Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Approximate Reasoning
41 (2006) 179–211
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

Fuzzy rough sets and multiple-premise gradual decision rules

Salvatore Greco [a,*], Masahiro Inuiguchi [b], Roman Slowinski [c,d]

[a] *Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy*
[b] *Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Toyonaka, Osaka 560-8531, Japan*
[c] *Institute of Computing Science, Poznan University of Technology, 60-965 Poznan, Poland*
[d] *Institute for Systems Research, Polish Academy of Sciences, 01-447 Warsaw, Poland*

Received 1 January 2004; received in revised form 1 May 2005; accepted 1 June 2005
Available online 16 September 2005

Abstract

We propose a new fuzzy rough set approach which, differently from most known fuzzy set extensions of rough set theory, does not use any fuzzy logical connectives (t-norm, t-conorm, fuzzy implication). As there is no rationale for a particular choice of these connectives, avoiding this choice permits to reduce the part of arbitrary in the fuzzy rough approximation. Another advantage of the new approach is that it is based on the ordinal properties of fuzzy membership degrees only. The concepts of fuzzy lower and upper approximations are thus proposed, creating a base for induction of fuzzy decision rules having syntax and semantics of gradual rules. The proposed approach to rule induction is also interesting from the viewpoint of philosophy supporting data mining and knowledge discovery, because it is concordant with the method of concomitant variations by John Stuart Mill. The decision rules are induced from lower and upper approximations defined for positive and negative relationships between credibility degrees of multiple premises, on one hand, and conclusion, on the other hand.

Keywords: Rough sets; Fuzzy sets; Decision rules; Gradual rules; Credibility

* Corresponding author.
E-mail addresses: salgreco@unict.it (S. Greco), inuiguti@sys.es.osaka-u.ac.jp (M. Inuiguchi), roman.slowisnki@cs.put.poznan.pl (R. Slowinski).

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2005.06.014

1. Introduction

It has been acknowledged by different studies that fuzzy set theory and rough set theory are complementary because of handling different kinds of uncertainty. Fuzzy sets deal with possibilistic uncertainty, connected with imprecision of states, perceptions and preferences [4]. Rough sets deal, in turn, with uncertainty following from ambiguity of information [18]. The two types of uncertainty can be encountered together in real-life problems. For this reason, many approaches have been proposed to combine fuzzy sets with rough sets (see, e.g., [6]).

Let us remember that fuzzy sets [21] are based on the idea that, given a universe U, the membership of $y \in U$ in a set X from U admits a graduality represented by means of function $\mu_X: U \to [0,1]$ such that $\mu_X(y) = 0$ means non-membership, $\mu_X(y) = 1$ means full membership, and for all intermediate values the greater $\mu_X(y)$, the more credible the membership of y in X. An analogous idea of graduality is introduced in fuzzy logic with respect to the truth value $v(p)$ of a proposition p such that $v(p) = 0$ means that p is definitely false, $v(p) = 1$ that p is definitely true and for all intermediate values the greater $v(p)$, the more credible the truth of p. In the context of fuzzy sets, fuzzy connectives, being functions from $[0,1] \times [0,1]$ to $[0,1]$, represent conjunction (t-norm), disjunction (t-conorm) or implication (fuzzy implications such as S-implication or R-implication) (see, e.g., [13]).

Let us also remember that rough set theory [17,18] is based 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.

The main preoccupation in almost all the studies conjugating rough sets with fuzzy sets was related to a fuzzy extension of Pawlak's definition of lower and upper approximations using fuzzy connectives. In fact, there is no rule for the choice of the "right" connective, so this choice is always arbitrary to some extent.

Another drawback of fuzzy extensions of rough sets involving fuzzy connectives is that they are based on cardinal properties of membership degrees. In consequence, the result of these extensions is sensitive to order preserving transformation of membership degrees. For example, consider the t-conorm of Lukasiewicz as fuzzy connective; it may be used in the definition of both fuzzy lower approximation (to build fuzzy implication) and fuzzy upper approximation (as a fuzzy counterpart of a union). The t-conorm of Lukasiewicz is defined as

$$T^*(\alpha, \beta) = \min\{\alpha + \beta, 1\}.$$

$T^*(\alpha, \beta)$ can be interpreted as follows. If $\alpha = \mu_X(z)$ represents the membership of $z \in U$ in set X and $\beta = \mu_Y(z)$ represents the membership of z in set Y, then $T^*(\alpha, \beta)$ expresses the membership of z in set $X \cup Y$. Given two fuzzy propositions p and q, putting $v(p) = \alpha$ and $v(q) = \beta$, $T^*(\alpha, \beta)$ can be interpreted also as $v(p \vee q)$, the truth value of the proposition $p \vee q$.

Let us consider the following values of arguments:

$$\alpha = 0.5, \quad \beta = 0.3, \quad \gamma = 0.2, \quad \delta = 0.1,$$

and their order preserving transformation:

$$\alpha' = 0.4, \quad \beta' = 0.3, \quad \gamma' = 0.2, \quad \delta' = 0.05.$$

The values of the t-conorm are in the two cases as follows:

$$T^*(\alpha, \delta) = 0.6, \quad T^*(\beta, \gamma) = 0.5, \quad T^*(\alpha', \delta') = 0.45, \quad T^*(\beta', \gamma') = 0.5.$$

One can see that the order of the results has changed after the order preserving transformation of the arguments. This means that the Lukasiewicz t-conorm takes into account not only the ordinal properties of the membership degrees, but also their cardinal properties. A natural question arises: is it reasonable to expect from membership degree a cardinal content instead of ordinal only? Or, in other words, is it realistic to claim that a human is able to say in a meaningful way not only that

(a) "object x belongs to fuzzy set X more likely than object y" (or "proposition p is more credible than proposition q")

but even something like

(b) "object x belongs to fuzzy set X two times more likely than object y" (or "proposition p is two times more credible than proposition q")?

We claim that it is safer to consider information of type (a), because information of type (b) is rather meaningless for a human (see [14]).

The above doubt about the cardinal content of the fuzzy membership degree shows the need for methodologies which consider the imprecision in perception typical for fuzzy sets but avoid as much as possible meaningless transformation of information through fuzzy connectives.

The approach we propose for fuzzy extension of rough sets takes into account the above request. It avoids arbitrary choice of fuzzy connectives and not meaningful operations on membership degrees. Our approach belongs to the minority of fuzzy extensions of the rough set concept that do not involve fuzzy connectives and cardinal interpretation of membership degrees. Within this minority, it is related to the approach of Nakamura and Gao [16] using α-cuts on fuzzy similarity relation between objects.

We propose a methodology of fuzzy rough approximation that infers the most cautious conclusion from available imprecise information. In particular, we observe that any approximation of knowledge about Y using knowledge about X is based on positive or negative relationships between premises and conclusions, i.e.,

(i) "the more x is X, the more it is Y" (positive relationship),
(ii) "the more x is X, the less it is Y" (negative relationship).

The following simple relationships illustrate (i) and (ii): "the larger the market share of a company, the greater its profit" (positive relationship) and "the greater the debt of a company, the smaller its profit" (negative relationship). These relationships have been already considered within fuzzy set theory under the name of *gradual decision rules* [5]. Recently, Greco et al. [7,8] proposed an approach for induction of gradual decision rules relating knowledge about X and knowledge about Y, represented by a single premise and a single conclusion, respectively. It handles ambiguity of information through fuzzy rough approximations. In this paper, we want to extend this approach to induction of gradual decision rules having multiple premises representing knowledge about X. Examples of these decision rules are: "if a car is speedy with credibility at least 0.8 and it has high fuel consumption with credibility at most 0.7, then it is a good car with a credibility at least 0.9" and "if a car is speedy with credibility at most 0.5 and it has high fuel consumption with credibility at least 0.8, then it is a good car with a credibility at most 0.6".

Remark that the syntax of gradual decision rules is based on monotonic relationship that can also be found in dominance-based decision rules induced from preference-ordered data. From this point of view, the fuzzy rough approximation proposed in this article is related to the dominance-based rough set approach [9–11].

For the reason of greater generality, one could drop the assumption of the monotonic relationship between premise and conclusion in gradual rules. For example, the gradual rule "the greater the temperature the better the weather" is true in some range of temperature only (say, up to 25 °C). In such cases, however, one can split the domain of the premise into sub-intervals, in which the monotonicity still holds, and represents the regularities observed in these sub-intervals by gradual rules. For example, we can split the range of the temperature into two open subintervals, under 25 °C and over 25 °C, obtaining the two gradual rules: "the greater the temperature the better the weather", which is valid in the first interval, and "the smaller the temperature the better the weather", which is valid in the second interval. Therefore, the concept of monotonicity in gradual rules is intrinsic to the idea of induction whose aim is to represent regularities according to the simplest law (see, Proposition 6.363 in [20]: "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 the monotonicity.

The above proposition of Wittgenstein is borrowed from the paper by Aragones et al. [1] on a similar subject. Remark, however, that these authors consider rules with non-monotonic relationships between premise and conclusion, and, moreover, their rule induction procedure is based on a cardinal concept of the credibility of information.

The model of rule induction proposed in this paper is interesting also from the viewpoint of data mining, knowledge discovery, machine learning and their philosophical background [3,2,19]. In fact, applications of data mining, knowledge discovery and machine learning requires a proper theory related to such questions as

- Can the whole process of knowledge discovery be automated or reduced to pure-logics?
- In what degree pieces of evidence found in data support a hypothesis? [12]
- How to choose an inductive strategy appropriate for the task one is facing?
- What is the relationship between machine learning and philosophy of science?
- "Is machine learning experimental philosophy of science?" [2]

In this paper, we focus on the kind of discoveries permitted by our methodology. The rule induction approach we are proposing is concordant with the method of concomitant variation proposed by John Stuart Mill. The general formulation of this method is the following: "Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or it is connected with it through some causation" [15]. In simpler words, the method of concomitant variation searches for positive or negative relations between magnitudes of considered variables. Mill's example concerned the tides and the position of the moon. In the above example of decision rules concerning evaluation of a car, the variations in evaluation of the car are positively related with variations in its speed and negatively related with variations in its fuel consumption. Cornish and Elliman [3] note that within current practice of data mining, the method of concomitant variation is the one which receives the least attention among the other methods proposed by Mill (method of agreement, method of difference, method of indirect difference and method of residues). However Cornish and Elliman [3] observe also that the method of concomitant variation "is believed to have the greatest potential for the discovery of knowledge, in such areas as biology and biomedicine, as it addresses parameters which are forever present and inseparable".

The plan of the article is the following. In Section 2, we are defining the syntax and semantics of considered gradual decision rules; we also show how they represent positive and negative relationships between fuzzy sets corresponding to multiple premises and to conclusion of a decision rule. In Section 3, we are introducing fuzzy rough approximations consistent with the considered gradual decision rules. Section 4 deals with rule induction based on rough approximations. In Section 5, we introduce fuzzy rough modus ponens and fuzzy rough modus tollens based on gradual decision rules. Section 6 is grouping conclusions and remarks on further research directions.

2. Gradual decision rules with positively or negatively related premises and conclusion

Let us consider condition attributes $X_1, \ldots, X_n$, related with decision attribute Y. More precisely, we shall denote by $X_i^{\uparrow}$ a fuzzy value of attribute X_i positively related with decision attribute Y, and by $X_i^{\downarrow}$, a fuzzy value of attribute X_i negatively related with decision attribute Y. We aim to obtain gradual decision rules of the following types:

- *lower-approximation rules* (L-rule): "if
  - $x \in X_{i1}^{\uparrow}$ with credibility $C(X_{i1}^{\uparrow}) \geqslant \alpha_{i1}$, $x \in X_{i2}^{\uparrow}$ with credibility $C(X_{i2}^{\uparrow}) \geqslant \alpha_{i2}, \ldots,$ and $x \in X_{ir}^{\uparrow}$ with credibility $C(X_{ir}^{\uparrow}) \geqslant \alpha_{ir}$, and
  - $x \in X_{j1}^{\downarrow}$ with credibility $C(X_{j1}^{\downarrow}) \leqslant \alpha_{j1}$, $x \in X_{j2}^{\downarrow}$ with credibility $C(X_{j2}^{\downarrow}) \leqslant \alpha_{j2}, \ldots,$ and $x \in X_{js}^{\downarrow}$ with credibility $C(X_{js}^{\downarrow}) \leqslant \alpha_{js}$,

  then decision $x \in Y$ has credibility $C(Y) \geqslant \beta$",
- *upper-approximation rule* (U-rule): "if
  - $x \in X_{i1}^{\uparrow}$ with credibility $C(X_{i1}^{\uparrow}) \leqslant \alpha_{i1}$, $x \in X_{i2}^{\uparrow}$ with credibility $C(X_{i2}^{\uparrow}) \leqslant \alpha_{i2}, \ldots,$ and $x \in X_{ir}^{\uparrow}$ with credibility $C(X_{ir}^{\uparrow}) \leqslant \alpha_{ir}$, and
  - $x \in X_{j1}^{\downarrow}$ with credibility $C(X_{j1}^{\downarrow}) \geqslant \alpha_{j1}$, $x \in X_{j2}^{\downarrow}$ with credibility $C(X_{j2}^{\downarrow}) \geqslant \alpha_{j2}, \ldots,$ and $x \in X_{js}^{\downarrow}$ with credibility $C(X_{js}^{\downarrow}) \geqslant \alpha_{js}$,

  then decision $x \in Y$ has credibility $C(Y) \leqslant \beta$".

