Variable Consistency Model of Dominance-based Rough Sets Approach

S. Greco[1], B. Matarazzo[1], R. Slowinski[2], J. Stefanowski[2]

[1]Faculty of Economics, University of Catania, 95129 Catania, Italy

[2]Institute of Computing Science, Poznan University of Technology, 60965 Poznan, Poland

Abstract. Consideration of preference-orders requires the use of an extended rough set model called Dominance-based Rough Set Approach (DRSA). The rough approximations defined within DRSA are based on consistency in the sense of dominance principle. It requires that objects having not-worse evaluation with respect to a set of considered criteria than a referent object cannot be assigned to a worse class than the referent object. However, some inconsistencies may decrease the cardinality of lower approximations to such an extent that it is impossible to discover strong patterns in the data, particularly when data sets are large. Thus, a relaxation of the strict dominance principle is worthwhile. The relaxation introduced in this paper to the DRSA model admits some inconsistent objects to the lower approximations; the range of this relaxation is controlled by an index called consistency level. The resulting model is called variable-consistency model (VC-DRSA). We concentrate on the new definitions of rough approximations and their properties, and we propose a new syntax of decision rules characterized by a confidence degree not less than the consistency level. The use of VC-DRSA is illustrated by an example of customer satisfaction analysis referring to an airline company.

1. Introduction

Rough sets theory introduced by Pawlak [6] is an approach for analysing information about objects described by attributes. It is particularly useful to deal with inconsistencies of input information caused by its granularity. The original rough set approach does not consider, however, the attributes with preference-ordered domains, i.e. *criteria*. Nevertheless, in many real-life problems the *ordering properties* of the considered attributes play an important role. For instance, such features of objects as product quality, market share, debt ratio are typically treated as criteria in economical problems. Motivated by this observation, Greco, Matarazzo and Slowinski [1,3] proposed a generalisation of the rough set approach to problems where ordering properties should be taken into account. Similarly to the original rough sets, this approach is based on approximations of partitions of the objects into pre-defined categories, however, differently to the original model, the categories are ordered from the best to the worst and the approximations are constructed using a *dominance relation* instead of an indiscernibility relation. The considered dominance relation is built on the basis of the information supplied by criteria. The new *Dominance-based Rough Set Approach* (DRSA) was applied to solve typical problems of *Multiple-Criteria Decision Aiding* (MCDA), i.e. choice, ranking and sorting (see e.g. [1,3]).

In this paper, we consider a variant of DRSA used to *multiple-criteria sorting problems*, which concerns an assignment of objects evaluated by a set of criteria to some pre-defined and preference-ordered decision classes. In this variant of DRSA, the sets to be approximated with the dominance relation are, so-called, *upward* and *downward unions of decision classes*. There are known encouraging results of its applications, e.g. to evaluation of bankruptcy risk [2].

The analysis of large real-life data tables shows, however, that for some multiple-criteria sorting problems the application of DRSA identifies large differences between lower and upper approximations of the unions of decision classes and, moreover, rather weak decision rules, i.e. supported by few objects from lower approximations. The reason is that inconsistency, in the sense of dominance principle, between objects x and y assigned to very distant classes, h and t, respectively, (x dominates y, while class h is worse than t) causes inconsistency (ambiguity) also with all objects belonging to intermediate classes (from h to t) and dominated by x. In such cases it seems reasonable to relax the conditions for assignment of objects to lower approximations of the unions of decision classes. Classically, only non-ambiguous objects can be included in lower approximations. The relaxation will admit some ambiguous objects as well; the range of this ambiguity will be controlled by an index called *consistency level*. The aim of this article is to present a generalization of DRSA to variable consistency model (VC-DRSA).

This kind of relaxation has been already considered within the classical indiscernibility-based rough set approach, by means of so-called *variable precision rough set model* (VPRS) [11]. VPRS allows defining lower approximations accepting a limited number of counterexamples controlled by pre-defined level of certainty.

The paper is organized as follows. In section 2, main concepts of VC-DRSA are introduced, including rough approximations, approximation measures and decision rules. An illustrative example presented in section 3 refers to a real problem of customer satisfaction analysis in an airline company. The final section groups conclusions.

2. Variable Consistency Dominance-based Rough Set Approach (VC-DRSA)

For algorithmic reasons, information about objects is represented in the form of an information table. The rows of the table are labelled by *objects*, whereas columns are labelled by *attributes* and entries of the table are *attribute-values*. Formally, by an *information table* we understand the 4-tuple $S=\langle U,Q,V,f\rangle$, where U is a finite set of objects, Q is a finite set of *attributes*, $V = \bigcup_{q\in Q} V_q$ and V_q is a domain of the attribute q, and $f: U\times Q\rightarrow V$ is a total function such that $f(x,q)\in V_q$ for every $q\in Q$, $x\in U$, called an *information function* [6]. The set Q is, in general, divided into set C of *condition attributes* and set D of *decision attributes*.

