Decision information functions for inconsistent decision tables analysis

Dominik Ślęzak

Institute of Mathematics of Warsaw University
Banacha 2 02097 Warsaw Poland
Phone:+48 22 658 34 49, Fax: +48 22 658 34 48
email: slezak@alfa.mimuw.edu.pl

Abstract. We consider the notion of decision function as acting over conditional frequency distributions computed from a data table. We draw the connection between decision functions and approaches to generating uncertain decision rules for the object classification. We introduce the notion of decision implicant with respect to a decision function and show properties of such implicants for exemplar types of functions. As a conclusion, we obtain a wide class of approximate implicants providing an intuitive and flexible tool for extracting information about dependencies from data.

1 Introduction

Concerning uncertain information about a distinguished feature (decision, output) conditioned by other features (conditional attributes, inputs), refers to the task of decision rule generation. In case of consistent data tables, where such conditional information is always deterministic, decision rules can be generated from implicants developed in rough sets theory ([1]) to describe subsets of features being appropriate for classification of cases from some domain with respect to fixed decision. There, according to minimal description length principle and some simple statistical observations concerning generated decision rules, the problem of finding minimal implicant is crucial. Although proved to be NP-hard with respect to the number of conditions, this problem is possible to be solved approximately by so far developed rough set based heuristics (see [3] for references). The consistency of data with respect to a distinguished decision is quite an artificial assumption in view of practical applications. Therefore, we focus on introducing possibly wide range of tools for implicant's notion generalization for inconsistent decision tables, by handling decision functions modeling different strategies of reasoning under uncertainty. Results analogous to those for deterministic implicants enable to apply the same algorithmic heuristics as mentioned above, for searching for conditions optimal with respect to different uncertainty representations. By introducing the intuitive formula for average precision of reasoning about decision with rules generated by particular subsets of conditions, we are also able to handle implicants which determine decision approximately up to some set degree being a number between 0 and 1. Presented properties of precision measures and corresponding approximate implicants enable to search for potentially much smaller subsets of conditions *almost* preserving the initial precision level. As a conclusion, we obtain a wide class of approximate implicants providing an intuitive and flexible tool for extracting information about dependencies from data.

2 Rough set representation of information

While reasoning about a domain specified by our needs, we are usually forced to base just on the information gathered by the analysis of some sample of objects. The main paradigm of rough set theory ([1]) states that such a universe of known objects is assumed to be the only source of knowledge able to be used for classification of cases outside the sample. In applications, reasoning is usually stated as a classification problem, concerning distinguished decision attribute to predict under given conditions. By a decision table we understand a triple $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, where each attribute $a \in A \cup \{d\}$ is identified with function $a : U \to V_a$ from the universe U of objects into the set V_a of all possible values on a. While classifying new objects outside U with respect to their membership to pairwise disjoint decision classes corresponding to values $v_d \in V_d$ for distinguished decision attribute $d \notin A$, we refer to equivalence classes of indiscernibility relation, defined, for arbitrary subset $B \subseteq A$, as $IND(B) = \{(u_1, u_2) \in U \times U : Inf_B(u_1) = Inf_B(u_2)\}$. We can see that information function $Inf_B(u) = \left(a_{i_1}(u), .., a_{i_{|B|}}(u)\right)$, yields a one-to-one correspondence between equivalence classes of $IND(B)$ and elements of the set $V_B^U = \left\{w_B \in V_B : Inf_B^{-1}(w_B) \neq \emptyset\right\}$ of all vector values on B supported by objects in U. Given the notion of indiscernibility relation, the very initial model of decision table is the following.

Definition 1. Decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ is called consistent iff for each $w_A \in V_A^U$ corresponding $IND(A)$-class $Inf_A^{-1}(w_A)$ is contained in decision class corresponding to some $v_d \in V_d$, i.e. exactly one decision value is possible for each subset of objects with the same values on A, within known universe U.