The above decision rules will be represented by $(r+s+2)$-tuples $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g\rangle$, respectively, where $f{:}[0,1]^{r+s}\to[0,1]$ and g: $[0,1]^{r+s}\to[0,1]$ are functions relating the credibility of membership in $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$, $X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$ with the credibility of membership in Y, in lower- and upper-approximation rules, respectively. More precisely, functions f and g permit to rewrite the conclusion part of above decision rules as follows:

- L-rule: "then decision $x\in Y$ has credibility $C(Y)\geqslant\beta=f(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$";
- U-rule: "then decision $x\in Y$ has credibility $C(Y)\leqslant\beta=g(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$".

If we have sufficient information about the lower boundary and upper boundary of credibility $C(Y)$, functions f and g would be obtained as functions which are monotonically non decreasing with $\alpha_{i1},\ldots,\alpha_{ir}$ and monotonically non increasing with $\alpha_{j1},\ldots,\alpha_{js}$. Otherwise, we cannot expect such monotonicity properties of functions f and g. Namely, under some partial information about those boundaries, functions f and g cannot be monotonically non decreasing with $\alpha_{i1},\ldots,\alpha_{ir}$ and monotonically non increasing with $\alpha_{j1},\ldots,\alpha_{js}$. In what follows, we assume only some partial information about the lower boundary and upper boundary of credibility $C(Y)$ so that functions f and g are not always monotonically non decreasing with $\alpha_{i1},\ldots,\alpha_{ir}$ and monotonically non increasing with $\alpha_{j1},\ldots,\alpha_{js}$.

Given an L-rule $LR=\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and an object z, taking into account that function f is not necessarily monotonic, we define the lower boundary of membership of z in Y with respect to LR, denoted by $C(z,LR,Y)$, as follows:

$$C(z,LR,Y)=\inf_{\alpha\in E^{+}(z)}f(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1}\ldots,\alpha_{js}),$$

where

$$E^{+}(z)=\Big\{\boldsymbol{\alpha}=(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})\in[0,1]^{r+s}:\alpha_h\geqslant\mu_{X_h}(z)\text{ for each } X_h\in\{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}\text{ and }\alpha_h\leqslant\mu_{X_h}(z)\text{ for each }X_h\in\{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}\Big\}.$$

Namely, with $f(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$, we revise the lower boundary by using the knowledge that credibility $C(Y)$ is monotonically non decreasing with credibilities $C(X_{i1}^{\uparrow}),\ldots,C(X_{ir}^{\uparrow})$ and monotonically non increasing with credibilities $C(X_{j1}^{\downarrow}),\ldots,C(X_{js}^{\downarrow})$. Note that this modification does not change the conclusion, i.e. $C(z,LR,Y)=f(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$, when function f is monotonically non decreasing with $\alpha_{i1},\ldots,\alpha_{ir}$ and monotonically non increasing with $\alpha_{j1},\ldots,\alpha_{js}$.

Intuitively, the lower boundary represents the lowest credibility we can assign to membership of object z in Y on the basis of an L-rule LR, given the hypothesis about the positive relationships with respect to membership in $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$ and the negative relationship with respect to membership in $X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$. The following example illustrates this idea.

Example 1 (*part* 1). Let us suppose that in order to evaluate the membership of a car in a set of good cars $Y_{\text{good\_cars}}$, we should take into account the membership of the car in a set of speedy cars $X^{\uparrow}_{\text{speedy\_cars}}$, and the membership of the cars in a set of expensive cars $X^{\downarrow}_{\text{expensive\_cars}}$. Of course the membership in set $Y_{\text{good\_cars}}$ of good cars is positively related with the membership in set $X^{\uparrow}_{\text{speedy\_cars}}$ of speedy cars and negatively related with the membership in set $X^{\downarrow}_{\text{expensive\_cars}}$ of expensive cars. Given a car z, $\mu_{\text{speedy\_cars}}(z)$ denotes the degree of membership of z in set $X^{\uparrow}_{\text{speedy\_cars}}$, $\mu_{\text{expensive\_cars}}(z)$ denotes the degree of membership of z in set $X^{\downarrow}_{\text{expensive\_car}}$ and $\mu_{\text{good\_cars}}(z)$ denotes the degree of membership of z in set $Y_{\text{good\_cars}}$.

Let us suppose that the membership of each car z in set $Y_{\text{good\_cars}}$ is based on L-rule $LR = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f\rangle$ where $f: [0,1]^2 \rightarrow [0,1]$ is defined as follows:

$$f(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.50, \\ 0.66 & \text{if } -0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0, \\ 0.33 & \text{if } 0 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.50, \\ 1 & \text{if } 0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

Let us observe that function f is not monotonically non decreasing with $\mu_{\text{speedy\_cars}}(z)$ and monotonically non increasing with $\mu_{\text{expensive\_cars}}(z)$ since we assume that only partial information is available.

Now, let us consider a car z such that $\mu_{\text{speedy\_cars}}(z) = 0.4$ and $\mu_{\text{expensive\_cars}}(z) = 0.8$. We have

$$C(z, LR, Y_{\text{good\_cars}}) = \inf_{\boldsymbol{\alpha} \in E^+(z)} f(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = 0.33,$$

where

$$E^+(z) = \Big\{ \boldsymbol{\alpha} = (\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2 : \alpha_{\text{speedy\_cars}} \geqslant \mu_{\text{speedy\_cars}}(z) = 0.4 \text{ and } \alpha_{\text{expensive\_cars}} \leqslant \mu_{\text{expensive\_cars}}(z) = 0.8 \Big\}.$$

Let us remark that

$$f[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)] = 0.66 > 0.33 = C(z, LR, Y_{\text{good\_cars}}).$$

This is explained by the principle of coherence with respect to the sign of relationships between condition attributes $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$, on one hand, and decision attribute $Y_{\text{good\_cars}}$, on the other hand. According to this principle, membership of car z in set $Y_{\text{good\_cars}}$ should not be greater than that of car w such that $\mu_{\text{speedy\_cars}}(w) \geqslant \mu_{\text{speedy\_cars}}(z)$ and $\mu_{\text{expensive\_cars}}(w) \leqslant \mu_{\text{expensive\_cars}}(z)$. Remark that, for example, in case of car w, for which $\mu_{\text{speedy\_cars}}(w) = 0.7$ and $\mu_{\text{expensive\_cars}}(w) = 0.5$, the above function f suggests a membership degree of w in $Y_{\text{good\_cars}}$ equal to $\mu_{\text{good\_cars}}(w) = f[\mu_{\text{speedy\_cars}}(w), \mu_{\text{expensive\_cars}}(w)] = 0.33$. Therefore, we should also have $\mu_{\text{good\_cars}}(z) \leqslant 0.33$. In this perspective, $C(z, LR, Y_{\text{good\_cars}})$ represents a prudent evaluation of $\mu_{\text{good\_cars}}(z)$ in such a way that for all cars w for which

$$\mu_{\text{speedy\_cars}}(w) \geqslant \mu_{\text{speedy\_cars}}(z) \quad \text{and} \quad \mu_{\text{expensive\_cars}}(w) \leqslant \mu_{\text{expensive\_cars}}(z), \tag{i}$$

we have

$$\mu_{\text{good\_cars}}(w) \geqslant \mu_{\text{good\_cars}}(z). \tag{ii}$$

More precisely, $C(z, LR, Y_{\text{good\_cars}})$ is the maximal value one can assign to $\mu_{\text{good\_cars}}(z)$ in such a way that (i) and (ii) hold.

For the sake of completeness let us observe that

$$C(z, LR, Y_{\text{good\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant -0.5, \\ 0.33 & \text{if } -0.5 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 0.5, \\ 1 & \text{if } 0.5 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 1. \end{cases}$$

Analogously, given an U-rule $UR = \langle X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y, g \rangle$ and an object z, we define the upper boundary of membership of z to Y with respect to UR, denoted by $C(z, UR, Y)$, as follows:

$$C(z, UR, Y) = \sup_{\alpha \in E^{-}(z)} g(\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js}),$$

where

$$E^{-}(z) = \Big\{ \boldsymbol{\alpha} = (\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js}) \in [0,1]^{r+s} : \alpha_h \leqslant \mu_{X_h}(z) \text{ for each}$$
$$X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\} \text{ and } \alpha_h \geqslant \mu_{X_h}(z) \text{ for each } X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\} \Big\}.$$

Namely, with $g(\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})$, we modify the upper boundary by using the knowledge that credibility $C(Y)$ is monotonically non decreasing with credibilities $C(X_{i1}^{\uparrow}), \ldots, C(X_{ir}^{\uparrow})$ and monotonically non increasing with credibilities $C(X_{j1}^{\downarrow}), \ldots, C(X_{js}^{\downarrow})$. Note that this modification does not change the conclusion, i.e. $C(z, UR, Y) = g(\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})$, when function g is monotonically non decreasing with $\alpha_{i1}, \ldots, \alpha_{ir}$ and monotonically non increasing with $\alpha_{j1}, \ldots, \alpha_{js}$.

Intuitively, the upper boundary represents the highest credibility we can assign to membership of object z in Y on the basis of an U-rule UR, given the hypothesis about the positive relationships with respect to membership in $X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}$ and the negative relationship with respect to membership in $X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}$. The following continuation of Example 1 illustrates this idea.

Example 1 (*part* 2). Let us suppose that the membership of each car in set $Y_{\text{good\_cars}}$ is based on U-rule $UR = \langle X_{\text{speedy\_cars}}^{\uparrow}, X_{\text{expensive\_cars}}^{\downarrow}, Y_{\text{good\_cars}}, g \rangle$ where $g: [0,1]^2 \to [0,1]$ has the same definition as above function f, i.e.,

$$g(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = f(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$$

for all $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$. Let us also suppose that we want to evaluate the membership in set $Y_{\text{good\_cars}}$ of a car u such that $\mu_{\text{speedy\_cars}}(u) = 0.8$ and $\mu_{\text{expensive\_cars}}(u) = 0.4$. We have

$$C(u, UR, Y_{\text{good\_cars}}) = \sup_{\alpha \in E^{-}(u)} g(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = 0.66,$$

where

$$E^{-}(u) = \Big\{ \boldsymbol{\alpha} = (\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2 : \alpha_{\text{speedy\_cars}} \leqslant \mu_{\text{speedy\_cars}}(u) = 0.8 \text{ and}$$
$$\alpha_{\text{expensive\_cars}} \geqslant \mu_{\text{expensive\_cars}}(u) = 0.4 \Big\}.$$

Let us remark that

$$f[\mu_{\text{speedy\_cars}}(u), \mu_{\text{expensive\_cars}}(u)] = 0.33 < 0.66 = C(u, UR, Y_{\text{good\_cars}}).$$

This is explained by the principle of coherence with respect to the sign of relationships between condition attributes, that are membership in $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$, on one hand, and decision attribute, that is membership in $Y_{\text{good\_cars}}$, on the other hand. According to this principle, membership in set $Y_{\text{good\_cars}}$ of car u should not be smaller than that of car v, such that $\mu_{\text{speedy\_cars}}(v) \leqslant \mu_{\text{speedy\_cars}}(u)$ and $\mu_{\text{expensive\_cars}}(v) \geqslant \mu_{\text{expensive\_cars}}(u)$. Remark that, for example, in case of car v, for which $\mu_{\text{speedy\_cars}}(v) = 0.5$ and $\mu_{\text{expensive\_cars}}(v) = 0.7$, the above function g suggests a membership degree of v in $Y_{\text{good\_cars}}$ equal to $\mu_{\text{good\_cars}}(v) = g[\mu_{\text{speedy\_cars}}(v), \mu_{\text{expensive\_cars}}(v)] = 0.66$. Therefore, we should have also $\mu_{\text{good\_cars}}(u) \geqslant 0.66$. In this perspective, $C(u, UR, Y_{\text{good\_cars}})$ represents a possible and "optimistic" evaluation of $\mu_{\text{good\_cars}}(u)$ in such a way that for all cars v for which

$$\mu_{\text{speedy\_cars}}(v) \leqslant \mu_{\text{speedy\_cars}}(u) \quad \text{and} \quad \mu_{\text{expensive\_cars}}(v) \geqslant \mu_{\text{expensive\_cars}}(u), \tag{iii}$$

we have

$$\mu_{\text{good\_cars}}(v) \leqslant \mu_{\text{good\_cars}}(u). \tag{iv}$$

More precisely $C(u, LR, Y_{\text{good\_cars}})$ is the minimal value one can assign to $\mu_{\text{good\_cars}}(u)$ in such a way that (iii) and (iv) hold.

For the sake of completeness let us observe that

$$C(z, UR, Y_{\text{good\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant -0.5, \\ 0.66 & \text{if } -0.5 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 0.5, \\ 1 & \text{if } 0.5 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 1. \end{cases}$$

Two L-rules $LR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f\rangle$ and $LR' = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f'\rangle$ are equivalent if for all possible objects z we have that $C(z, LR, Y) = C(z, LR', Y)$.

Two U-rules $UR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g\rangle$ and $UR' = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g'\rangle$ are equivalent if for all possible objects z we have that $C(z, UR, Y) = C(z, UR', Y)$.