Assuming that all condition attributes $q \in C$ are *criteria*, let $\succeq_q$ be an *weak preference relation* on U with respect to criterion q such that $x \succeq_q y$ means "x is at least as good as y with respect to criterion q". We suppose that $\succeq_q$ is a total preorder, i.e. a strongly complete and transitive binary relation, defined on U on the basis of evaluations $f(\cdot,q)$.

Furthermore, assuming that the set of decision attributes D (possibly a singleton $\{d\}$) makes a partition of U into a finite number of decision classes, let $\boldsymbol{Cl}=\{Cl_t, t \in T\}$, $T=\{1,...,n\}$, be a set of these classes such that each $x \in U$ belongs to one and only one class $Cl_t \in \boldsymbol{Cl}$. We suppose that the classes are preference-ordered, i.e. for all $r,s \in T$, such that $r>s$, the objects from Cl_r are preferred to the objects from Cl_s. The above assumptions are typical for consideration of a *multiple-criteria sorting problem*.

The sets to be approximated are called *upward union* and *downward union* of classes, respectively:

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

The statement $x \in Cl_t^{\geq}$ means "x belongs at least to class Cl_t", while $x \in Cl_t^{\leq}$ means "x belongs at most to class Cl_t".

Let us remark that $Cl_1^{\geq} = Cl_n^{\leq} = U$, $Cl_n^{\geq} = Cl_n$ and $Cl_1^{\leq} = Cl_1$. Furthermore, for $t=2,...,n$, we have:

$$Cl_{t-1}^{\leq} = U - Cl_t^{\geq} \quad \text{and} \quad Cl_t^{\geq} = U - Cl_{t-1}^{\leq}.$$

The key idea of rough sets is approximation of one knowledge by another knowledge. In classical rough set approach (CRSA), the knowledge approximated is a partition of U into classes generated by a set of decision attributes; the knowledge used for approximation is a partition of U into elementary sets of objects that are indiscernible by a set of condition attributes. The elementary sets are seen as "*granules of knowledge*" used for approximation.

In DRSA approach, where condition attributes are criteria and classes are preference-ordered, the knowledge approximated is a collection of *upward* and *downward unions of classes* and the "granules of knowledge" are sets of objects defined using a dominance relation instead of an indiscernibility relation. This is the main difference between CRSA and DRSA. Let us define now the dominance relation.

We say that x *dominates* y with respect to $P \subseteq C$, denoted by xD_Py, if $x \succeq_q y$ for all $q \in P$. Given $P \subseteq C$ and $x \in U$, the "granules of knowledge" used for approximation in DRSA are:

- a set of objects dominating x, called *P-dominating set*, $D_P^+(x)=\{y \in U: yD_Px\}$,
- a set of objects dominated by x, called *P-dominated set*, $D_P^-(x)=\{y \in U: xD_Py\}$.

For any $P \subseteq C$ we say that $x \in U$ belongs to $Cl_t^{\geq}$ *with no ambiguity at consistency level* $l \in (0, 1]$, if $x \in Cl_t^{\geq}$ and at least $l*100\%$ of all objects $y \in U$ dominating x with respect to P also belong to $Cl_t^{\geq}$, i.e.

$$\frac{card\left(D_P^+(x) \cap Cl_t^{\geq}\right)}{card\left(D_P^+(x)\right)} \geq l.$$

The level l is called *consistency level* because it controls the degree of consistency between objects qualified as belonging to $Cl_t^{\geq}$ without any ambiguity. In other words, if $l<1$, then $(1-l)*100\%$ of all objects $y \in U$ dominating x with respect to P do not belong to $Cl_t^{\geq}$ and thus contradict the inclusion of x in $Cl_t^{\geq}$.

Analogously, for any $P \subseteq C$ we say that $x \in U$ belongs to $Cl_t^{\leq}$ *with no ambiguity at consistency level* $l \in (0, 1]$, if $x \in Cl_t^{\leq}$ and at least $l*100\%$ of all the objects $y \in U$ dominated by x with respect to P also belong to $Cl_t^{\leq}$, i.e.

$$\frac{card\left(D_P^-(x) \cap Cl_t^{\leq}\right)}{card\left(D_P^-(x)\right)} \geq l.$$

Thus, for any $P \subseteq C$, each object $x \in U$ is either ambiguous or non-ambiguous at consistency level l with respect to the upward union $Cl_t^{\geq}$ (t=2,...,n) or with respect to the downward union $Cl_t^{\leq}$ (t=1,...,n -1).

