{\geq\alpha}= Y^{\leq\alpha}$, their approximations are different due to the different directions of cutting the membership function of sets X and Y. Of course, a similar remark holds also for $X^{<\alpha}$ and $Y^{>\alpha}$.

The following theorem states some properties of the rough approximations in a pairwise fuzzy information base.

Theorem. Given a *fuzzy pairwise information base* $\boldsymbol{B}$=< U, F,σ > and a fuzzy set X in U with membership function $\mu_X(\cdot)$, the following properties hold for any $E \subseteq F$:

1. For any α, $0 \leq \alpha \leq 1$,
$\underline{E}_\sigma(X^{\leq\alpha}) \subseteq X^{\leq\alpha}\times X^{\leq\alpha} \subseteq \overline{E}_\sigma(X^{\leq\alpha}), \underline{E}_\sigma(X^{\geq\alpha}) \subseteq X^{\geq\alpha}\times X^{\geq\alpha} \subseteq \overline{E}_\sigma(X^{\geq\alpha}),$
$\underline{E}_\sigma(X^{<\alpha}) \subseteq X^{<\alpha}\times X^{<\alpha} \subseteq \overline{E}_\sigma(X^{<\alpha}), \underline{E}_\sigma(X^{>\alpha}) \subseteq X^{>\alpha}\times X^{>\alpha} \subseteq \overline{E}_\sigma(X^{>\alpha}).$

2. For any α, $0 \leq \alpha\leq 1$,
$\underline{E}_\sigma(X^{\leq\alpha}) = U \times U - \overline{E}_\sigma(X^{>\alpha})$, $\underline{E}_\sigma(X^{\geq\alpha}) = U \times U - \overline{E}_\sigma(X^{<\alpha})$.
3. For any α, β, $0\leq\alpha\leq\beta\leq 1$,

$$\begin{array}{ll} \underline{E}_\sigma(X^{\leq\alpha}) \subseteq \underline{E}_\sigma(X^{\leq\beta}), & \underline{E}_\sigma(X^{<\alpha}) \subseteq \underline{E}_\sigma(X^{<\beta}), \\ \underline{E}_\sigma(X^{\geq\alpha}) \supseteq \underline{E}_\sigma(X^{\geq\beta}), & \underline{E}_\sigma(X^{>\alpha}) \supseteq \underline{E}_\sigma(X^{>\beta}), \\ \overline{E}_\sigma(X^{\leq\alpha}) \subseteq \overline{E}_\sigma(X^{\leq\beta}), & \overline{E}_\sigma(X^{<\alpha}) \subseteq \overline{E}_\sigma(X^{<\beta}), \\ \overline{E}_\sigma(X^{\geq\alpha}) \supseteq \overline{E}_\sigma(X^{\geq\beta}), & \overline{E}_\sigma(X^{>\alpha}) \supseteq \overline{E}_\sigma(X^{>\beta}). \end{array}$$

4. For any $x,y,w,z \in U$ and for any α, $0\leq\alpha\leq 1$,

$$[(y,x)D_E(w,x) \text{ and } (w,x) \in\underline{E}_\sigma(X^{\geq\alpha})]\Rightarrow (y,x) \in\underline{E}_\sigma(X^{\geq\alpha}),$$

$$[(y,x)D_E(w,x) \text{ and } (w,x) \in\underline{E}_\sigma(X^{>\alpha})]\Rightarrow (y,x) \in\underline{E}_\sigma(X^{>\alpha}),$$

$$[(y,x)D_E(w,x) \text{ and } (w,x) \in \overline{E}_\sigma(X^{\geq\alpha})]\Rightarrow (y,x) \in \overline{E}_\sigma\ (X^{\geq\alpha}),$$

$$[(y,x)D_E(w,x) \text{ and } (w,x) \in \overline{E}_\sigma\ (X^{>\alpha})]\Rightarrow (y,x) \in \overline{E}_\sigma\ (X^{>\alpha}),$$

$$[(w,x)D_E(y,x) \text{ and } (w,x) \in\underline{E}_\sigma(X^{\leq\alpha})]\Rightarrow (y,x) \in\underline{E}_\sigma(X^{\leq\alpha}),$$

$$[(w,x)D_E(y,x) \text{ and } (w,x) \in\underline{E}_\sigma(X^{<\alpha})]\Rightarrow (y,x) \in\underline{E}_\sigma(X^{<\alpha}),$$

$$[(w,x)D_E(y,x) \text{ and } (w,x) \in \overline{E}_\sigma\ (X^{\leq\alpha})]\Rightarrow (y,x) \in \overline{E}_\sigma\ (X^{\leq\alpha}),$$

$$[(w,x)D_E(y,x) \text{ and } (w,x) \in \overline{E}_\sigma\ (X^{<\alpha})]\Rightarrow (y,x) \in \overline{E}_\sigma\ (X^{<\alpha}).$$