Example 1 (*part* 3). Let us suppose that we want to evaluate the membership of each car in set $Y_{\text{good\_cars}}$ on the basis of L-rule $LR' = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f'\rangle$, where $f' : [0,1]^2 \to [0,1]$ is defined as follows:

$$f'[\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}] = \begin{cases} 0 & -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.50, \\ 0.50 & -0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0, \\ 0.33 & 0 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.50, \\ 1 & 0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

We have that $C(z, LR, Y) = C(z, LR', Y)$ and, therefore, L-rules LR and LR' are equivalent. The reason is that f and f' differ in a part of their domain only, where the monotonicity of membership in $Y_{\text{good\_cars}}$ with respect to membership in $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ is not satisfied. In fact, $f'[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)]$ differs from $f[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)]$ in case

$$-0.50 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 0, \tag{v}$$

where $f'[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)] = 0.5$ and $f[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)] = 0.66$, while in case

$$0 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 0.50, \tag{vi}$$

we have $f'[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)] = f[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)] = 0.33$. Let us remark that, for each car h for which

$$-0.50 < \mu_{\text{speedy\_cars}}(h) - \mu_{\text{expensive\_cars}}(h) \leqslant 0,$$

and therefore $f'[\mu_{\text{speedy\_cars}}(h), \mu_{\text{expensive\_cars}}(h)] = 0.5$ and $f[\mu_{\text{speedy\_cars}}(h), \mu_{\text{expensive\_cars}}(h)] = 0.66$, there exists at least one other car k for which $\mu_{\text{speedy\_cars}}(k) \geqslant \mu_{\text{speedy\_cars}}(h)$ and $\mu_{\text{expensive\_cars}}(k) \leqslant \mu_{\text{expensive\_cars}}(h)$ such that

$$0 < \mu_{\text{speedy\_cars}}(k) - \mu_{\text{expensive\_cars}}(k) \leqslant 0.50,$$

and therefore $f'[\mu_{\text{speedy\_cars}}(k), \mu_{\text{expensive\_cars}}(k)] = f[\mu_{\text{speedy\_cars}}(k), \mu_{\text{expensive\_cars}}(k)] = 0.33$. Thus, for the principle of coherence with respect to the sign of relationships between condition attributes $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$, on one hand, and decision attribute $Y_{\text{good\_cars}}$ on the other hand, it is cautious with respect to f as well as to f' to conclude that

$$\begin{aligned}\mu_{\text{good\_cars}}(h) = \mu_{\text{good\_cars}}(k) &= f'[\mu_{\text{speedy\_cars}}(k), \mu_{\text{expensive\_cars}}(k)] \\ &= f[\mu_{\text{speedy\_cars}}(k), \mu_{\text{expensive\_cars}}(k)] = 0.33,\end{aligned}$$

without taking into account that $f'[\mu_{\text{speedy\_cars}}(h), \mu_{\text{expensive\_cars}}(h)] = 0.5$ and $f[\mu_{\text{speedy\_cars}}(h), \mu_{\text{expensive\_cars}}(h)] = 0.66$.

Now, let us suppose that we want to evaluate the membership of each car in set $Y_{\text{good\_cars}}$ on the basis of U-rule $UR' = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g' \rangle$, where $g' : [0,1]^2 \to [0,1]$ is defined as follows:

$$g'[\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}] = \begin{cases} 0 & -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.50, \\ 0.66 & -0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0, \\ 0.40 & 0 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.50, \\ 1 & 0.50 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

We have that $C(z, UR, Y) = C(z, UR', Y)$ and therefore U-rules UR and UR' are equivalent. Again, the reason is that g and g' differ in a part of their domain only (when $0 < \mu_{\text{speedy\_cars}}(z) - \mu_{\text{expensive\_cars}}(z) \leqslant 0.5$) where the monotonicity of membership in $Y_{\text{good\_cars}}$ with respect to membership in $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ is not satisfied.

Theorem 1. *For each L-rule $LR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f \rangle$ there exists an equivalent L-rule $LR' = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f' \rangle$ with functions $f'(\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})$ non-decreasing in each of its first r arguments and non-increasing in its last s arguments.*

For each U-rule $UR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g \rangle$ there exists an equivalent U-rule $UR' = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g' \rangle$ with functions $g'(\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})$ non-decreasing in each of its first r arguments and non-increasing in its last s arguments.

Proof. Let us suppose that the L-rule $LR = \langle X^{\uparrow}_{i1},\ldots,X^{\uparrow}_{ir},X^{\downarrow}_{j1},\ldots,X^{\downarrow}_{js},Y,f\rangle$ does not satisfies the property to be non-decreasing in each of its first r arguments and/or non-increasing in its last s arguments. Let us consider the L-rule $LR' = \langle X^{\uparrow}_{i1},\ldots,X^{\uparrow}_{ir},X^{\downarrow}_{j1},\ldots,X^{\downarrow}_{js},Y,f'\rangle$ with function $f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$ defined as follows, for each $(\alpha'_{i1},\ldots,\alpha_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) \in [0,1]^{r+s}$

$$f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js}) = \inf\Big\{f(\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) : (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) \in [0,1]^{r+s}, \alpha'_{i1} \leqslant \alpha_{i1},\ldots,\alpha'_{ir} \leqslant \alpha_{ir},\alpha'_{j1} \geqslant \alpha_{j1},\ldots,\alpha'_{js} \geqslant \alpha_{js}\Big\}.$$

Let us prove that function $f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$ is non-decreasing in each of its first r arguments and non-increasing in its last s arguments. In fact, on the basis of definition of function $f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$, for each $(\alpha''_{i1},\ldots,\alpha''_{ir},\alpha''_{j1},\ldots,\alpha''_{js})$, $(\alpha'''_{i1},\ldots,\alpha'''_{ir},\alpha'''_{j1},\ldots,\alpha'''_{js}) \in [0,1]^{r+s}$ such that

$$\alpha''_{i1} \leqslant \alpha'''_{i1},\ldots\alpha''_{ir} \leqslant \alpha'''_{ir}, \quad \alpha''_{j1} \geqslant \alpha'''_{j1},\ldots,\alpha''_{js} \geqslant \alpha'''_{js}, \tag{i}$$

we have that

$$\begin{aligned} f'(\alpha''_{i1},\ldots,\alpha''_{ir},\alpha''_{j1},\ldots,\alpha''_{js}) &= \inf\Big\{f(\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) : (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) \\ &\quad \in [0,1]^{r+s},\alpha'_{i1} \geqslant \alpha''_{i1},\ldots\alpha'_{ir} \geqslant \alpha''_{ir},\alpha'_{j1} \leqslant \alpha''_{j1},\ldots,\alpha'_{js} \leqslant \alpha''_{js}\Big\} \\ &\leqslant \inf\Big\{f(\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) : (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) \in [0,1]^{r+s},\alpha'_{i1} \\ &\quad \geqslant \alpha'''_{i1},\ldots\alpha'_{ir} \geqslant \alpha'''_{ir},\alpha'_{j1} \leqslant \alpha'''_{j1},\ldots,\alpha'_{js} \leqslant \alpha'''_{js}\Big\} = f'(\alpha'''_{i1},\ldots,\alpha'''_{ir},\alpha'''_{j1},\ldots,\alpha'''_{js}). \end{aligned} \tag{ii}$$

Now we prove that L-rules LR and LR' are equivalent. For each possible object z

$$\begin{aligned} C(z,LR,Y) &= \inf_{\alpha\in E^{+}(z)} f(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js}) \\ &= \inf\Big\{f(\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) : (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js}) \in [0,1]^{r+s},\alpha'_{i1} \\ &\qquad \geqslant \mu_{X_{i1}}(z),\ldots\alpha'_{ir} \geqslant \mu_{X_{ir}}(z),\alpha'_{j1} \leqslant \mu_{X_{j1}}(z),\ldots,\alpha'_{js} \leqslant \mu_{X_{js}}(z)\Big\} \\ &= f'(\mu_{X_{i1}}(z),\ldots,\mu_{X_{ir}}(z),\mu_{X_{j1}}(z),\ldots,\mu_{X_{js}}(z)). \end{aligned} \tag{iii}$$

On the basis of monotonicity of function $f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})$

$$\begin{aligned} f'(\mu_{X_{i1}}(z),\ldots,\mu_{X_{ir}}(z),\mu_{X_{j1}}(z),\ldots,\mu_{X_{js}}(z)) &= \inf_{\alpha\in E^{+}(z)} f'(\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js}) \\ &= C(z,LR',Y). \end{aligned} \tag{iv}$$

Thus from (iii) and (iv) we have

$$C(z,LR,Y) = C(z,LR',Y).$$

Thus, we proved the Theorem with respect to L-rules. With respect to U-rules an analogous proof holds. □

Example 1 (*part* 4). With respect to the L-rule $LR = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f\rangle$ presented above, consider the following L-rule $LR^* = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^*\rangle$, where

$$f^*(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.33 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

We have $C(z, LR, Y) = C(z, LR^*, Y)$ and thus L-rule LR and L-rule LR^* are equivalent. Moreover, according to Theorem 1, $f^*(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ is non-decreasing with respect to $\alpha_{\text{speedy\_cars}}$ and non-increasing with respect to $\alpha_{\text{expensive\_cars}}$, so it satisfies monotonicity of membership in $Y_{\text{good\_cars}}$ with respect to membership in $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$.

Analogously, with respect to the U-rule $UR = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g\rangle$ presented above, consider the following U-rule $UR^* = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g^*\rangle$, where

$$g^*(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.66 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

We have $C(z, UR, Y) = C(z, UR^*, Y)$ and thus U-rule UR and U-rule UR^* are equivalent. Moreover, according to Theorem 1, $g^*(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ is non-decreasing with respect to $\alpha_{\text{speedy\_cars}}$ and non-increasing with respect to $\alpha_{\text{expensive\_cars}}$, so it satisfies monotonicity of membership in $Y_{\text{good\_cars}}$ with respect to membership in $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$.

An L-rule can be regarded as a gradual rule [5]; indeed, it can be interpreted as

"the more object x is $X_{i1}, \ldots, X_{ir}$ and the less object x is $X_{j1}, \ldots, X_{js}$, the more it is Y".

Analogously, the U-rule can be interpreted as

"the less object x is $X_{i1}, \ldots, X_{ir}$ and the more object x is $X_{j1}, \ldots, X_{js}$, the less it is Y".

On the other hand, the syntax of L- and U-rules is more general than that of usual gradual rules introduced in [5]. Indeed, while the usual gradual rules are statements of the type "if $\mu_X(x) \geqslant \alpha$, then $\mu_Y(x) \geqslant \alpha$", the simplest L-rule states "if $\mu_{X_i^{\uparrow}}(x) \geqslant \alpha_i$, then $\mu_Y(x) \geqslant \beta$" or "if $\mu_{X_j^{\downarrow}}(x) \leqslant \alpha_j$, then $\mu_Y(x) \geqslant \beta$". Therefore, the L- and U-rules permit to consider different degrees of credibility in premises and conclusion, which is not the case of the gradual rules.

Let us also remark that the syntax of L- and U-rules is similar to the syntax of "at least" and "at most" decision rules induced from dominance-based rough approximations of preference-ordered decision classes [9–11].

Example 1 (*part* 5). In terms of gradual rules, the L-rule LR and U-rule UR has the following structure:

"the more car z is $X^{\uparrow}_{\text{speedy\_cars}}$ and the less it is $X^{\downarrow}_{\text{expensive\_cars}}$, the more it is $Y_{\text{good\_cars}}$".

This is equivalent to

"the less car z is $X^{\uparrow}_{\text{speedy\_cars}}$ and the more it is $X^{\downarrow}_{\text{expensive\_cars}}$, the less it is $Y_{\text{good\_cars}}$".

An example of a usual gradual rule is the following:

$$r \equiv \text{"if } \mu_{\text{speedy\_cars}}(z) \geqslant \xi \text{ and } N[\mu_{\text{expensive\_cars}}(z)] \geqslant \xi, \text{ then } \mu_{\text{good\_cars}}(z) \geqslant \xi\text{"},$$

where $N(\cdot)$ is a negation, i.e., a decreasing function $N:[0,1] \to [0,1]$ such that $N(0) = 1$ and $N(1) = 0$, so that if $\alpha \in [0,1]$ is the credibility of proposition p, then $N(\alpha)$ is the credibility of $\neg p$ (the negation of p). Let us observe that conditions "$\mu_{\text{speedy\_cars}}(z) \geqslant \xi$" and "$N[\mu_{\text{expensive\_cars}}(z)] \geqslant \xi$" and the conclusion "$\mu_{\text{good\_cars}}(z) \geqslant \xi$" are all related to the same threshold ξ. This is not the case of the gradual rules corresponding to above L-rule *LR* and U-rule *UR*. For example, according to above L-rule *LR* and considering $N(\alpha) = 1 - \alpha$,

$$r' \equiv \text{"if } \mu_{\text{speedy\_cars}}(z) \geqslant 0.8 \text{ and } N[\mu_{\text{expensive\_cars}}(z)] \geqslant 0.6, \text{ then } \mu_{\text{good\_cars}}(z) \geqslant 0.33\text{"}$$

(in fact $N[\mu_{\text{expensive\_cars}}(z)] = 1 - \mu_{\text{expensive\_cars}}(z) \geqslant 0.6$ implies $\mu_{\text{expensive\_cars}}(z) \leqslant 0.4$ and for $\mu_{\text{speedy\_cars}}(z) \geqslant 0.8$ and $\mu_{\text{expensive\_cars}}(z) \leqslant 0.4$ we have $C(z, LR, Y_{\text{good\_cars}}) \geqslant 33$). Let us remark that in rule r' there are different thresholds for $\mu_{\text{speedy\_cars}}(z)$ (0.8), $N[\mu_{\text{expensive\_cars}}(z)]$ (0.6) and $\mu_{\text{good\_cars}}(z)$ (0.33). Analogous arguments hold with respect to above U-rule *UR*.