The concept of non-ambiguous objects at some consistency level l leads naturally to the definition of P-lower approximations of the unions of classes $Cl_t^{\geq}$ and $Cl_t^{\leq}$.

$$\underline{P}^l\left(Cl_t^{\geq}\right)=\{x \in Cl_t^{\geq} : \frac{card\left(D_P^+(x) \cap Cl_t^{\geq}\right)}{card\left(D_P^+(x)\right)} \geq l\},\ \underline{P}^l\left(Cl_t^{\leq}\right)=\{x \in Cl_t^{\leq} : \frac{card\left(D_P^-(x) \cap Cl_t^{\leq}\right)}{card\left(D_P^-(x)\right)} \geq l\}.$$

Given $P \subseteq C$ and consistency level l, we can define the P-*upper approximations* of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, denoted by $\overline{P}^l\left(Cl_t^{\geq}\right)$ and $\overline{P}^l\left(Cl_t^{\leq}\right)$, by complementation of $\underline{P}^l\left(Cl_{t-1}^{\leq}\right)$ and $\underline{P}^l\left(Cl_{t+1}^{\geq}\right)$ with respect to U:

$$\overline{P}^l\left(Cl_t^{\geq}\right)=U - \underline{P}^l\left(Cl_{t-1}^{\leq}\right),\ \overline{P}^l\left(Cl_t^{\leq}\right)=U - \underline{P}^l\left(Cl_{t+1}^{\geq}\right).$$

$\overline{P}^l\left(Cl_t^{\geq}\right)$ can be interpreted as the set of all the objects belonging to $Cl_t^{\geq}$, *possibly ambiguous* at consistency level l. Analogously, $\overline{P}^l\left(Cl_t^{\leq}\right)$ can be interpreted as the set of all the objects belonging to $Cl_t^{\leq}$, *possibly ambiguous* at consistency level l. The P-*boundaries* (P-doubtful regions) of $Cl_t^{\geq}$ and $Cl_t^{\leq}$ are defined as:

$$Bn_P(Cl_t^{\geq})=\overline{P}^l\left(Cl_t^{\geq}\right)-\underline{P}^l\left(Cl_t^{\geq}\right), \quad Bn_P(Cl_t^{\leq})=\overline{P}^l\left(Cl_t^{\leq}\right)-\underline{P}^l\left(Cl_t^{\leq}\right), \quad \text{for } t=1,...,n.$$

The *variable consistency* model of the dominance-based rough set approach provides some degree of flexibility in assigning objects to lower and upper approximations of the unions of decision classes. It can easily be demonstrated that for $0<l'<l\leq 1$ and t=2,...,n,

$$\underline{P}^l\left(Cl_t^{\geq}\right) \subseteq \underline{P}^{l'}\left(Cl_t^{\geq}\right) \quad \text{and} \quad \overline{P}^{l'}\left(Cl_t^{\geq}\right) \subseteq \overline{P}^l\left(Cl_t^{\geq}\right).$$

The *variable consistency* model is inspired by Ziarko's model of the *variable precision* rough set approach [11,12], however, there is a significant difference in the definition of rough approximations because $\underline{P}^l\left(Cl_t^{\geq}\right)$ and $\overline{P}^l\left(Cl_t^{\geq}\right)$ are composed of non-ambiguous and ambiguous objects at consistency level *l*, respectively, while Ziarko's $\underline{P}^l\left(Cl_t\right)$ and $\overline{P}^l\left(Cl_t\right)$ are composed of *P*-indiscernibility sets such that at least *l*∗100% of these sets are included in Cl_t or have an non-empty intersection with Cl_t, respectively. If one would like to use Ziarko's definition of variable precision rough approximations in the context of multiple-criteria sorting, then the *P*-indiscernibility sets should be substituted by *P*-dominating sets $D_P^+(x)$, however, then the notion of ambiguity that naturally leads to the general definition of rough approximations (see [9]) looses its meaning. Moreover, bad side effect of a direct use of Ziarko's definition is that a lower approximation $\underline{P}^l\left(Cl_t^{\geq}\right)$ may include objects *y* assigned to Cl_h, where *h* is much less than *t*, if *y* belongs to $D_P^+(x)$ that was included in $\underline{P}^l\left(Cl_t^{\geq}\right)$. When the decision classes are preference ordered, it is reasonable to expect that objects assigned to far worse classes than the considered union are not counted to the lower approximation of this union.