In case of such consistency we classify new cases by analogy to those from the universe, i.e., given some $new \notin U$ indiscernible from $old \in U$, we predict that it would have decision value $d(old)$, since this is the only supported choice. If for such $old \in U$ there is $Inf_A(old) = w_A$ and $d(old) = v_d$, then we classify $new \notin U$ basing on decision rule $A = w_A \Rightarrow d = v_d$, stating that if any object is equal to w_A on A, then it is going to have decision value v_d. We can also consider boolean implication $A \Rightarrow d$ which holds iff each $w_A \in V_A$ implies some $v_d \in V_d$ in the above sense. Thus, decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ is consistent iff boolean implication $A \Rightarrow d$ is satisfied, what means that decision d can be completely determined by the conditions. The fact that the whole A determines d implies the question whether all conditions are necessary to keep this property.

Definition 2. Given consistent decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, subset $B \subseteq A$ is called an implicant iff it satisfies boolean implication $B \Rightarrow d$ understood as above.

Actually, it seems to be natural in practical applications to ask for possibly minimal implicants, what would make the process of classification by analogy as efficient as possible. Such understanding of efficiency, finding additional support in minimal description length principle and some simple statistical observations, can be interpreted in terms of the following notion.

Definition 3. We say that decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ is applicable to an object $new \notin U$ with respect to subset $B \subseteq A$ iff new fits some indiscernibility class of $IND(B)$ i.e. value vector taken by new on B belongs to V_B^U.

In view of the classification task we would like to maximize applicability to new examples, keeping in mind that it is somehow opposite to the precision of classification. Indeed, for any $w_A \in V_A$ we have implication

$$w_A \in V_A^U \Longrightarrow w_A^{\downarrow B} \in V_B^U \tag{1}$$

(where $w_A^{\downarrow B}$ denotes the projection of w_A onto the value vector over attributes from B) and thus we can claim that applicability to new objects is potentially more probable with respect to smaller subsets of conditions. On the other hand, however, we cannot consider too small attribute sets because they may not preserve precision of classification. From this point of view, implicants with possibly small number of attributes seem to be optimal with respect to the balance between applicability and precision.

3 Reasoning in inconsistent decision tables

In a consistent decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, where each class of $IND(A)$ is contained in one of the decision classes, preserving an information about decision is equivalent to its determination. There, decision rules generated by implicant $B \subseteq A$ form the bunch of the form $B = w_B \Rightarrow d = v_d$, $w_B \in V_B^U$, $Inf_B^{-1}(w_B) \subseteq d^{-1}(v_d)$. Since in inconsistent decision tables we cannot expect all decision rules to be deterministic, we must settle a way of dealing with possibly indeterministic knowledge about decision, conditioned by information about other attribute values similarly as above. Let us focus on conditional representation based on rough membership functions (compare with [2]) $\mu_{d/B} : V_d \times V_B^U \to [0,1]$, defined, for any fixed $B \subseteq A$, by formula

$$\mu_{d/B}(v_d/w_B) = \frac{\left|Inf_B^{-1}(w_B) \cap Inf_d^{-1}(v_d)\right|}{\left|Inf_B^{-1}(w_B)\right|} \tag{2}$$

where the denominator equals to the number of objects with vector value w_B on B, and the nominator equals to the number of objects which additionally have the decision value equal to v_d. Under fixed linear ordering $V_d = \langle v_{d,1}, ..., v_{d,|V_d|} \rangle$ (where $|V_d|$ denotes the cardinality of d's domain), one can consider rough membership distributions $\mu_{d/B} : V_B^U \to \triangle_{|V_d|-1}$ defined by

$$\mu_{d/B}(w_B) = \left(\mu_{d/B}(v_{d,1}/w_B), ..., \mu_{d/B}\left(v_{d,|V_d|}/w_B\right)\right) \tag{3}$$