5. For any $E_1 \subseteq E_2 \subseteq F$ and for any α, $0\leq\alpha\leq 1$,

$$\begin{array}{l} \underline{E}_{1\sigma}(\mathrm{X}^{\leq\alpha}) \subseteq \underline{E}_{2\sigma}(\mathrm{X}^{\leq\alpha}),\ \underline{E}_{1\sigma}(\mathrm{X}^{<\alpha}) \subseteq \underline{E}_{2\sigma}(\mathrm{X}^{<\alpha}), \\ \underline{E}_{1\sigma}(\mathrm{X}^{\geq\alpha}) \subseteq \underline{E}_{2\sigma}(\mathrm{X}^{\geq\alpha}),\ \underline{E}_{1\sigma}(X^{>\alpha}) \subseteq \underline{E}_{2\sigma}(X^{>\alpha}), \\ \overline{E_1}_\sigma(\mathrm{X}^{\leq\alpha}) \supseteq \overline{E_2}_\sigma(X^{\leq\alpha}),\ \overline{E_1}_\sigma(X^{<\alpha}) \supseteq \overline{E_2}_\sigma(X^{<\alpha}), \\ \overline{E_1}_\sigma(X^{\geq\alpha}) \supseteq \overline{E_2}_\sigma(X^{\geq\alpha}),\ \overline{E_1}_\sigma(X^{>\alpha}) \supseteq \overline{E_2}_\sigma(X^{>\alpha}). \end{array}$$

Proof. 1. For all $x, y \in U$,

$$x \in D_E^+(y,x). \quad (i)$$

Thus, $D_E^+(y,x) \subseteq X^{\geq\alpha}$ implies $x \in X^{\geq\alpha}$ and $y \in X^{\geq\alpha}$. For the definition of $\underline{E}_\sigma(X^{\geq\alpha})$, we have that $(y,x) \in \underline{E}_\sigma(X^{\geq\alpha})$ if $D_E^+(y,x) \subseteq X^{\geq\alpha}$, thus we conclude that, $\forall (y,x) \in U \times U$,

$$(y,x) \in \underline{E}_\sigma(X^{\geq\alpha}) \Rightarrow (y,x) \in X^{\geq\alpha} \times X^{\geq\alpha}$$

i.e.

$$\underline{E}_\sigma(X^{\geq\alpha}) \subseteq X^{\geq\alpha} \times X^{\geq\alpha}.$$

Moreover, from (i) we get that for all $(y,x) \in X^{\geq\alpha} \times X^{\geq\alpha}$, $y \in D_E^-(y,x)$. For the definition of $\overline{E}_\sigma(X^{\geq\alpha})$ we have that $(y,x) \in \overline{E}_\sigma(X^{\geq\alpha})$ if $D_E^-(y,x) \cap X^{\geq\alpha} \neq \emptyset$, thus we conclude that, $\forall (y,x) \in U \times U$,

$$(y,x) \in X^{\geq\alpha} \times X^{\geq\alpha} \Rightarrow (y,x) \in \overline{E}_\sigma(X^{\geq\alpha})$$

i.e.

$$X^{\geq\alpha} \times X^{\geq\alpha} \subseteq \overline{E}_\sigma(X^{\geq\alpha}).$$

Consequently, we proved that

$$\underline{E}_\sigma(X^{\geq\alpha}) \subseteq X^{\geq\alpha} \times X^{\geq\alpha} \subseteq \overline{E}_\sigma(X^{\geq\alpha}).$$

Other cases can be proved analogously.

2. Remembering that $X^{<\alpha} = U - X^{\geq\alpha}$ and observing that

$$D_E^+(y,x) \subseteq X^{\geq\alpha} \Leftrightarrow D_E^+(y,x) \cap (U - X^{\geq\alpha}) = \emptyset \Leftrightarrow D_E^+(y,x) \cap X^{<\alpha} = \emptyset$$

we get

$$\underline{E}_\sigma(X^{\geq\alpha}) = \{(y,x) \in U \times U : D_E^+(y,x) \subseteq X^{\geq\alpha}\} =$$
$$= U \times U - \{(y,x) \in U \times U : D_E^+(y,x) \cap X^{<\alpha} \neq \emptyset\} =$$
$$= U \times U - \overline{E}_\sigma(X^{<\alpha}).$$

Analogous proof holds for $\underline{E}_\sigma(X^{\geq\alpha}) = U \times U - \overline{E}_\sigma(X^{<\alpha})$.