Furthermore, the following properties can be proved:

1) $$\overline{P}^l\left(Cl_t^{\geq}\right) = Cl_t^{\geq} \cup \{x \in Cl_{t-1}^{\leq} : \frac{card\left(D_P^-(x) \cap Cl_{t-1}^{\leq}\right)}{card\left(D_P^-(x)\right)} < l\} =$$

$$= Cl_t^{\geq} \cup \{x \in Cl_t^{\leq} : \frac{card\left(D_P^-(x) \cap Cl_t^{\geq}\right)}{card\left(D_P^-(x)\right)} > 1\text{-}l\},$$

$$\overline{P}^l\left(Cl_t^{\leq}\right) = Cl_t^{\leq} \cup \{x \in Cl_{t+1}^{\geq} : \frac{card\left(D_P^-(x) \cap Cl_{t+1}^{\geq}\right)}{card\left(D_P^-(x)\right)} < l\} =$$

$$= Cl_t^{\leq} \cup \{x \in Cl_t^{\geq} : \frac{card\left(D_P^+(x) \cap Cl_t^{\leq}\right)}{card\left(D_P^+(x)\right)} > 1\text{-}l\},$$

2) $\underline{P}^l\left(Cl_t^{\geq}\right) \subseteq Cl_t^{\geq} \subseteq \overline{P}^l\left(Cl_t^{\geq}\right)$, $\underline{P}^l\left(Cl_t^{\leq}\right) \subseteq Cl_t^{\leq} \subseteq \overline{P}^l\left(Cl_t^{\leq}\right)$.

Due to complementarity of the rough approximations [3], also the following property holds:

$$Bn_P(Cl_t^{\geq}) = Bn_P(Cl_{t-1}^{\leq}), \text{ for } t=2,...,n, \text{ and } Bn_P(Cl_t^{\leq}) = Bn_P(Cl_{t+1}^{\geq}), \text{ for } t=1,...,n-1.$$

For every $t \in T$ and for every $P \subseteq C$ we define the *quality of approximation of partition* ***Cl*** by set of criteria *P*, or in short, *quality of sorting*:

$$\gamma_P(\boldsymbol{Cl}) = \frac{\text{card}\left(U - \left(\bigcup_{t\in T} Bn_P\left(Cl_t^{\leq}\right)\right)\right)}{card\,(U)} = \frac{\text{card}\left(U - \left(\bigcup_{t\in T} Bn_P\left(Cl_t^{\geq}\right)\right)\right)}{card\,(U)}.$$

The quality expresses the ratio of all P-correctly sorted objects to all objects in the table.

Each minimal subset $P\subseteq C$ such that $\gamma_P(\boldsymbol{Cl})=\gamma_C(\boldsymbol{Cl})$ is called a *reduct* of $\boldsymbol{Cl}$ and denoted by $RED_{\boldsymbol{Cl}}$. Let us remark that an information table can have more than one reduct. The intersection of all reducts is called the *core* and denoted by $CORE_{\boldsymbol{Cl}}$.

Let us remind that the dominance-based rough approximations of upward and downward unions of classes can serve to induce a generalized description of objects contained in the information table in terms of "*if..., then...*" decision rules. For a given upward or downward union of classes, $Cl_t^{\geq}$ or $Cl_s^{\leq}$, the decision rules induced under a hypothesis that objects belonging to $\underline{P}^l\left(Cl_t^{\geq}\right)$ (or $\underline{P}^l\left(Cl_s^{\leq}\right)$) are *positive* and all the others *negative*, suggest an assignment to "at least class Cl_t" (or to "at most class Cl_s"). They are called $D_{\geq}$- (or $D_{\leq}$) *certain decision rules* because they assign objects to classes without any ambiguity. Next, if upper approximations differ from lower approximations, another kind of decision rules can be induced under the hypothesis that objects belonging to $\overline{P}^l\left(Cl_t^{\geq}\right)$ (or to $\overline{P}^l\left(Cl_s^{\leq}\right)$) are *positive* and all the others *negative*. These rules are called $D_{\geq}$- (or $D_{\leq}$) *possible decision* rules suggesting that an object *could belong* to "at least class $Cl_t^{\geq}$" (or "at most class $Cl_s^{\leq}$"). Yet another option is to induce $D_{\geq\leq}$-approximate decision rules from the intersection $\overline{P}^l(Cl_s^{\leq})\cap\overline{P}^l(Cl_t^{\geq})$ instead of possible rules. For more discussion see [8].

Within VC-DRSA, decision rules are induced from examples belonging to extended approximations. So, it is necessary to assign to each decision rules an additional parameter α, called *confidence of the rule*. It controls the discrimination ability of the rule.