where $\triangle_{|V_d|-1}$ denotes ($|V_d|$-1)-dimensional simplex. Such conditional frequency distributions, widely studied e.g. in statistics as estimators of conditional probability distributions (see e.g. [5] or [6] for further references), may be regarded as decision functions which attach to each vector value $w_B \in V_B^U$ over considered conditions $B \subseteq A$

an information concerning degrees of hitting of $Inf_B^{-1}(w_B)$ into particular decision classes $Inf_d^{-1}(v_{d,i})$, $i = 1,..,|V_d|$. Distributions $\mu_{d/B}$ are just one of the examples of decision functions which, for arbitrary $B \subseteq A$, specify conditional information about decision attribute. This is because one may not need so complete information, corresponding for a given $B \subseteq A$ to the whole bunch of decision rules of the form $B = w_B \Rightarrow_{\mu_{d/B}(v_d/w_B)} d = v_d$, $w_B \in V_B^U$, $v_d \in V_d$, where $\mu_{d/B}(v_d/w_B)$ is interpreted as a rule's precision. For example, one may reason about a new case with values w_B on B basing on generalized decision function $\partial_{d/B} : V_B^U \rightarrow 2^{V_d}$ defined by

$$\partial_{d/B}(w_B) = \left\{v_d \in V_d : Inf_B^{-1}(w_B) \cap Inf_d^{-1}(v_d) \neq \emptyset\right\} \tag{4}$$

which corresponds to the reasoning with the set of possible decision values $\partial_{d/B}(w_B)$, without considering accurate values of $\mu_{d/B}$. Yet another example can be, e.g., connected with so called majority strategy, where we concern just the most probable decision values, i.e. such $v_{\max} \in V_d$ that the quantity of $\mu_{d/B}(v_{\max}/w_B)$ is the highest possible over V_d, not needing information about other decision values. Nevertheless, we can say that rough membership distributions express the most accurate knowledge about dependencies of the decision on conditions, unless some additional information outside **A** is provided. So, it should be possible to model particular reasoning strategies as functions acting over $\mu_{d/B}$, by "forgetting" some part of frequency based information which is not necessary for a given approach.

Definition 4. Given a simplex $\triangle_{n-1}$ consisting of elements $s = \langle s_1, ..., s_n \rangle$, $s_i \geq 0, i = 1,..,n$, $\sum_{i=1}^n s_i = 1$, we call as a decision function any $\phi : \triangle_{n-1} \rightarrow \triangle_{n-1}$, represented as $\phi(s) = \langle \phi(s_1), ..., \phi(s_n) \rangle$, satisfying assumptions:
(i) $\forall_{i=1,..,n} [(s_i = 0) \Rightarrow (\phi(s_i) = 0)]$
(ii) $\forall_{i,j=1,..,n} [(s_i \leq s_j) \Rightarrow (\phi(s_i) \leq \phi(s_j))]$
Given a decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, $B \subseteq A$ and ϕ as above, function $\phi_{d/B} : V_B^U \rightarrow \triangle_{|V_d|-1}$ defined by $\phi_{d/B}(w_B) = \phi\left(\mu_{d/B}(w_B)\right)$ is called a $\phi_{d/B}$-decision function.

The motivation for such definition is that by considering decision functions as irrelevant from decision table structure, unless combined with rough membership distributions $\mu_{d/B}$, we focus just on the strategy of reasoning with conditional frequency information provided. Here, assumptions Def. 4(i) and Def. 4(ii) mean, respectively, that we cannot attach a positive chance to not supported events and that the chances provided by reasoning strategy cannot contradict to frequencies derived from the information source. Thus, we can say that both above assumptions express the idea that we are not allowed to base on any additional knowledge besides that derived from data. Obviously, Definition 4 cannot refer to all methods of adjusting numerical values to conditioned chances of particular events (compare with [5]). Still, the expressive power of the above characteristics is enough to describe a reasonably large group of strategies. For illustration, let us go back to the above mentioned examplar techniques based, respectively, on majority and generalized approaches.