3. Fuzzy rough approximations

The functions f and g introduced in the previous section are related to specific definitions of lower and upper approximations considered within rough set theory [18]. Let us consider a universe of discourse U and $r + s + 1$ fuzzy sets, $X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$ and Y, defined on U by means of membership functions $\mu_{X_h} : U \to [0,1]$, $h \in \{i1, \ldots, ir, j1, \ldots, js\}$ and $\mu_Y : U \to [0,1]$. Suppose that we want to approximate knowledge contained in Y using knowledge about $X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$, under the hypothesis that $X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}$ are positively related with Y and $X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$ are negatively related with Y.

Then, the lower approximation of Y given the information on $X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$ is a fuzzy set $\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y)$, whose membership function for each $x \in U$, denoted by $\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]$, is defined as follows:

$$\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] = \inf_{z \in D\uparrow(x)} \{\mu_Y(z)\}, \tag{1}$$

where for each $x \in U$, $D\uparrow(x)$ is a non-empty set defined by

$$D\uparrow(x) = \Big\{ z \in U : \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h = X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, \text{ and } \mu_{X_h}(z) \leqslant \mu_{X_h}(x) \text{ for each } X_h = X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js} \Big\}.$$

Lower approximation $\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]$ can be interpreted as follows: in the universe U the following implication holds:

"If $\mu_{X_h}(z) \geqslant \mu_{X_h}(x)$ for each $X_h = X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$, and $\mu_{X_h}(z) \leqslant \mu_{X_h}(x)$ for each $X_h = X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$, then $\mu_Y(z) \geqslant \mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]$".

Interpretation of lower approximation (1) is based on a specific meaning of the concept of ambiguity. According to knowledge about $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$, the membership of object $x \in U$ to fuzzy set Y is ambiguous if there exists an object $z \in U$ such that $\mu_{X_h}(z) \geqslant \mu_{X_h}(x)$ for each $X_h = X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$, and $\mu_{X_h}(z) \leqslant \mu_{X_h}(x)$ for each $X_h = X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$, however, $\mu_Y(x) > \mu_Y(z)$.

Remark that the above meaning of ambiguity is concordant with the dominance principle introduced in rough set theory in order to deal with preference-ordered data [9–11]. In this case, the dominance principle says that, having an object with some membership degrees in X and Y, its modification consisting in an increase of its membership in X should not decrease its membership in Y; otherwise, the original object and the modified object are ambiguous.

Analogously, the upper approximation of Y given the information on $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$, $X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$ is a fuzzy set $\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y)$, whose membership function for each $x \in U$, denoted by $\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]$, is defined as follows:

$$\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x] = \sup_{z \in D\downarrow(x)} \{\mu_Y(z)\}, \tag{2}$$

where for each $x \in U$, $D\downarrow(x)$ is a non-empty set defined by

$$D\downarrow(x) = \Big\{ z \in U : \mu_{X_h}(z) \leqslant \mu_{X_h}(x) \text{ for each } X_h = X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow} \\ \text{and } \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h = X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow} \Big\}.$$

Upper approximation $\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]$ can be interpreted as follows: in the universe U the following implication holds:

"If $\mu_{X_h}(z) \leqslant \mu_{X_h}(x)$ for each $X_h = X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}$, and $\mu_{X_h}(z) \geqslant \mu_{X_h}(x)$ for each $X_h = X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}$, then $\mu_Y(z) \leqslant \mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]$".

Example 1 (*part* 6). Let us consider a universe of discourse U composed of all the cars z such that $(\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)) \in [0,1]^2$, i.e., of all possible and imaginable cars. Suppose that for all $z \in U$ membership of z in $Y_{\text{good\_cars}}$ is given as

$$\mu_{\text{good\_cars}}(z) = f[\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)],$$

where function f (which in our didactic example could represent customer preferences) is defined as in above part 1 of this example. We want to approximate knowledge contained in $Y_{\text{good\_cars}}$ using knowledge about $X_{\text{speedy\_cars}}^{\uparrow}$ and $X_{\text{expensive\_cars}}^{\downarrow}$ under the hypothesis that membership in set $Y_{\text{good\_cars}}$ of good cars is positively related with the membership in set $X^{\uparrow}_{\text{speedy\_cars}}$ of speedy cars and negatively related with the membership in set $X^{\downarrow}_{\text{expensive\_cars}}$ of expensive cars.

The lower approximation of $Y_{\text{good\_cars}}$ given the information on $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ is a fuzzy set $\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}})$, whose membership function for each $x \in U$, denoted by $\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]$, is defined as follows:

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = \inf_{z \in D\uparrow(x)} \{\mu_{\text{good\_cars}}(z)\}$$

where for each $x \in U$, $D\uparrow(x)$ is a non-empty set defined by

$$D\uparrow(x) = \{z \in U : \mu_{\text{speedy\_cars}}(z) \geqslant \mu_{\text{speedy\_cars}}(x) \text{ and } \mu_{\text{expensive\_cars}}(z) \leqslant \mu_{\text{expensive\_cars}}(x)\}.$$

Thus, for $x \in U$ such that $\mu_{\text{speedy\_cars}}(x) = 0.7$ and $\mu_{\text{expensive\_cars}}(x) = 0.9$, we have that

$$\begin{aligned}
&\mu[\underline{App}(X^{\uparrow}_{\text{speedy_cars}}, X^{\downarrow}_{\text{expensive_cars}}, Y_{\text{good_cars}}), x] \\
&\quad = \inf_{z \in D\uparrow(x)} \{\mu_{\text{good_cars}}(z)\} = \inf \{\mu_{\text{good_cars}}(z) : \mu_{\text{speedy_cars}}(z) \\
&\quad \geqslant 0.7 \text{ and } \mu_{\text{expensive_cars}}(z) \leqslant 0.9\} = 0.33.
\end{aligned}$$

Let us observe that, in general, we have the following *explicit formulation of lower approximation*:

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant -0.5, \\ 0.33 & \text{if } -0.5 < \mu_{\text{speedy\_cars}}(\mathrm{x}) - \mu_{\text{expensive\_cars}}(x) \leqslant 0.5, \\ 1 & \text{if } 0.5 < \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant 1. \end{cases}$$

The upper approximation of $Y_{\text{good\_cars}}$ given the information on $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ is a fuzzy set $\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}})$, whose membership function for each $x \in U$, denoted by $\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]$, is defined as follows:

$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = \sup_{z \in D\downarrow(x)} \{\mu_{\text{good\_cars}}(z)\},$$

where for each $x \in U$, $D\downarrow(x)$ is a non-empty set defined by

$$D\downarrow(x) = \{z \in U : \mu_{\text{speedy\_cars}}(z) \leqslant \mu_{\text{speedy\_cars}}(x) \text{ and } \mu_{\text{expensive\_cars}}(z) \geqslant \mu_{\text{expensive\_cars}}(x)\}.$$

Thus, for $x \in U$ such that $\mu_{\text{speedy\_cars}}(x) = 0.7$ and $\mu_{\text{expensive\_cars}}(x) = 0.9$, we have that

$$\begin{aligned}
&\mu[\overline{App}(X^{\uparrow}_{\text{speedy_cars}}, X^{\downarrow}_{\text{expensive_cars}}, Y_{\text{good_cars}}), x] \\
&\quad = \sup_{z \in D\downarrow(x)} \{\mu_{\text{good_cars}}(z)\} = \inf \{\mu_{\text{good_cars}}(z) : \mu_{\text{speedy_cars}}(z) \\
&\quad \leqslant 0.7 \text{ and } \mu_{\text{expensive_cars}}(z) \geqslant 0.9\} = 0.66.
\end{aligned}$$

Let us observe that in general we have the following *explicit formulation of upper approximation*:

$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant -0.5, \\ 0.66 & \text{if } -0.5 < \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant 0.5, \\ 1 & \text{if } 0.5 < \mu_{\text{speedy\_cars}}(\text{x}) - \mu_{\text{expensive\_cars}}(x) \leqslant 1. \end{cases}$$

Theorem 2. *Let us consider fuzzy sets* $X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$ *and* Y *defined on* U. *The following properties are satisfied*:

(1) *for each* $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] \leqslant \mu_Y(x) \leqslant \mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]$$

(2) *for any negation* $N(\cdot)$, *being a strictly decreasing function* $N: [0, 1] \to [0, 1]$ *such that* $N(1) = 0$ *and* $N(0) = 1$, *for each fuzzy set* $X_h = X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}$ *and* Y *defined on* U, *and for each* $x \in U$

(2.1) $\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y^{c}), x] = N(\mu[\overline{App}(X^{c\uparrow}_{i1}, \ldots, X^{c\uparrow}_{ir}, X^{c\downarrow}_{j1}, \ldots, X^{c\downarrow}_{js}, Y), x])$,

(2.2) $\mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y^{c}), x] = N(\mu[\underline{App}(X^{c\uparrow}_{i1}, \ldots, X^{c\uparrow}_{ir}, X^{c\downarrow}_{j1}, \ldots, X^{c\downarrow}_{js}, Y), x])$,

(2.3) $N(\mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]) = \mu[\underline{App}(X^{c\uparrow}_{i1}, \ldots, X^{c\uparrow}_{ir}, X^{c\downarrow}_{j1}, \ldots, X^{c\downarrow}_{js}, Y^{c}), x]$,

(2.4) $N(\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]) = \mu[\overline{App}(X^{c\uparrow}_{i1}, \ldots, X^{c\uparrow}_{ir}, X^{c\downarrow}_{j1}, \ldots, X^{c\downarrow}_{js}, Y^{c}), x]$,

where for a given fuzzy set W, *the fuzzy set* W^{c} *is its complement defined by* $\mu_{W^{c}}(x) = N(\mu_W(x))$.

(3) *for each* $\{X^{\uparrow}_{h1}, \ldots, X^{\uparrow}_{hv}\} \subseteq \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}$ *and* $\{X^{\downarrow}_{k1}, \ldots, X^{\downarrow}_{kw}\} \subseteq \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\}$

(3.1) $\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] \geqslant \mu[\underline{App}(X^{\uparrow}_{h1}, \ldots, X^{\uparrow}_{hv}, X^{\downarrow}_{k1}, \ldots, X^{\downarrow}_{kw}, Y), x]$,

(3.2) $\mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] \leqslant \mu[\overline{App}(X^{\uparrow}_{h1}, \ldots, X^{\uparrow}_{hv}, X^{\downarrow}_{k1}, \ldots, X^{\downarrow}_{kw}, Y), x]$.

(4) *for each* $x, y \in U$, *such that* $\mu_{X_h}(x) \geqslant \mu_{X_h}(y)$ *for each* $X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}$, *and* $\mu_{X_h}(x) \leqslant \mu_{X_h}(y)$ *for each* $X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\}$, *we have*

(4.1) $\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] \geqslant \mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), y]$,

(4.2) $\mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] \geqslant \mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), y]$.

Proof. Obviously, for all $x \in U$, $\mu_{X_h}(x) \geqslant \mu_{X_h}(x)$ for each $X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}$, and $\mu_{X_h}(x) \leqslant \mu_{X_h}(x)$ for each $X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\}$ and, therefore, $x \in D{\uparrow}(x)$. Thus we have

$$\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x] = \inf_{z \in D\uparrow(x)} \{\mu_Y(z)\} \leqslant \mu_Y(x). \tag{i}$$

Analogously, $x \in D{\downarrow}(x)$ and, therefore, we have that

$$\mu_Y(x) \leqslant \sup_{z \in D\downarrow(x)} \{\mu_Y(z)\} = \mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]. \tag{ii}$$

From (i) and (ii) we obtain

$$\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x] \leqslant \mu_Y(x) \leqslant \mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x].$$

Thus, we proved (1).

According to the above definition of rough approximation, we have

$$\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y^{c}),x] = \inf_{z\in D\uparrow(x)}\{N(\mu_y(z))\} = N(\sup_{z\in D\uparrow(x)}\{\mu_y(z)\}). \qquad \text{(iii)}$$

Now, to each $x \in U$ and to each $X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}$ let us associate the negation of the membership functions $\mu_{X_h}(x)$, i.e., $N[\mu_{X_h}(x)]$. Remembering that $N(\cdot)$ is strictly decreasing we obtain

$$\begin{aligned} D\uparrow(x) &= \Big\{z \in U : \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(z) \\ &\qquad \leqslant \mu_{X_h}(x) \text{ for each } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}\Big\} \\ &= \Big\{z \in U : N(\mu_{X_h}(z)) \leqslant N(\mu_{X_h}(x)) \text{ for each } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and} \\ &\qquad N(\mu_{X_h}(z)) \geqslant N(\mu_{X_h}(x)) \text{ for each } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}\Big\} = D\downarrow^{c}(x), \end{aligned} \qquad \text{(iv)}$$

where the index "c" in $D\downarrow^{c}(x)$ denotes that we are considering the negation of the membership functions $\mu_{X_h}(x)$. On the basis of (iv) we can write

$$\begin{aligned} N\left(\sup_{z\in D\uparrow(x)}\{\mu_y(z)\}\right) &= N\left(\sup_{z\in D\downarrow^{c}(x)}\{\mu_y(z)\}\right) \\ &= N\Big(\mu[\overline{App}(X_{i1}^{c\uparrow},\ldots,X_{ir}^{c\uparrow},X_{j1}^{c\downarrow},\ldots,X_{js}^{c\downarrow},Y),x]\Big). \end{aligned} \qquad \text{(v)}$$

From (iii) and (v) we obtain

$$\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y^{c}),x] = N\Big(\mu[\overline{App}(X_{i1}^{c\uparrow},\ldots,X_{ir}^{c\uparrow},X_{j1}^{c\downarrow},\ldots,X_{js}^{c\downarrow},Y),x]\Big).$$

Thus we proved (2.1). (2.2)–(2.4) can be proved analogously.