Assuming that for each $q\in C$, $V_q\subseteq\mathbf{R}$ (i.e. V_q is quantitative) and that for each $x,y\in U$, $f(x,q)\geq f(y,q)$ implies $x\succeq_q y$ (i.e. V_q is preference-ordered), the following two basic types of variable-consistency decision rules can be considered:

1) $D_{\geq}$-*decision rules* with the following syntax:

if $f(x,q_1)\geq r_{q1}$ *and* $f(x,q_2)\geq r_{q2}$ *and* ...$f(x,q_p)\geq r_{qp}$, *then* $x\in Cl_t^{\geq}$ with confidence α,

where $P=\{q_1,...,q_p\}\subseteq C$, $(r_{q1},...,r_{qp})\in V_{q1}\times V_{q2}\times...\times V_{qp}$ and $t\in T$;

2) $D_{\leq}$-*decision rules* with the following syntax:

if $f(x,q_1)\leq r_{q1}$ *and* $f(x,q_2)\leq r_{q2}$ *and* ... $f(x,q_p)\leq r_{qp}$, *then* $x\in Cl_t^{\leq}$ with confidence α,

where $P=\{q_1,...,q_p\}\subseteq C$, $(r_{q1},...,r_{qp})\in V_{q1}\times V_{q2}\times...\times V_{qp}$ and $t\in T$;

We say that an object *supports* a decision rule if it matches both condition and decision parts of the rule. On the other hand, an object is *covered* by a decision rule if it matches the condition part of the rule. More formally, given a $D_{\geq}$-rule ρ : *if* $f(x,q_1)\geq r_{q1}$ *and* $f(x,q_2)\geq r_{q2}$ *and* $...f(x,q_p)\geq r_{qp}$, *then* $x\in Cl_t^{\geq}$, an object $y\in U$ supports decision rule ρ iff $f(y,q_1)\geq r_{q1}$ and $f(y,q_2)\geq r_{q2}$ and ... $f(y,q_p)\geq r_{qp}$ and $y\in Cl_t^{\geq}$, while y is covered by ρ iff $f(y,q_1)\geq r_{q1}$ and $f(y,q_2)\geq r_{q2}$ and ... $f(y,q_p)\geq r_{qp}$. Similar definitions hold for $D_{\leq}$-decision rules.

Let *Cover(ρ)* denote the set of all objects covered by the rule ρ. Thus, the confidence α of $D_{\geq}$-decision rule ρ is defined as: $\frac{card\left(Cover(\rho)\cap Cl_t^{\geq}\right)}{card\left(Cover(\rho)\right)}$. For $D_{\leq}$-decision rule the confidence is defined in a similar way.

Let us remark that the decision rules are induced from P-lower approximations whose composition is controlled by user-specified consistency level l. In consequence, the value of confidence α for the rule should be constrained from the bottom. It seems reasonable to require that the smallest accepted confidence of the rule should not be lower than the currently used consistency level l. Indeed, in the worst case, some objects from the P-lower approximation may create a rule using all criteria from P thus giving a confidence $\alpha\geq l$. The user may have a possibility of increasing this lower bound for confidence of the rule but then decision rules may not cover all objects from the approximations.

Moreover, we require that each decision rule is minimal. Since a decision rule is an implication, by a *minimal* decision rule we understand such an implication that there is no other implication with an antecedent of at least the same weakness (in other words, rule using a subset of elementary conditions or/and weaker elementary conditions) and a consequent of at least the same strength (in other words, rule assigning objects to the same union or sub-union of classes) with a not worse confidence $\alpha\geq l$.

Consider a $D_{\geq}$-decision rule "*if* $f(x,q_1)\geq r_{q1}$ *and* $f(x,q_2)\geq r_{q2}$ *and* $...f(x,q_p)\geq r_{qp}$, *then* $x\in Cl_t^{\geq}$" with confidence α. If there exists an object $y\in \underline{P}^l\left(Cl_t^{\geq}\right)$, $P=\{q_1, q_2, \ldots, q_p\}$ and $l\leq\alpha$, such that $f(y,q_1)=r_{q1}$ *and* $f(y,q_2)=r_{q2}$ *and* $...f(y,q_p)=r_{qp}$, then y is called *basis* of the rule. Each $D_{\geq}$-decision rule having a basis is called *robust* because it is "founded" on an object existing in the data table. Analogous definition of robust decision rules holds for $D_{\leq}$-decision rules.

The induction of variable-consistency decision rules can be done using properly modified algorithms proposed for DRSA. Let us remind that in DRSA, decision rules should have confidence equal to 1. The key modification of rule induction algorithms for VC-DRSA consists in accepting as rules such conjunctions of elementary conditions that yield confidence $\alpha\geq l$. Let us also notice that different strategies of rule induction could be used [10]. For instance, one can wish to induce a minimal and complete set of rules covering all input examples, or all minimal rules, or a subset of rules satisfying some user's pre-defined requirements, e.g. generality or support. The details of one of the rule induction algorithms for VC-DRSA can be found in [4].

3. Illustrative example

Let us illustrate the above concepts on a didactic example. The example refers to a real problem of customer satisfaction analysis [7] in an airline company. The company has diffused a questionnaire to its customers in order to get opinion about the quality of its services. Among the questions of the questionnaire there are three items concerning specific aspects of the aircraft comfort: space for hand luggage (q_1), seat width (q_2) and leg room (q_3). Moreover, there is also a question about an overall evaluation of the aircraft comfort (d). A customer's answer on each of these questions gives an evaluation on a three grade ordinal scale: poor, average, good.