3. Let us observe that for any α, β, $0{\leq}\alpha{\leq}\beta{\leq}1$

$$X^{\geq\alpha} = \{x \in U : \mu(x) \geq \alpha\} \supseteq \{x \in U : \mu(x) \geq \beta\} = X^{\geq\beta}.$$

Taking this into account, we get

$$\{(y,x) \in U \times U : D_E^+(y,x) \subseteq X^{\geq\alpha}\} \subseteq \{(y,x) \in U \times U : D_E^+(y,x) \subseteq X^{\geq\beta}\}$$

i.e.

$$\underline{E}_\sigma(X^{\geq\alpha}) \subseteq \underline{E}_\sigma(X^{\geq\beta}).$$

Moreover, we also obtain

$$\{(y,x) \in U \times U : D_E^-(y,x) \cap X^{\geq\alpha} \neq \emptyset\} \subseteq \{(y,x) \in U \times U : D_E^+(y,x) \cap X^{\geq\beta} \neq \emptyset\}$$

i.e.

$$\overline{E}_\sigma(X^{\geq\alpha}) \subseteq \overline{E}_\sigma(X^{\geq\beta}).$$

Other cases can be proved analogously.

4. Let us observe that for the transitivity of D_E, for any x,y,w,$z \in U$ and for any $E \subseteq F$

$$[(z,x)D_E(y,x) \text{ and } (y,x)D_E(w,x) \Rightarrow (z,x)D_E(w,x)]$$
$$\Leftrightarrow$$
$$[(z,x) \in D_E^+(y,x) \text{ and } (y,x)D_E(w,x) \Rightarrow (z,x) \in D_E^+(w,x)]$$
$$\Leftrightarrow$$
$$[(y,x)D_E(w,x) \Rightarrow D_E^+(y,x) \subseteq D_E^+(w,x)].$$

From this we get that if $(y,x)D_E(w,x)$, then

$$D_E^+(w,x) \subseteq X^{\geq\alpha} \Rightarrow D_E^+(y,x) \subseteq X^{\geq\alpha}$$

i.e.

$$[(y,x)D_E(w,x) \text{ and } (w,x) \in \underline{E}_\sigma(X^{\geq\alpha})] \Rightarrow (y,x) \in \underline{E}_\sigma(X^{\geq\alpha}).$$

Other cases can be proved analogously.

5. For any $E_1 \subseteq E_2 \subseteq F$ and for any $x, y, w, z \in U$

$$(x,y)D_{E_2}(w,z) \Rightarrow (x,y)D_{E_1}(w,z)$$

and thus

$$D_{E_1}^+(x,y) \supseteq D_{E_2}^+(x,y) \text{ and } D_{E_1}^-(x,y) \supseteq D_{E_2}^-(x,y).$$

From this we get that for all α, $0 \leq \alpha \leq 1$,

$$D_{E_1}^+(y,x) \subseteq X^{\geq\alpha} \Rightarrow D_{E_2}^+(y,x) \subseteq X^{\geq\alpha}$$

and

$$D_{E_2}^-(y,x) \cap X^{\geq\alpha} \neq \emptyset \Rightarrow D_{E_1}^-(y,x) \subseteq X^{\geq\alpha} \neq \emptyset,$$

which give, respectively,

$$\underline{E}_{1\sigma}(X^{\geq\alpha}) \subseteq \underline{E}_{2\sigma}(X^{\geq\alpha}) \text{ and } \overline{E}_{1\sigma}(X^{\geq\alpha}) \supseteq \overline{E}_{2\sigma}(X^{\geq\alpha}).$$

Other cases can be proved analogously. $\diamond$

3 Conclusions

We presented a model of case-based reasoning using Dominance-based Rough Set Approach (DRSA). This model is based only on ordinal properties of similarity relations and membership functions of fuzzy sets. Moreover, we did not impose any specific aggregation functional based on specific axioms (see for example [4]), but we considered a set of decision rules based on the very general monotonicity property of comprehensive similarity with respect to similarity of single features. From this viewpoint our approach to case-based reasoning is as much “neutral” and “objective” as possible and it is very little “invasive” comparing to many other existing approaches. Future research on rough set approach to case-based reasoning can be focused on

– comparison of our approach with other case-based reasoning methodologies and
– the use of our approach for extension of other concepts, results and methodologies of rough set theory.