Now we consider $R = \{X_{h1}^{\uparrow},\ldots,X_{hv}^{\uparrow}\}$, $S = \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}$, $T = \{X_{k1}^{\downarrow},\ldots,X_{kw}^{\downarrow}\}$ and $V = \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}$ such that $R \subseteq S$ and $T \subseteq V$.

We consider also

$$\begin{aligned} &D(R\cup T)\uparrow(x) \\ &\quad = \big\{z \in U : \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h \in R, \text{ and } \mu_{X_h}(z) \leqslant \mu_{X_h}(x) \text{ for each } X_h \in T\big\}, \\ &D(S\cup V)\uparrow(x) \\ &\quad = \big\{z \in U : \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h \in S, \text{ and } \mu_{X_h}(z) \leqslant \mu_{X_h}(x) \text{ for each } X_h \in V\big\}, \\ &D(R\cup T)\downarrow(x) \\ &\quad = \big\{z \in U : \mu_{X_h}(z) \leqslant \mu_{X_h}(x) \text{ for each } X_h \in R, \text{ and } \mu_{X_h}(z) \geqslant \mu_{X_h}(x) \text{ for each } X_h \in T\big\}, \end{aligned}$$

and

$$D(S\cup V)\downarrow(x)=\big\{z\in U:\mu_{X_h}(z)\leqslant\mu_{X_h}(x)\text{ for each }X_h\in S,\text{ and }\mu_{X_h}(z)\geqslant\mu_{X_h}(x)\text{ for each }X_h\in V\big\}.$$

Since $R\subseteq S$ and $T\subseteq V$, we have that

$$D(S\cup V)\uparrow(x)\subseteq D(R\cup T)\uparrow(x) \tag{vi}$$

and

$$D(S\cup V)\downarrow(x)\subseteq D(R\cup T)\downarrow(x). \tag{vii}$$

On the basis of (vi) we have that

$$\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]=\inf_{z\in D(S\cup V)\uparrow(x)}\{\mu_y(z)\}\geqslant\inf_{z\in D(R\cup T)\uparrow(x)}\{\mu_y(z)\}$$
$$=\mu[\underline{App}(X_{h1}^{\uparrow},\ldots,X_{hv}^{\uparrow},X_{k1}^{\downarrow},\ldots,X_{kw}^{\downarrow},Y),x].$$

Thus we proved (3.1).

On the basis of (vii) we have that

$$\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]=\sup_{z\in D(S\cup V)\downarrow(x)}\{\mu_y(z)\}\leqslant\sup_{z\in D(R\cup T)\downarrow(x)}\{\mu_y(z)\}$$
$$=\mu[\overline{App}(X_{h1}^{\uparrow},\ldots,X_{hv}^{\uparrow},X_{k1}^{\downarrow},\ldots,X_{kw}^{\downarrow},Y),x].$$

Thus we proved (3.2).

Now let us consider $x,y\in U$ such that $\mu_{X_h}(x)\geqslant\mu_{X_h}(y)$ for each $X_h\in\{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}$, and $\mu_{X_h}(x)\leqslant\mu_{X_h}(y)$ for each $X_h\in\{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}$. We have that

$$D\uparrow(x)\subseteq D\uparrow(y) \tag{viii}$$

and

$$D\downarrow(x)\supseteq D\downarrow(y). \tag{ix}$$

From (viii) we obtain

$$\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]=\inf_{z\in D\uparrow(x)}\{\mu_Y(z)\}\geqslant\inf_{z\in D\uparrow(y)}\{\mu_y(z)\}$$
$$=\mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),y].$$

Thus we proved (4.1).

From (ix) we obtain

$$\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x]=\sup_{z\in D\downarrow(x)}\{\mu_Y(z)\}\geqslant\sup_{z\in D\downarrow(y)}\{\mu_y(z)\}$$
$$=\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),y].$$

Thus we proved (4.2). □

Results (1), (2) and (3) of Theorem 2 can be read as fuzzy counterparts of results well-known within the classical rough set theory. More precisely, (1) says that fuzzy set Y includes its lower approximation and is included in its upper approximation; (2) represents complementarity properties of the proposed fuzzy rough approximations; (3) expresses the fact that when we approximate Y, if we pass from a set of attributes to its subset, for any

$x \in U$, the membership to the lower approximation of Y does not increase while the membership to the upper approximation of Y does not decrease. Result (4) is more related with the specific context in which we are defining rough approximation: it says that lower and upper approximations respect monotonicity with respect to fuzzy membership functions $\mu_{X_h}(x)$, and more precisely, that they are non-decreasing operators with respect to $\mu_{X_h}(x)$ for $X_h \in \{X^{\uparrow}_{i1},\ldots,X^{\uparrow}_{ir}\}$ and non-increasing operators with respect to $\mu_{X_h}(x)$ for $X_h \in \{X^{\downarrow}_{j1},\ldots,X^{\downarrow}_{js}\}$.

Example 1 (*part* 7). Taking into account car x already introduced in part 6 of this example (let us remember that $\mu_{\text{speedy\_cars}}(x) = 0.7$ and $\mu_{\text{expensive\_cars}}(x) = 0.9$), we can see that, according to point (1) of Theorem 2,

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] \leqslant \mu_{\text{good\_cars}}(x) \leqslant \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]$$

(let us remember that $\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = 0.33$, $\mu_{\text{good\_cars}}(x) = 0.33$ and $\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] = 0.66$).

Let us suppose now that we want to approximate the membership in set $(Y_{\text{good\_cars}})^{\text{c}} = Y_{\text{bad\_cars}}$ such that for each $x \in U$, we have that $\mu_{\text{bad\_cars}}(x) = N[\mu_{\text{good\_cars}}(x)]$, where $N(\alpha) = 1 - \alpha$ (but the results we obtain in this part of the example do not depend on the specific formulation of N). Let us also suppose that we approximate the knowledge contained in $Y_{\text{bad\_cars}}$ using knowledge about $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ under the hypothesis that membership in set $Y_{\text{bad\_cars}}$ of bad cars is positively related with the membership in set $X^{\uparrow}_{\text{speedy\_cars}}$ of speedy cars and negatively related with the membership in set $X^{\downarrow}_{\text{expensive\_cars}}$ of expensive cars. Since this hypothesis is not well founded ("the more a car is speedy and the less it is expensive, the worse it is" is a paradoxical hypothesis), the final results are not interesting. In fact, we have that for each $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x] = 0$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x] = 1.$$

Let us also try to approximate knowledge contained in $Y_{\text{good\_cars}}$ using knowledge about $(X^{\uparrow}_{\text{speedy\_cars}})^{\text{c}} = X^{\uparrow}_{\text{slow\_cars}}$ and $(X^{\downarrow}_{\text{expensive\_cars}})^{\text{c}} = X^{\downarrow}_{\text{cheap\_cars}}$ ($\mu_{\text{slow\_cars}}(x) = N[\mu_{\text{speedy\_cars}}(x)]$ and $\mu_{\text{expensive\_cars}}(x) = N[\mu_{\text{cheap\_cars}}(x)]$) under the hypothesis that membership in set $Y_{\text{good\_cars}}$ of good cars is positively related with the membership in set $X^{\uparrow}_{\text{slow\_cars}}$ of slow cars and negatively related with the membership in set $X^{\downarrow}_{\text{cheap\_cars}}$ of cheap cars. Since this hypothesis is also not well founded ("the more a car is slow and the less it is cheap, the better it is" is a paradoxical hypothesis) the final results are again not interesting. In fact, we have that for each $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{good\_cars}}), x] = 0$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{good\_cars}}), x] = 1.$$

Let us remark that, according to points (2.1) and (2.2) of Theorem 2, we have that

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x] = N\left(\mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]\right)$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x] = N\left(\mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]\right).$$

Now, let us approximate knowledge contained in $Y_{\text{bad\_cars}}$ using knowledge about $X^{\uparrow}_{\text{slow\_cars}}$ and $X^{\downarrow}_{\text{cheap\_cars}}$, under the hypothesis that membership in set $Y_{\text{bad\_cars}}$ of good cars is positively related with the membership in set $X^{\uparrow}_{\text{slow\_cars}}$ of slow cars and negatively related with the membership in set $X^{\downarrow}_{\text{cheap\_cars}}$ of cheap cars. This hypothesis is meaningful, of course ("the more a car is slow and the less it is cheap, the worse it is" is somehow equivalent to "the more a car is speedy and the less it is expensive, the better it is"), and the final results are quite interesting. In fact we have that for each $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_,cars}}), x] = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) < -0.5, \\ 0.34 & \text{if } -0.5 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) < 0.5, \\ 1 & \text{if } 0.5 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) \leqslant 1 \end{cases}$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_,cars}}), x] = \begin{cases} 0 & \text{if } -1 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) < -0.5, \\ 0.67 & \text{if } -0.5 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) < 0.5, \\ 1 & \text{if } 0.5 \leqslant \mu_{\text{slow\_cars}}(x) - \mu_{\text{cheap\_cars}}(x) \leqslant 1. \end{cases}$$

Remembering that $\mu_{\text{slow\_cars}}(x) = 1 - \mu_{\text{speedy\_cars}}(x)$ and $\mu_{\text{cheap\_cars}}(x) = 1 - \mu_{\text{expensive\_cars}}(x)$ we can rewrite above membership functions of rough approximations as

$$\mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_cars}}), x] = \begin{cases} 0 & \text{if } 0.5 < \mu_{\text{speedy\_cars}} - \mu_{\text{expensive\_cars}} \leqslant 1, \\ 0.34 & \text{if } -0.5 < \mu_{\text{speedy\_cars}} - \mu_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}} - \mu_{\text{expensive\_cars}} \leqslant -0.5 \end{cases}$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_cars}}), x] = \begin{cases} 0 & \text{if } 0.5 < \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant 1, \\ 0.67 & \text{if } -0.5 < \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant 0.5, \\ 1 & \text{if } -1 \leqslant \mu_{\text{speedy\_cars}}(x) - \mu_{\text{expensive\_cars}}(x) \leqslant -0.5. \end{cases}$$

Let us remark that, according to points (2.3) and (2.4) of Theorem 2, we have that

$$N\left(\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]\right) = \mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_cars}}), x]$$

and

$$N\left(\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]\right) = \mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{cheap\_cars}}, Y_{\text{bad\_cars}}), x].$$

These two equalities can be interpreted as follows:

- the credibility that car x is certainly not a good car – $N(\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x])$ – is equivalent to the credibility that x could be a bad car – $\mu[\overline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x]$;
- the credibility that it is false that car x could be a good car – $N(\mu[\overline{App}\ (X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x])$ – is equivalent to the credibility that x is certainly a bad car – $\mu[\underline{App}(X^{\uparrow}_{\text{slow\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{bad\_cars}}), x]$.

Now, let us approximate knowledge contained in $Y_{\text{good\_cars}}$ using only knowledge about $X^{\uparrow}_{\text{speedy\_cars}}$ under the hypothesis that $X^{\uparrow}_{\text{speedy\_cars}}$ is positively related with $Y_{\text{good\_cars}}$. In other words, we do not consider knowledge about $X^{\downarrow}_{\text{expensive\_cars}}$. We obtain the following rough approximations: for all $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, Y_{\text{good\_cars}}), x] = 0,$$
$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, Y_{\text{good\_cars}}), x] = 1.$$

It is clear that removing information about $X^{\downarrow}_{\text{expensive\_cars}}$ reduces drastically the accuracy of the approximation; in fact, according to point (3) of Theorem 2, we have that for all $x \in U$

$$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, Y_{\text{good\_cars}}), x] \leqslant \mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x]$$

and

$$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, Y_{\text{good\_cars}}), x] \geqslant \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x].$$

Approximating knowledge contained in $Y_{\text{good\_cars}}$ using only knowledge about $X^{\downarrow}_{\text{expensive\_cars}}$, i.e., removing information about $X^{\uparrow}_{\text{speedy\_cars}}$, we obtain analogous results.

Let us consider two cars, x and y, such that $\mu_{\text{speedy\_cars}}(x) = 0.2$, $\mu_{\text{expensive\_cars}}(x) = 0.8$, $\mu_{\text{speedy\_cars}}(y) = 0.8$ and $\mu_{\text{expensive\_cars}}(y) = 0.4$. We have

$$\begin{aligned} 0 &= \mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] \\ &\leqslant \mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), y] = 0.33 \end{aligned}$$

and

$$\begin{aligned} 0 &= \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x] \\ &\leqslant \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), y] = 0.66. \end{aligned}$$

This result agrees with point (4) of Theorem 3 which states that increasing the value of the condition attributes being positively related with the decision attributes ($\mu_{\text{speedy\_cars}}(x) \leqslant \mu_{\text{speedy\_cars}}(y)$) and decreasing the value of the attributes being negatively related with the decision attributes ($\mu_{\text{expensive\_cars}}(x) \geqslant \mu_{\text{expensive\_cars}}(y)$), the value of the rough approximations increases or, at least, does not decrease (the membership in both lower and upper approximations of y are greater, or at least not smaller, than the analogous membership of x). More generally, from the explicit formulation of lower and upper approximations presented in part 6 of the example, it is clear that the membership in rough approximations is monotonic with respect to attributes positively and negatively related with the decision attribute.