The data table contains 50 objects (questionnaires) described by the set $C=\{q_1, q_2, q_3\}$ of 3 criteria corresponding to the considered aspects of the aircraft comfort and the overall evaluation $D=\{d\}$. All criteria are to be maximized. The scale of criteria is number-coded: 1=poor, 2=average, 3=good. The overall evaluation d creates three decision classes, which are preference-ordered according to increasing class number, i.e. Cl_1=poor, Cl_2=average, Cl_3=good. The analysed data are presented in Table 1.

Table 1. Customer satisfaction data table

Cust.	q_1	q_2	q_3	d	Cust.	q_1	q_2	q_3	d
1	1	3	2	2	26	2	2	1	2
2	1	3	1	1	27	1	2	2	1
3	3	3	1	1	28	3	2	2	2
4	3	3	2	3	29	1	3	2	2
5	3	1	3	1	30	2	3	1	2
6	2	3	1	2	31	1	1	1	1
7	2	1	2	2	32	1	2	2	2
8	1	1	3	2	33	3	1	1	1
9	2	3	3	3	34	2	2	1	2
10	3	3	1	1	35	3	2	1	2
11	1	3	3	2	36	2	2	2	2
12	2	1	1	2	37	1	1	2	1
13	1	1	1	1	38	3	1	2	2
14	1	3	3	3	39	3	3	1	1
15	1	1	1	1	40	1	1	1	1
16	2	2	3	2	41	3	1	1	1
17	2	1	1	2	42	1	2	1	1
18	1	2	3	2	43	1	3	2	2
19	3	2	1	3	44	3	1	2	1
20	2	2	2	1	45	2	2	3	1
21	2	3	3	3	46	3	2	2	1
22	1	2	3	2	47	2	1	3	1
23	3	2	2	2	48	1	2	2	3
24	2	2	1	2	49	2	2	3	3
25	1	3	1	1	50	3	3	3	3

The marketing department of the airline company wants to analyse the influence of the three specific aspects on the overall evaluation of the aircraft comfort. Thus, a sample of questionnaires was analysed using VC-DRSA. As the decision classes are ordered, the following downward and upward unions of classes are to be considered:

at most poor: $Cl_1^{\leq} = \{2,3,5,10,13,15,20,24,25,27,31,33,37,39,40,41,42,44,45,46,47\}$,

at most average: $Cl_2^{\leq} = \{1,2,3,5,6,7,8,10,11,12,13,15,16,17,18,20,22,23,24,25,26,27,$ $28,29,30,31, 32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47\}$;

at least average: $Cl_2^{\geq} = \{1,4,6,7,8,9,11,12,14,16,17,18,19,21,22,23,26,28,29,30,32,$ $34,35,36,38,43,48,49,50\}$,

at least good - $Cl_3^{\geq} = \{4,9,14,19,21,48,49,50\}$.

Let us observe that in the data table there are several inconsistencies. For instance, object #3 dominates object #6, because its evaluations on all criteria q_1, q_2, q_3 are not worse, however, it is assigned to the decision class Cl_1 worse than Cl_2 to which belongs object #6. This means that the customer #3 gave an evaluation for all the considered aspects not worse than the evaluation given by customer # 6 and, on another hand, customer #3 gave an overall evaluation of the aircraft comfort worse than the overall evaluation of customer #6. There are 99 inconsistent pairs in the data table violating the dominance principle in this way.

The data table has been analysed by VC-DRSA assuming the confidence level l=0.8. In this case, the approximations of upward and downward unions of decision classes are the following (the objects present in the lower approximations obtained for confidence level l=1 are in bold):

$\underline{C}^{0.8}(Cl_1^{\leq}) = \{\mathbf{2},\mathbf{13},\mathbf{15},\mathbf{25},\mathbf{31},\mathbf{37},\mathbf{40},\mathbf{42}\}$,

$\overline{C}^{0.8}(Cl_1^{\leq}) = \{2,3,5,6,7,8,10,12,13,15,17,19,20,24,25,26,27,30,31,33,34,35,36,37,38,$ $39,40,41,42,43,44,45,46,47\}$,

$Bn_C^{0.8}(Cl_1^{\leq}) = \{3,5,6,7,8,10,12,17,19,20,24,26,27,30,33,34,35,36,38,39,40,41,43,44,$ $45,46,47\}$;