With respect to comparison of our approach with other case-based reasoning methodologies, an important future research concerns axiomatic considerations. As observed by Gilboa and Schmeidler [4] the interest of axiomatic consideration can be summarized in the following points:

1) meta-scientific reasons: axiomatization provides a link between theoretical terms and observable terms in order to rend the latter meaningful;
2) descriptive reasons: it supplies the basis for testing the empirical validity of the theory because axioms permit to conceive experiments able to falsify the theory rendering the theory falsifiable as requested by Popper [15];
3) normative reasons: a simple set of axioms is often more understandable than the mathematical formulation of the theory and from this viewpoint can be the basis for a discussion with a decision maker about acceptance or rejection of the theory.

With respect to our approach to case-based reasoning, the axiomatic considerations have the further merits of permitting a comparison with the axiomatization of Gilboa and Schmeidler [4] and of pointing out the fact that only monotonic properties of similarity measures are considered.

The research field seems very promising also with respect to rough set theory and we envisage interesting developments with respect to three following issues:

1) generalizations of other rough set fundamental concepts such as reducts and core [13];
2) algebraic properties of the proposed rough approximations (for a general introduction of algebraic properties of classical rough set approach see [14]);
3) application of the absolute and relative rough membership concept (see [8]) in a generalized variable precision model based on the proposed rough approximations in order to admit decision rules with a limited number of counterexamples, which is particularly useful when dealing with large data sets.

Acknowledgements. The research of the first two authors has been supported by the Italian Ministry of Education, University and Scientific Research (MIUR). The third author wishes to acknowledge financial support from the Polish Ministry of Education and Science.

References

1. Dubois, D., Prade, H., Gradual inference rules in approximate reasoning, *Information Sciences*, 61 (1992) 103-122
2. Dubois, D., Prade, H., Putting rough sets and fuzzy sets together, in: R. Slowinski (ed.), *Intelligent Decision Support: Handbook of Applications and Advances of the Sets Theory*, Kluwer, Dordrecht, 1992, pp. 203-232
3. Dubois, D., Prade, H., Esteva, F., Garcia, P., Godo, L., Lopez de Mantara, R., Fuzzy Set Modelling in Case-based Reasoning, *International Journal of Intelligent Systems*, 13 (1998) 345-373
4. Gilboa, I., Schmeidler, D., *A Theory of Case-Based Decisions*, Cambridge University Press, 2001
5. Greco, S., Matarazzo, B., Slowinski R., Rough set theory for multicriteria decision analysis, *European Journal of Operational Research*, 129 (2001) 1-47
6. Greco, S., Matarazzo, B., Slowinski R., Decision rule approach, in: J. Figueira, S. Greco and M. Ehrgott (eds.) *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer, Berlin, 2005, pp. 507-563
7. S. Greco, B. Matarazzo, R. Slowinski. Generalizing Rough Set Theory through Dominance-based Rough Set Approach. In D. Slezak, J. Yao, J. Peters, W. Ziarko, X. Hu (eds.), *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, LNAI 3642, Springer, Berlin, 2005, pp. 1-11
8. Greco, S., Matarazzo, B., Slowinski R., Rough membership and Bayesian confirmation measures for parametrized rough sets, In D. Slezak, J. Yao, J. Peters, W. Ziarko, X. Hu (eds.), *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, LNAI 3641, Springer, Berlin, 2005, pp. 314-324
9. Hume, D., *An Enquiry concerning Human Understanding*, Oxford, Clarendon Press, 1748
10. Kolodner, J., *Case-Based Reasoning*, Morgan Kaufmann, San Mateo, CA, 1993
11. Leake, D. B., CBR in Context: The Present and Future, in: D. Leake (ed.) *Case-Based Reasoning: Experiences, Lessons, and Future Directions*, AAAI Press/MIT Press, Menlo Park, 1996, pp. 1-30
12. Pawlak, Z., Rough sets, *International Journal of Computer and Information Sciences*, 11 (1982) 341-356
13. Pawlak, Z., *Rough Sets*, Kluwer, Dordrecht, 1991
14. Polkowski, L., *Rough Sets: Mathematical Foundations*, Physica-Verlag, 2002
15. Popper, K. R., *The Logic of Scientific Discovery*, Hutchinson and Co., London, 1958
16. Slowinski, R., Rough set processing of fuzzy information, in: T. Y. Lin, A. Wildberger (eds.), *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledege Discovery*, Simulation Councils, Inc., San Diego, CA, 1995, pp. 142-145
17. Wittgenstein, L., *Tractatus Logico-Philosophicus*, Routledge and Kegan Paul, London, 1922; fifth impression 1951
18. Zadeh, L., Fuzzy Set, *Information and Control*, 8 (1965) 338-353