4. Decision rule induction from fuzzy rough approximations

The lower and upper approximations defined above can serve to induce L-rules and U-rules, respectively. Let us remark that inferring L-rules $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and U-rules $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g\rangle$ is equivalent to finding functions $f(\cdot)$ and $g(\cdot)$. Since we want to induce decision rules representing the considered universe U, the following conditions of correct representation must be satisfied by the L-rule $\langle X_{i1}^{\uparrow},\ldots,X_{ir},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and U-rule $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g\rangle$ searched for:

- correct representation with respect to the lower approximation: for all $x\in U$ and for each $\boldsymbol{\alpha}\in[0,1]^{r+s}$,

$$[\mu_{X_h}(x)\leqslant\alpha_h \text{ for each } X_h\in\{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x)\geqslant\alpha_h \text{ for each } X_h\in\{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}]$$
$$\Rightarrow\ f(\boldsymbol{\alpha})\geqslant\mu[\underline{App}(X_{i1}^{\uparrow},s\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x],$$

- correct representation with respect to the upper approximation: for all $x\in U$ and for each $\boldsymbol{\alpha}\in[0,1]^{r+s}$,

$$[\mu_{X_h}(x)\geqslant\alpha_h \text{ for each } X_h\in\{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x)\leqslant\alpha_h \text{ for each } X_h\in\{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}]$$
$$\Rightarrow\ g(\boldsymbol{\alpha})\leqslant\mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),x].$$

These conditions of correct representation are concordant with the idea that lower and upper approximation are reference values for a cautious lower and upper evaluation of membership in set Y on the basis of the membership in $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots$ and $X_{js}^{\downarrow}$.

In general, there are more than one L-rule $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and more than one U-rule $\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g\rangle$ satisfying the correct representation condition. Thus, how to choose "the best L-rule and the best U-rule"? To answer this question, we propose the following conditions of prudence:

- given two L-rules $LR=\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f\rangle$ and $LR'=\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,f'\rangle$ we say that LR is more prudent than LR' if for all $\boldsymbol{\alpha}\in[0,1]^{r+s}$, $f(\boldsymbol{\alpha})\leqslant f'(\boldsymbol{\alpha})$,
- given two U-rules $UR=\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g\rangle$ and $UR'=\langle X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y,g'\rangle$ we say that UR is more prudent than UR' if for all $\boldsymbol{\alpha}\in[0,1]^{r+s}$, $g(\boldsymbol{\alpha})\geqslant g'(\boldsymbol{\alpha})$.

These conditions of prudence are concordant with the idea of presenting the most cautious evaluation of membership in set Y on the base of the membership in $X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,$ and $X_{js}^{\downarrow}$. In this sense, the "lower evaluation" of the membership in set Y should be the smallest possible while the "upper evaluation" should be the greatest possible.

Example 1 (*part* 8). Let us consider the following L-decision rule $LR_1=\langle X_{\text{speedy\_cars}}^{\uparrow},X_{\text{expensive\_cars}}^{\downarrow},Y_{\text{good\_cars}},f_1\rangle$, where $f_1:[0,1]\times[0,1]\rightarrow[0,1]$ is defined as follows:

$$f_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0.16 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.50 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

Comparing rule LR_1 with the lower approximation of $Y_{\text{good\_cars}}$ (more precisely, comparing f_1 with the explicit formulation of the lower approximation of $Y_{\text{good\_cars}}$ presented in part 6) and taking into account that for all $x \in U$ and for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$[\mu_{\text{speedy\_cars}}(x) \leqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(x) \geqslant \alpha_{\text{expensive\_cars}}] \\ \Rightarrow f_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \geqslant \mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x],$$

we can conclude that rule LR_1 satisfies the property of correct representation with respect to the lower approximation. This means that using rule LR_1, we are considering the lower approximation of $Y_{\text{good\_cars}}$ as a reference for the minimum value we can give to decision attribute $Y_{\text{good\_cars}}$ on the basis of condition attributes $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$. Thus, there is no case for which rule LR_1 gives to attribute $Y_{\text{good\_cars}}$ an evaluation smaller than the lower approximation of any $x \in U$ such that x has a smaller evaluation on the attributes positively related with the decision attribute (in our example $\mu_{\text{speedy\_cars}}(x) \leqslant \alpha_{\text{speedy\_car}}$) and a larger evaluation on the attributes negatively related with the decision attribute (in our example $\mu_{\text{expensive\_cars}}(x) \geqslant \alpha_{\text{expensive\_cars}}$).

Let us consider now the following U-decision rule $UR_1 = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g_1\rangle$ where $g_1 : [0,1] \times [0,1] \to [0,1]$

$$g_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.50 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 0.75 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

Comparing rule UR_1 with the upper approximation of $Y_{\text{good\_cars}}$ (more precisely, comparing g_1 with the explicit formulation of the upper approximation of $Y_{\text{good\_cars}}$ presented in part 6) and taking into acount that for all $x \in U$ and for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$[\mu_{\text{speedy\_cars}}(x) \geqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(x) \leqslant \alpha_{\text{expensive\_cars}}] \\ \Rightarrow g_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \geqslant \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), x],$$

we can conclude that rule UR_1 satisfies the property of correct representation with respect to the upper approximation. This means that using rule UR_1, we are considering the upper approximation of $Y_{\text{good\_cars}}$ as a reference for the maximum value we can give to decision attribute $Y_{\text{good\_cars}}$ on the basis of condition attributes $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$. Thus, there is no case for which rule UR_1 gives to attribute $Y_{\text{good\_cars}}$ an evaluation greater than the upper approximation of any $x \in U$ such that x has a greater evaluation on the attributes positively related with the decision attribute (in our example $\mu_{\text{speedy\_cars}}(x) \geqslant \alpha_{\text{speedy\_cars}}$) and a smaller evaluation on the attributes negatively related with the decision attribute (in our example $\mu_{\text{expensive\_cars}}(x) \leqslant \alpha_{\text{expensive\_cars}}$).

Now, let us consider the L-decision rule $LR_2 = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f_2\rangle$, where $f_2 : [0,1] \times [0,1] \to [0,1]$ is defined as follows:

$$f_2(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0.05 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.45 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

One can easily verify that also L-decision rule LR_2 satisfies the property of correct representation with respect to the lower approximation. Moreover, since for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$f_2(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \leqslant f_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}),$$

we can conclude that rule LR_2 is more prudent than rule LR_1.

Let us also consider the U-decision rule $UR_2 = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g_2 \rangle$ where $g_2 : [0,1] \times [0,1] \to [0,1]$ is defined as follows:

$$g_2(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.55 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 0.90 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

One can easily verify that also U-decision rule UR_2 satisfies the property of correct representation with respect to the upper approximation. Moreover, since for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$g_2(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \geqslant g_1(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}),$$

we can conclude that rule UR_2 is more prudent than rule UR_1.

Let CLR be the set of all the L-rules $LR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f \rangle$ satisfying the condition of correct representation. We say that the L-rule $LR^{\#} = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f^{\#} \rangle$ is maximally prudent if $LR^{\#}$ is more prudent than all other LR rules in CLR.

Let also CUR be the set of all the U-rules $UR = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g \rangle$ satisfying the condition of correct representation. We say that the U-rule $UR^{\#} = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g^{\#} \rangle$ is maximally prudent if $UR^{\#}$ is more prudent than all other UR rules in CUR.

Theorem 3. *If $LR^{\#} = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, f^{\#} \rangle$ is an L-rule maximally prudent and $UR^{\#} = \langle X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y, g \rangle$ is an U-rule maximally prudent, then: for each $\boldsymbol{\alpha} \in [0,1]^{r+s}$,*

$$f^{\#}(\boldsymbol{\alpha}) = \inf_{LR \in CLR} f(\alpha) = \begin{cases} \sup_{x \in A^-(\alpha)} \{\mu[\underline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]\} & \text{if } A^-(\alpha) \neq \emptyset, \\ 0 & \text{if } A^-(\alpha) = \emptyset \end{cases}$$

and

$$g^{\#}(\boldsymbol{\alpha}) = \sup_{UR \in CUR} g(\alpha) = \begin{cases} \inf^{+}_{x \in A}(\alpha) \{\mu[\overline{App}(X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}, X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}, Y), x]\} & \text{if } A^+(\alpha) \neq \emptyset, \\ 1 & \text{if } A^+(\alpha) = \emptyset. \end{cases}$$

where

$$A^-(\boldsymbol{\alpha}) = \Big\{x \in U : \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}\Big\},$$

$$A^+(\boldsymbol{\alpha}) = \Big\{x \in U : \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}\Big\}.$$

Moreover, for any $z \in U$

$$f^{\#}(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)) = \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), z],$$

$$g^{\#}(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)) = \mu[\overline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), z].$$

Proof. Let us start by proving that $f^{\#}(\boldsymbol{\alpha})$ satisfies the correct representation with respect to the lower approximation. Let us consider $\boldsymbol{\alpha} \in [0,1]^{r+s}$ and $x \in U$ such that

$$\mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h = X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow} \quad \text{and} \quad \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h = X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}. \tag{i}$$

For (i), $x \in A^-(\boldsymbol{\alpha})$ and, therefore, $A^-(\boldsymbol{\alpha}) \neq \emptyset$. Consequently, for the definition of function $f^{\#}(\cdot)$ we have that

$$f^{\#}(\boldsymbol{\alpha}) = \sup_{y \in A^-(\alpha)} \{\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), y]\}. \tag{ii}$$

$x \in A^-(\boldsymbol{\alpha})$ implies also that

$$\sup_{y \in A^-(\alpha)} \{\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), y]\} \geqslant \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x]. \tag{iii}$$

From (ii) and (iii) we obtain

$$f^{\#}(\boldsymbol{\alpha}) \geqslant \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x]. \tag{iv}$$

(iv) means that $f^{\#}(\cdot)$ satisfies the correct representation with respect to the lower approximation.

Now, we prove that $LR^{\#} = \langle X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y, f^{\#}\rangle$ is the L-rule maximally prudent. For contradiction, let us suppose that there exists an L-rule $LR = \langle X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y, f\rangle$ and $\boldsymbol{\alpha} \in [0,1]^{r+s}$ such that

$$f(\boldsymbol{\alpha}) < f^{\#}(\boldsymbol{\alpha}). \tag{vi}$$

(vi) would mean that $LR^{\#}$ is not more prudent than LR and therefore $LR^{\#}$ would not be the L-rule maximally prudent.

Let us observe that $A^-(\boldsymbol{\alpha}) \neq \emptyset$. Otherwise, for the definition of $f^{\#}(\boldsymbol{\alpha})$, $A^-(\boldsymbol{\alpha}) = \emptyset$ would imply $f^{\#}(\boldsymbol{\alpha}) = 0$ and since $f(\boldsymbol{\alpha}) \geqslant 0$, (vi) could not hold.

Considering the definition of $f^{\#}(\cdot)$, (vi) gives

$$f(\boldsymbol{\alpha}) < \sup_{x \in A^-(\alpha)} \{\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x]\}. \tag{vii}$$

(vii) means that there exists $x \in U$ such that $\mu_{X_h}(x) \leqslant \alpha_h$ for each $X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}$, and $\mu_{X_h}(x) \geqslant \alpha_h$ for each $X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}$ and

$$f(\boldsymbol{\alpha}) < \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x]. \tag{viii}$$

(viii) says that L-rule $LR = \langle X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y, f\rangle$ does not satisfy correct representation with respect to the lower approximation. This means that for any L-rule LR, if LR is not more prudent than $LR^{\#}$, then LR is not in CLR. Thus we proved that $LR^{\#}$ is more prudent than all other LR rules in CLR. This completes the proof with respect to $LR^{\#}$. With respect to $UR^{\#}$ an analogous proof holds.

Now, we prove that for any $z \in U$

$$f^{\#}(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)) = \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), z]. \tag{ix}$$

We have that

$$A^-(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)) = \Big\{x \in U : \mu_{X_h}(x) \leqslant \mu_{X_h}(z) \text{ for each } X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x) \geqslant \mu_{X_h}(z) \text{ for each } X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}\Big\}.$$

Let us remark that for any $z \in U$

$$z \in A^-(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z))$$

and, therefore,

$$A^-(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)) \neq \emptyset. \tag{x}$$

For point (4) of Theorem 2, we have that for all $x, z \in U$ such that $\mu_{X_h}(x) \leqslant \mu_{X_h}(z)$ for each $X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}$ and $\mu_{X_h}(x) \geqslant \mu_{X_h}(z)$ for each $X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}$ we have that

$$\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x] \leqslant \mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), z]. \tag{xi}$$

Since for any $x \in A^-(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z))$ we have $\mu_{X_h}(x) \leqslant \mu_{X_h}(z)$ for each $X_h \in \{X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}\}$ and $\mu_{X_h}(x) \geqslant \mu_{X_h}(z)$ for each $X_h \in \{X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}\}$ we can conclude that (xi) holds for all $x \in A^-(\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z))$ and therefore

$$\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), z] = \sup_{x \in A^-(\boldsymbol{\mu}(z))} \Big\{\mu[\underline{App}(X_{i1}^{\uparrow}, \ldots, X_{ir}^{\uparrow}, X_{j1}^{\downarrow}, \ldots, X_{js}^{\downarrow}, Y), x]\Big\}, \tag{xii}$$

where $\boldsymbol{\mu}(z) = [\mu_{X_{i1}^{\uparrow}}(z), \ldots, \mu_{X_{ir}^{\uparrow}}(z), \mu_{X_{j1}^{\downarrow}}(z), \ldots, \mu_{X_{js}^{\downarrow}}(z)]$.