$\underline{C}^{0.8}(Cl_2^{\leq}) = \{1,\mathbf{2},3,\mathbf{5},\mathbf{6},\mathbf{7},\mathbf{8},10,11,\mathbf{12},\mathbf{13},\mathbf{15},16,\mathbf{17},18,20,22,23,\mathbf{24},\mathbf{25},\mathbf{26},27,28,29,\mathbf{30},\mathbf{31},$ $32,\mathbf{33},\mathbf{34},35,36,\mathbf{37},\mathbf{38},39,\mathbf{40},\mathbf{41}, \mathbf{42},43,\mathbf{44},45,46,\mathbf{47}\}$,

$\overline{C}^{0.8}(Cl_2^{\leq}) = \{1,2,3,5,6,7,8,10,11,12,13,15,16,17,18,19,20,22,23,24,25,26,27,28,29,30,$ $31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49\}$,

$Bn_C^{0.8}(Cl_2^{\leq}) = \{19,48,49\}$;

$\underline{C}^{0.8}(Cl_2^{\geq}) = \{1,\mathbf{4},\mathbf{9},\mathbf{11},\mathbf{14},16,18,\mathbf{21},22,23,28,\mathbf{29},32,48,49,\mathbf{50}\}$,

$\overline{C}^{0.8}(Cl_2^{\geq}) = \{1,3,4,5,6,7,8,9,10,11,12,14,16,17,18,19,20,21,22,23,24,26,272,28,29,30,$ $32,33,34,35,36,38,39,40,41,43, 44,45,46,47, 48,49,50\}$,

$Bn_C^{0.8}(Cl_2^{\geq}) = \{3,5,6,7,8,12,17,19,20,24,26,27,30,33,34,35,36,38,39,40,41,43,$ $44,45,46,47\}$;

$\underline{C}^{0.8}(Cl_3^{\geq}) = \{\mathbf{4},\mathbf{9},14,\mathbf{21},\mathbf{50}\}$,

$\overline{C}^{0.8}(Cl_3^{\geq}) = \{4,9,14,19,21,48,49,50\}$,

$Bn_C^{0.8}(Cl_3^{\geq}) = \{19,48,49\}$.

The set of all robust decision rules having a confidence level α≥0.8 was induced from the above approximations. Let us remark that rules having confidence α=1 are the same as obtained with the DRSA rule induction algorithm. The induced rules are listed below:

Rule 1. *if* (f(x,q_1) ≤1) *and* (f(x,q_2) ≤1), *then* x∈ $Cl_1^{\leq}$ [α=.83]

Rule 2. *if* (f(x,q_1) ≤1) *and* (f(x,q_3) ≤1), *then* x∈ $Cl_1^{\leq}$ [α=1]

Rule 3. *if* (f(x,q_1) ≤1) *and* (f(x,q_2) ≤1) *and* (f(x,q_3) ≤2), *then* x∈ $Cl_1^{\leq}$ [α=1]

Rule 4. *if* ((f(x,q_1) ≤1), *then* x∈ $Cl_2^{\leq}$ [α=0.89]

Rule 5. *if* ((f(x,q_1) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.85]

Rule 6. *if* ((f(x,q_2) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.91]

Rule 7. *if* ((f(x,q_2) ≤1), *then* x∈ $Cl_2^{\leq}$ [α=1]

Rule 8. *if* (f(x,q_1) ≤2) *and* (f(x,q_2) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.92]

Rule 9. *if* ((f(x,q_3) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.92]

Rule 10. *if* ((f(x,q_3) ≤1), *then* x∈ $Cl_2^{\leq}$ [α=0.95]

Rule 11. *if* (f(x,q_1) ≤2) *and* (f(x,q_3) ≤1), *then* x∈ $Cl_2^{\leq}$ [α=1]

Rule 12. *if* (f(x,q_1) ≤2) *and* (f(x,q_3) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.96]

Rule 13. *if* (f(x,q_2) ≤2) *and* (f(x,q_3) ≤2), *then* x∈ $Cl_2^{\leq}$ [α=0.93]

Rule 14. *if* (f(x,q_2)≥2) *and* (f(x,q_3)≥2), *then* x∈ $Cl_2^{\geq}$ [α=0.82]

Rule 15. *if* (f(x,q_2)≥3) *and* (f(x,q_3)≥2), *then* x∈ $Cl_2^{\geq}$ [α=1]

Rule 16. *if* (f(x,q_2)≥2) *and* (f(x,q_3)≥3), *then* x∈ $Cl_2^{\geq}$ [α=0.9]

Rule 17. *if* (f(x,q_1)≥3) *and* (f(x,q_2)≥2) *and* (f(x,q_3)≥2), *then* x∈ $Cl_2^{\geq}$ [α=0.83]

Rule 18. *if* (f(x,q_2)≥3) *and* (f(x,q_3)≥3), *then* x∈ $Cl_3^{\geq}$ [α=0.8]