On the basis of the definition of $f^{\#}(\cdot)$ and for (x), (xii) gives (ix).

Analogously, we can prove that for any $z \in U$

$$g^{\#}(\mu_{X_{i1}^{\uparrow}}(z),\ldots,\mu_{X_{ir}^{\uparrow}}(z),\mu_{X_{j1}^{\downarrow}}(z),\ldots,\mu_{X_{js}^{\downarrow}}(z)) = \mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),z]. \quad \square$$

Theorem 3 is a characterization of the decision rules obtained through our fuzzy rough approach: there is only one L-rule and one U-rule maximally prudent in the set of L-rules and U-rules satisfying the property of correct representation and these are the L-rule $LR^{\#}$ and the U-rule $UR^{\#}$. Let us also remark the importance of lower and upper approximations obtained through our fuzzy rough approach for the definition of L-rule $LR^{\#}$ and U-rule $UR^{\#}$. The last part of Theorem 3 says that L-rule $LR^{\#}$ and U-rule $UR^{\#}$ permit an exact reclassification of any object $z \in U$. More precisely, function $f^{\#}$ reassigns z its lower approximation, i.e.,

$$f^{\#}(\mu_{X_{i1}^{\uparrow}}(z),\ldots,\mu_{X_{ir}^{\uparrow}}(z),\mu_{X_{j1}^{\downarrow}}(z),\ldots,\mu_{X_{js}^{\downarrow}}(z)) = \mu[\underline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),z],$$

while function $g^{\#}$ reassigns z its upper approximation, i.e.,

$$g^{\#}(\mu_{X_{i1}^{\uparrow}}(z),\ldots,\mu_{X_{ir}^{\uparrow}}(z),\mu_{X_{j1}^{\downarrow}}(z),\ldots,\mu_{X_{js}^{\downarrow}}(z)) = \mu[\overline{App}(X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow},X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow},Y),z].$$

Example 1 (*part* 9). According to Theorem 3, the maximally prudent L-rule $LR^{\#}$ is the rule $\langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^{\#}\rangle$, where for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$f^{\#}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \begin{cases} \sup_{x\in A^{-}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})}\left\{\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}},X^{\uparrow}_{\text{expensive\_cars}},y_{\text{good\_cars}}),x]\right\} \\ \quad \text{if } A^{-}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) \neq \emptyset, \\ 0 \ \text{ if } A^{-}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \emptyset \end{cases}$$

with

$$A^{-}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \{x \in U : \mu_{\text{speedy\_cars}}(x) \leqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(x) \geqslant \alpha_{\text{expensive\_cars}}\}$$

Writing $f^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ directly in terms of $\alpha_{\text{speedy\_cars}}$ and $\alpha_{\text{expensive\_cars}}$ we get

$$f^{\#}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.33 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

Analogously, according to Theorem 3 the maximally prudent U-rule $UR^{\#}$ is the rule $\langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g^{\#}\rangle$, where for each $(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \in [0,1]^2$

$$g^{\#}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \begin{cases} \inf_{x\in A^{+}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})}\{\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}},X^{\downarrow}_{\text{expensive\_cars}},Y_{\text{good\_cars}}),x]\} \\ \text{if } \ A^{+}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) \neq \emptyset, \\ 1 \text{ if } \ A^{+}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \emptyset, \end{cases}$$

where

$$A^{+}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}}) = \{x \in U : \mu_{\text{speedy\_cars}}(x) \geqslant \alpha_1 \text{ and } \mu_{\text{expensive\_cars}}(x) \leqslant \alpha_2\}$$

Writing $g^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ directly in terms of $\alpha_{\text{speedy\_cars}}$ and $\alpha_{\text{expensive\_cars}}$, we get

$$g^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) = \begin{cases} 0 & \text{if } -1 \leqslant \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant -0.5, \\ 0.66 & \text{if } -0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 0.5, \\ 1 & \text{if } 0.5 < \alpha_{\text{speedy\_cars}} - \alpha_{\text{expensive\_cars}} \leqslant 1. \end{cases}$$

Comparing $f^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ with the explicit formulation of the lower approximation introduced in part 6 of this example, one can easily verify that for any $z \in U$,

$$f^{\#}(\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)) = \mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), z].$$

Analogously, comparing $g^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}})$ with the explicit formulation of the upper approximation introduced in part 6 of this example, one can easily verify that for any $z \in U$,

$$g^{\#}(\mu_{\text{speedy\_cars}}(z), \mu_{\text{expensive\_cars}}(z)) = \mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), z].$$

5. Fuzzy rough modus-ponens and fuzzy rough modus tollens

The L-rule and the U-rule can be used to evaluate objects, possibly not belonging to U, by means of a proper generalization of *modus ponens* (MP) and *modus tollens* (MT) in order to infer a conclusion from gradual rules. Classically, the MP has the following form:

$$\begin{array}{lll} \text{if} & X \to Y & \text{is true} \\ \text{and} & X & \text{is true} \\ \hline \text{then} & Y & \text{is true} \end{array}$$

MP has the following interpretation: assuming an implication $X \to Y$ (decision rule) and a fact X (premise), we obtain another fact Y (conclusion). If we replace the classical decision rule above by our L-rules and U-rules, then we obtain the following two generalized fuzzy-rough MP:

$$\left.\begin{array}{ll} \text{if} & \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}, \text{ and } \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\} \\ & \quad \to \ \mu_Y(x) \geqslant f(\boldsymbol{\alpha}) \quad [\boldsymbol{\alpha} = (\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})] \\ \text{and} & \mu_{X_h}(x) \geqslant \alpha'_h \text{ for each } X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}, \text{ and } \mu_{X_h}(x) \leqslant \alpha'_h \text{ for each } X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\} \\ \hline \text{then} & \mu_Y(x) \geqslant f(\boldsymbol{\alpha}') \quad [\boldsymbol{\alpha}' = (\alpha'_{i1}, \ldots, \alpha'_{ir}, \alpha'_{j1}, \ldots, \alpha'_{js})], \end{array}\right\}$$

$$\left.\begin{array}{ll} \text{if} & \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}, \text{ and } \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\} \\ & \quad \to \ \mu_Y(x) \leqslant g(\boldsymbol{\alpha}) \quad [\boldsymbol{\alpha} = (\alpha_{i1}, \ldots, \alpha_{ir}, \alpha_{j1}, \ldots, \alpha_{js})] \\ \text{and} & \mu_{X_h}(x) \leqslant \alpha'_h \text{ for each } X_h \in \{X^{\uparrow}_{i1}, \ldots, X^{\uparrow}_{ir}\}, \text{ and } \mu_{X_h}(x) \geqslant \alpha'_h \text{ for each } X_h \in \{X^{\downarrow}_{j1}, \ldots, X^{\downarrow}_{js}\} \\ \hline \text{then} & \mu_Y(x) \leqslant g(\boldsymbol{\alpha}') \quad [\boldsymbol{\alpha}' = (\alpha'_{i1}, \ldots, \alpha'_{ir}, \alpha'_{j1}, \ldots, \alpha'_{js})]. \end{array}\right\}$$

Classically, the MT has the following form:

$$\begin{array}{lll} \text{if} & X \to Y & \text{is true} \\ \text{and} & Y & \text{is false} \\ \hline \text{then} & X & \text{is false} \end{array}$$

MT has the following interpretation: assuming an implication $X \to Y$ (decision rule) and a fact *not* Y (premise), we obtain another fact *not* X (conclusion). If we replace the classical decision rule above by our L-rules and U-rules, then we obtain the following two generalized fuzzy-rough MT:

$$\left.\begin{array}{l}\text{if} \quad \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\} \\ \quad \to \ \mu_Y(x) \geqslant f(\boldsymbol{\alpha}) \quad [\boldsymbol{\alpha} = (\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})] \\ \text{and} \ \ \mu_Y(x) < f(\boldsymbol{\alpha}') \quad [\boldsymbol{\alpha}' = (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js})] \\ \hline \text{then} \ \ \mu_{X_h}(x) < \alpha'_h \text{ for at least one } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ or} \\ \qquad \mu_{X_h}(x) > \alpha'_h \text{ for at least one } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}. \end{array}\right\}$$

$$\left.\begin{array}{l}\text{if} \quad \mu_{X_h}(x) \leqslant \alpha_h \text{ for each } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ and } \mu_{X_h}(x) \geqslant \alpha_h \text{ for each } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\} \\ \quad \to \ \mu_Y(x) \leqslant g(\boldsymbol{\alpha}) \quad [\boldsymbol{\alpha} = (\alpha_{i1},\ldots,\alpha_{ir},\alpha_{j1},\ldots,\alpha_{js})] \\ \text{and} \ \ \mu_Y(x) > g(\boldsymbol{\alpha}')[\boldsymbol{\alpha}' = (\alpha'_{i1},\ldots,\alpha'_{ir},\alpha'_{j1},\ldots,\alpha'_{js})] \\ \hline \text{then} \ \ \mu_{X_h}(x) > \alpha'_h \text{ for at least one } X_h \in \{X_{i1}^{\uparrow},\ldots,X_{ir}^{\uparrow}\}, \text{ or} \\ \qquad \mu_{X_h}(x) < \alpha'_h \text{ for at least one } X_h \in \{X_{j1}^{\downarrow},\ldots,X_{js}^{\downarrow}\}. \end{array}\right\}$$

Example 1 (*part* 10). Let us consider L-rule $LR^{\#} = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^{\#}\rangle$ presented in part 9. Let us also consider car x such that $\mu_{\text{speedy\_cars}}(x) = 0.7$ and $\mu_{\text{expensive\_cars}}(x) = 0.3$. Applying L-rule $LR^{\#}$ to car x we obtain the following generalized fuzzy-rough MP:

$$\left.\begin{array}{l}\text{if} \quad \mu_{\text{speedy\_cars}}(x) \geqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(x) \leqslant \alpha_{\text{expensive\_cars}} \\ \quad \to \ \mu_{\text{good\_cars}}(x) \geqslant f^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \\ \text{and} \ \ \mu_{\text{speedy\_cars}}(x) \geqslant 0.7 \text{ and } \mu_{\text{expensive\_cars}}(x) \leqslant 0.3 \\ \hline \text{then} \ \ \mu_{\text{good\_cars}}(x) \geqslant f^{\#}(0.7, 0.3) = 0.33. \end{array}\right\}$$

Let us consider now U-rule $UR^{\#} = X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g^{\#}\rangle$ presented in part 9 too. Let us consider again car x. Applying U-rule $UR^{\#}$ to car x we obtain the following generalized fuzzy-rough MP:

$$\left.\begin{array}{l}\text{if} \quad \mu_{\text{speedy\_cars}}(x) \leqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(x) \geqslant \alpha_{\text{expensive\_cars}} \\ \quad \to \ \mu_{\text{good\_cars}}(x) \leqslant \ g^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \\ \text{and} \ \ \mu_{\text{speedy\_cars}}(x) \leqslant 0.7 \text{ and } \mu_{\text{expensive\_cars}}(x) \geqslant 0.3 \\ \hline \text{then} \ \ \mu_{\text{good\_cars}}(x) \leqslant \ g^{\#}(0.7, 0.3) = 0.66. \end{array}\right\}$$

Now, let us also consider car y such that $\mu_{\text{good\_cars}}(y) = 0.3$. Applying above L-rule $LR^{\#} = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^{\#}\rangle$ to car y we obtain the following generalized fuzzy-rough MT:

$$\left.\begin{array}{l}\text{if} \quad \mu_{\text{speedy\_cars}}(y) \geqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(y) \leqslant \alpha_{\text{expensive\_cars}} \\ \quad \to \ \mu_{\text{good\_cars}}(y) \geqslant f^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \\ \text{and} \ \ \mu_{\text{good\_cars}}(y) = 0.3 < f^{\#}(0.7, 0.4) = 0.33 \\ \hline \text{then} \ \ \mu_{\text{speedy\_cars}}(y) < 0.7 \text{ or } \mu_{\text{expensive\_cars}}(y) > 0.4. \end{array}\right\}$$

Applying U-rule $UR^{\#} = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^{\#}\rangle$ to car z such that $\mu_{\text{good\_cars}}(z) = 0.8$. we obtain the following generalized fuzzy-rough MT:

$$\left.\begin{array}{ll} \text{if} & \mu_{\text{speedy\_cars}}(z) \leqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(z) \geqslant \alpha_{\text{expensive\_cars}} \\ & \quad \rightarrow \ \mu_{\text{good\_cars}}(z) \leqslant \ f^{\#}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \\ \text{and} & \mu_{\text{good\_cars}}(z) = 0.8 > g^{\#}(0.6, 0.2) = 0.66 \\ \hline \text{then} & \mu_{\text{speedy\_cars}}(z) > 0.6 \text{ or } \mu_{\text{expensive\_cars}}(z) < 0.2. \end{array}\right\}$$

Example 2. In this example we show that all the concepts introduced until now can also be applied to finite sets of objects. In a certain sense this is a more natural application of the introduced concepts. An infinite universe of discourse as that one considered in Example 1, is interesting for didactic reasons, but it is not appropriate for real life applications. Therefore, let us suppose that, more realistically, U is a finite set of cars described in Table 1.

Let us approximate knowledge contained in $Y_{\text{good\_cars}}$ using knowledge about $X^{\uparrow}_{\text{speedy\_cars}}$ and $X^{\downarrow}_{\text{expensive\_cars}}$ under the hypothesis that membership in $X^{\uparrow}_{\text{speedy\_cars}}$ is positively related and membership in $X^{\downarrow}_{\text{expensive\_cars}}$ is negatively related with membership in $Y_{\text{good\_cars}}$. The results of the approximations are in Table 2.

On the basis of the rough approximations presented in Table 2 we can induce the maximally prudent L-rule with

$$LR^{\#*} = \langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, f^{\#*}\rangle,$$

and the maximally prudent U-rule with

$$UR^{\#*}\langle X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}, g^{\#*}\rangle,$$

Table 1
Data table about cars

Car	$\mu_{\text{speedy\_cars}}(\cdot)$	$\mu_{\text{expensive\_cars}}(\cdot)$	$\mu_{\text{good\_cars}}(\cdot)$
C1	0.9	0.8	0.4
C2	0.7	0.5	0.7
C3	0.5	0.3	0.5
C4	0.4	0.4	0.6
C5	0.8	0.2	0.8

Table 2
Rough approximations

Car	$\mu[\underline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), \cdot]$	$\mu[\overline{App}(X^{\uparrow}_{\text{speedy\_cars}}, X^{\downarrow}_{\text{expensive\_cars}}, Y_{\text{good\_cars}}), \cdot]$
C1	0.4	0.4
C2	0.7	0.7
C3	0.5	0.6
C4	0.5	0.6
C5	0.8	0.8

where

$$f^{\#*}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})=\begin{cases}0.4 & \text{if } \alpha_{\text{speedy\_cars}}\geqslant 0.9 \text{ and } 0.5<\alpha_{\text{expensive\_cars}}\leqslant 0.8,\\ 0.5 & \text{if } 0.4\leqslant \alpha_{\text{speedy\_cars}}<0.7 \text{ and } \alpha_{\text{expensive\_cars}}\leqslant 0.4,\\ 0.7 & \text{if } \alpha_{\text{speedy\_cars}}\geqslant 0.7 \text{ and } 0.2<\alpha_{\text{expensive\_cars}}\leqslant 0.5,\\ 0.7 & \text{if } 0.7\leqslant \alpha_{\text{speedy\_cars}}<0.8 \text{ and } \alpha_{\text{expensive\_cars}}\leqslant 0.2,\\ 0.8 & \text{if } \alpha_{\text{speedy\_cars}}\geqslant 0.8 \text{ and } \alpha_{\text{expensive\_cars}}\leqslant 0.2,\\ 0 & \text{otherwise}\end{cases}$$

and

$$g^{\#*}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})=\begin{cases}0.4 & \text{if } \alpha_{\text{speedy\_cars}}\leqslant 0.9 \text{ and } \alpha_{\text{expensive\_cars}}\geqslant 0.8,\\ 0.6 & \text{if } \alpha_{\text{speedy\_cars}}\leqslant 0.5 \text{ and } 0.3\leqslant\alpha_{\text{expensive\_cars}}<0.8,\\ 0.7 & \text{if } 0.5<\alpha_{\text{speedy\_cars}}\leqslant 0.7 \text{ and } 0.5\leqslant\alpha_{\text{expensive\_cars}}<0.8,\\ 0.8 & \text{if } 0.5<\alpha_{\text{speedy\_cars}}\leqslant 0.8 \text{ and } 0.2\leqslant\alpha_{\text{expensive\_cars}}<0.5,\\ 0.8 & \text{if } 0.7<\alpha_{\text{speedy\_cars}}\leqslant 0.8 \text{ and } 0.5\leqslant\alpha_{\text{expensive\_cars}}<0.8,\\ 0.8 & \text{if } \alpha_{\text{speedy\_cars}}\leqslant 0.5 \text{ and } 0.2\leqslant\alpha_{\text{expensive\_cars}}<0.3,\\ 1 & \text{otherwise.}\end{cases}$$

Let us also consider car w such that $\mu_{\text{speedy\_cars}}(w)=0.6$ and $\mu_{\text{expensive\_cars}}(w)=0.3$. Applying L-rule $LR^{\#*}$ to car w, we obtain the following generalized fuzzy-rough MP:

$$\left.\begin{array}{ll}\text{if} & \mu_{\text{speedy\_cars}}(w)\geqslant\alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(w)\leqslant\alpha_{\text{expensive\_cars}}\\ & \quad\to\ \mu_{\text{good\_cars}}(w)\geqslant f^{\#*}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})\\ \text{and} & \mu_{\text{speedy\_cars}}(w)\geqslant 0.6 \text{ and } \mu_{\text{expensive\_cars}}(w)\leqslant 0.3\\ \hline \text{then} & \mu_{\text{good\_cars}}(w)\geqslant f^{\#*}(0.6,0.3)=0.5.\end{array}\right\}$$

Let us consider in turn U-rule $UR^{\#*}=\langle X^{\uparrow}_{\text{speedy\_cars}},X^{\downarrow}_{\text{expensive\_cars}},Y_{\text{good\_cars}},g^{\#*}\rangle$ and car w. Applying U-rule $UR^{\#*}$ to car w we obtain the following generalized fuzzy-rough MP:

$$\left.\begin{array}{ll}\text{if} & \mu_{\text{speedy\_cars}}(w)\leqslant\alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(w)\geqslant\alpha_{\text{expensive\_cars}}\\ & \quad\to\ \mu_{\text{good\_cars}}(w)\leqslant\ g^{\#*}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})\\ \text{and} & \mu_{\text{speedy\_cars}}(w)\leqslant 0.6 \text{ and } \mu_{\text{expensive\_cars}}(w)\geqslant 0.3\\ \hline \text{then} & \mu_{\text{good\_cars}}(w)\leqslant g^{\#*}(0.6,0.3)=0.8.\end{array}\right\}$$

Now, let us also consider another car z such that $\mu_{\text{good\_cars}}(z)=0.55$. Applying above L-rule $LR^{\#*}=\langle X^{\uparrow}_{\text{speedy\_cars}},X^{\downarrow}_{\text{expensive\_cars}},Y_{\text{good\_cars}},f^{\#*}\rangle$ to car z, we obtain the following generalized fuzzy-rough MT:

$$\left.\begin{array}{ll}\text{if} & \mu_{\text{speedy\_cars}}(z)\geqslant\alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(z)\leqslant\alpha_{\text{expensive\_cars}}\\ & \quad\to\ \mu_{\text{good\_cars}}(z)\geqslant f^{\#*}(\alpha_{\text{speedy\_cars}},\alpha_{\text{expensive\_cars}})\\ \text{and} & \mu_{\text{good\_cars}}(z)=0.55<f^{\#*}(0.8,0.3)=0.7\\ \hline \text{then} & \mu_{\text{speedy\_cars}}(z)<0.8 \text{ or } \mu_{\text{expensive\_cars}}(z)>0.3.\end{array}\right\}$$

Applying U-rule $UR^{\#}=\langle X^{\uparrow}_{\text{speedy\_cars}},X^{\downarrow}_{\text{expensive\_cars}},Y_{\text{good\_cars}},g^{\#*}\rangle$ to car z such that $\mu_{\text{good\_cars}}(z)=0.55$, we obtain the following generalized fuzzy-rough MT:

$$\left.\begin{array}{ll} \text{if} & \mu_{\text{speedy\_cars}}(z) \leqslant \alpha_{\text{speedy\_cars}} \text{ and } \mu_{\text{expensive\_cars}}(z) \geqslant \alpha_{\text{expensive\_cars}} \\ & \quad \rightarrow \mu_{\text{good\_cars}}(z) \leqslant g^{\#*}(\alpha_{\text{speedy\_cars}}, \alpha_{\text{expensive\_cars}}) \\ \text{and} & \mu_{\text{good\_cars}}(z) = 0.55 > g^{\#*}(0.80, 0.85) = 0.4 \\ \hline \text{then} & \mu_{\text{speedy\_cars}}(z) > 0.80 \text{ or } \mu_{\text{expensive\_cars}}(z) < 0.85. \end{array}\right\}$$

6. Conclusions and further research directions

In this paper we presented a new fuzzy rough set approach. The main advantage of this new approach is that it infers the most cautious conclusions from available imprecise information, without using neither fuzzy connectives nor specific parameters, whose choice are always subjective to some extent. Another advantage of our approach is that it uses only ordinal properties of membership degrees. We noticed that our approach is related to:

- gradual rules, with respect to syntax and semantics of considered decision rules,
- dominance-based rough set approach, with respect to the idea of monotonic relationship between credibility degrees of multiple premises and conclusion,
- Mill's method of concomitant variation with respect to the philosophy of data mining and knowledge discovery.

We think that this approach gives a new prospect for applications of fuzzy rough approximations in real-world decision problems. More precisely, we envisage the following two extensions of this methodology:

(1) Variable precision fuzzy rough approximation: in this paper we propose to calculate the degree of membership to the fuzzy lower approximation on the basis of non-ambiguous objects only, however, it might be useful in practical applications to allow a limited number of ambiguous objects as well; in this way we may get less specific rules of the type: "the larger the market share of a company, the greater its profit, in l% of the cases", where l is a parameter controlling the proportion of ambiguous objects in the definition of the lower approximation.
(2) Imprecise input data represented by fuzzy numbers and missing values: the evaluation of the objects in the universe U from which the rough approximations and the gradual decision rules are induced may include imprecise values, represented by fuzzy numbers, or missing values.

Acknowledgements

The research of the first author has been supported by the Italian Ministry of Education, University and Scientific Research (MIUR). The second author acknowledges financial support by the Grant-in-Aid for Scientific Research (B) No. 17310098. The third author wishes to acknowledge financial support from the Ministry of Scientific Research and Information Technology.

References

[1] E. Aragones, I. Gilboa, A. Postlewaite, D. Schmeidler, Accuracy vs. simplicity: A complex trade-off, PIER working paper 02-027, 2002.

[2] H. Bensusan, Is machine learning experimental philosophy of science? in: A. Aliseda, D. Pearce (Eds.), ECAI2000 Workshop notes on Scientific Reasoning in Artificial Intelligence and the Philosophy of Science, 2000, pp. 9–14.

[3] T.A.O. Cornish, A.D. Elliman, What has Mill to say about Data Mining? in: Proceedings of the 11th Conference for AI Applications, LA, February 20–22, 1995, IEEE Computer Society Press, 1995, pp. 347–353.

[4] D. Dubois, H. Prade, R. Yager, Information Engineering, John Wiley, New York, 1997.

[5] D. Dubois, H. Prade, Gradual inference rules in approximate reasoning, Information Sciences 61 (1992) 103–122.

[6] D. Dubois, H. Prade, 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.

[7] S. Greco, M. Inuiguchi, R. Slowinski, Rough sets and gradual decision rules, in: G. Wang, O. Lin, Y. Yao, A. Skowron (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Lecture Notes in Computer Science, vol. 2639, Springer-Verlag, Berlin, 2003, pp. 156–164.

[8] S. Greco, M. Inuiguchi, R. Slowinski, A new proposal for rough fuzzy approximations and decision rule representation, in: D. Dubois, J. Grzymala-Busse, M. Inuiguchi, L. Polkowski (Eds.), Transaction on Rough Sets II: Rough Sets and Fuzzy Sets, Lecture Notes in Computer Science, vol. 3135, Springer-Verlag, Berlin, 2004, pp. 156–164.

[9] S. Greco, B. Matarazzo, R. Slowinski, The use of rough sets and fuzzy sets in MCDM, in: T. Gal, T. Stewart, T. Hanne (Eds.), Advances in Multiple Criteria Decision Making, Kluwer Academic Publishers, Boston, 1999, pp. 14.1–14.59 (Chapter 14).

[10] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (1) (2001) 1–47.

[11] S. Greco, B. Matarazzo, R. Slowinski, Decision rule approach, in: J. Figueira, S. Greco, M. Ehrgott (Eds.), Multiple Criteria Decision Analysis, State of the art Surveys, Springer, 2005, pp. 507–561 (Chapter 13).

[12] S. Greco, Z. Pawlak, R. Slowinski, Can Bayesian confirmation measures be useful for rough set decision rules? Engineering Applications of Artificial Intelligence 17 (2004) 345–361.

[13] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publishers, Dordrecht, 2000.

[14] T. Marchant, The measurement of membership by comparisons, Fuzzy Sets and Systems 148 (2004) 157–177.

[15] J.S. Mill, A system of logic, ratiocinative and inductive: being a connected view of the principles of evidence and the methods of scientific investigation, in: J.M. Robson (Ed.), Collected Works, vol. VII, Books l-m, University of Toronto Press, RKP, 1843/1973.

[16] A. Nakamura, J.M. Gao, A logic for fuzzy data analysis, Fuzzy Sets and Systems 39 (1991) 127–132.

[17] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.

[18] Z. Pawlak, Rough Sets, Kluwer, Dordrecht, 1991.

[19] J. Williamson, A dynamic interaction between machine learning and the philosophy of science, Minds and Machine 14 (4) (2004) 539–549.

[20] L. Wittgenstein, Tractatus Logico-Philosophicus, Routledge and Kegan Paul, London, 1922, fifth impression, 1951.

[21] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.