Rule 19. *if* (f(x,q_1)≥3) *and* (f(x,q_2)≥3) *and* (f(x,q_3)≥2), *then* x∈ $Cl_3^{\geq}$ [α=1]

Rule 20. *if* (f(x,q_1)≥2) *and* (f(x,q_2)≥3) *and* (f(x,q_3)≥3), *then* x∈ $Cl_3^{\geq}$ [α=1]

Managers of the airline company appreciated the easy verbal interpretation of the above rules. For example rule 16 says that if seat width is at least average and leg room is (at least) good, then the overall evaluation of the comfort is at least average with a confidence of 90%, independently of the space for hand luggage. This rule covers 10 examples, i.e. 20% of all questionnaires. Let us remark that such a strong pattern would not be discovered by DRSA because of one negative example (#45) being inconsistent with four other positive examples (#16,18,22,49).

Let us remark that relaxation of the confidence level has at least the following two positive consequences:

1) it enlarges lower approximations, permitting to regain many objects that were inconsistent with some marginal objects from outside of the considered union of

classes: for instance six objects, #16,18,22,32,48,49, inconsistent with object #45 entered the lower approximation $\underline{C}^{0.8}(Cl_2^{\geq})$;

2) it discovers strong rule patterns that did not appear when the dominance was strictly observed in rule induction: for instance rule 16 above.

The above aspects are very useful when dealing with large real data sets, which is usually the case of customer satisfaction analysis.

4. Conclusions

The relaxation of the dominance principle introduced in the dominance-based rough set approach results in a more flexible approach insensitive to marginal inconsistencies encountered in data sets. The variable-consistency model thus obtained maintains all basic properties of the rough sets theory, like inclusion, monotonicity with respect to supersets of criteria and with respect to the consistency level. The rough approximations resulting from this model are the basis for construction of decision rules with a required confidence. The variable-consistency model is particularly useful for analysis of large data sets where marginal inconsistencies may considerably reduce the lower approximations and prevent discovery of strong rule patterns.

Acknowledgement

The research of S. Greco and B. Matarazzo has been supported by the Italian Ministry of University and Scientific Research (MURST). R. Slowinski and J. Stefanowski wish to acknowledge financial support of the KBN research grant no. 8T11F 006 19 from the State Committee for Scientific Research.

References

1. Greco S., Matarazzo B., Slowinski R., *Rough Approximation of Preference Relation by Dominance Relations*, ICS Research Report 16/96, Warsaw University of Technology, Warsaw, 1996. Also in *European Journal of Operational Research* 117 (1999) 63-83.

2. Greco S., Matarazzo B., Slowinski R., A new rough set approach to evaluation of bankruptcy risk. In C. Zopounidis (eds.), *Operational Tools in the Management of Financial Risk*, Kluwer Academic Publishers, Dordrecht, Boston, 1998, 121-136.

3. Greco S., Matarazzo B., Slowinski R., The use of rough sets and fuzzy sets in MCDM. In T. Gal, T. Stewart and T. Hanne (eds.) *Advances in Multiple Criteria Decision Making*, chapter 14, Kluwer Academic Publishers, Boston, 1999, 14.1-14.59.

4. Greco S., Matarazzo B., Slowinski R., Stefanowski J., An algorithm for induction of decision rules consistent with dominance principle. In: *Proc. 2nd Int. Conference on Rough Sets and Current Trends in Computing*, Banff, October 16-19, 2000 (to appear).

5. Pawlak, Z., Rough sets*, International Journal of Information & Computer Sciences* 11 (1982) 341-356.

6. Pawlak, Z., *Rough Sets. Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.

7. Siskos Y., Grigoroudis E., Zopounidis C., Sauris O., Measuring customer satisfaction using a collective preference disaggregation model, *Journal of Global Optimization*, 12 (1998) 175-195.

8. Slowinski R., Stefanowski J, Greco, S., Matarazzo, B., Rough sets processing of inconsistent information. *Control and Cybernetics* 29 (2000) no.1, 379-404.

9. Slowinski R., Vanderpooten D., A generalized definition of rough approximations based on similarity. *IEEE Transactions on Data and Knowledge Engineering*, 12 (2000) no. 2, 331-336.

10. Stefanowski J., On rough set based approaches to induction of decision rules. In Polkowski L., Skowron A. (eds.) *Rough Sets in Data Mining and Knowledge Discovery*, vol. 1, Physica-Verlag, Heidelberg, 1998, 500-529.

11. Ziarko W. Variable precision rough sets model. *Journal of Computer and Systems Sciences* 46 (1993) no. 1, 39-59.

12. Ziarko W. Rough sets as a methodology for data mining. In Polkowski L., Skowron A. (eds.) *Rough Sets in Data Mining and Knowledge Discovery*, vol. 1, Physica-Verlag, Heidelberg, 1998, 554-576.