Example 1. Let us consider function $\partial : \triangle_{n-1} \rightarrow \triangle_{n-1}$, defined, for $i = 1,..,n$, by $\partial(s_i) = |\partial(s)|^{-1}$ for $s_i \in \partial(s)$ and $\partial(s_i) = 0$ otherwise, where $\partial(s) = \{s_i : s_i > 0, i = 1,..,n\}$. For a given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, for any $B \subseteq A$, we obtain that function $\partial_{d/B} : V_B^U \rightarrow \triangle_{|V_d|-1}$, for $i = 1,..,|V_d|$, is equal to $\left|\partial_{d/B}(w_B)\right|^{-1}$ for $v_{d,i} \in \partial_{d/B}(w_B)$ and 0 otherwise, where $\partial_{d/B}(w_B)$ is understood in terms of (4). Thus, we obtain a new interpretation of $\partial_{d/B}$ as the function labeling each $w_B \in V_B^U$ with the uniform distribution over the subset of decision values with positive frequencies conditioned by w_B. Depending on the context, values of $\partial_{d/B}$ can be interpreted as the subsets of V_d or as simplex elements similar to those of the form (3).

Example 2. Let us consider function $m : \triangle_{n-1} \rightarrow \triangle_{n-1}$, defined for $i = 1,..,n$, by $m(s_i) = |m(s)|^{-1}$ and $m(s_i) = 0$ otherwise, where $m(s) = \{s_i : s_i = \max_{j=1,..,n} s_j, i = 1,..,n\}$. Combining it with $\mu_{d/B}$ just like above, for $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, $B \subseteq A$, we obtain function $m_{d/B} : V_B^U \rightarrow \triangle_{|V_d|-1}$, with marginal values denoted by $m_{d/B}(v_{d,i}/w_B)$, $i = 1,..,|V_d|$, which labels each $w_B \in V_B^U$ with the uniform distribution over the subset of decision values with maximal possible frequencies conditioned by w_B.

The significance of the above examples turns out to be connected with the notion of precision for so called ϕ-based boolean implications of the form $B \Rightarrow_{\phi_{d/B}} d$, where to each decision rule with prefix "$B = w_B$" and suffix "$d = v_d$" we attach precision $\phi_{d/B}(v_d/w_B)$. Since it is quite uncomfortable to handle the whole bunches of such local precision values while the global analysis, a solution would be to label subsets $B \subseteq A$ with numerical coefficients reflecting the average precision of corresponding rules. We propose to set such coefficients as the expected measures of local precision values computed over the *a priori* frequency distribution derived from a given decision table.

Definition 5. Given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ and decision function ϕ, we define the ϕ-precision measure $E_\phi(d/\cdot) : 2^A \to [0,1]$ as

$$E_\phi(d/B) = \sum_{w_B \in V_B^U} \sum_{v_d \in V_d} \mu_{B,d}(w_B, v_d)\, \phi_{d/B}(v_d/w_B) \tag{5}$$

where for each $w_B \in V_B^U$, $v_d \in V_d$ we put

$$\mu_{B,d}(w_B, v_d) = \frac{\left|Inf_B^{-1}(w_B) \cap Inf_d^{-1}(v_d)\right|}{|U|} \tag{6}$$

Taking the average local precision with respect to frequencies of the form (6) corresponds to the idea of defining the support of the uncertain decision rule with prefix "$B = w_B$" and suffix "$d = v_d$" as the number of objects for which this rule is both applicable and valid. Thus, we can say that quantity $E_\phi(d/B)$ expresses the expected degree of precision of a decision rule generated by $B \subseteq A$, applicable and well classifying randomly chosen object from the universe. It turns out that by using ϕ-precision measures one can obtain intuitive properties of average precision connected with ϕ-based boolean implications. The following result points at possibility of comparison of different ϕ-related reasoning strategies in terms of precision factors defined by (5).

Proposition 6. *Given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ and fixed $B \subseteq A$, for each decision function ϕ satisfying Def. 4(i) and Def. 4(ii) we have inequalities*

$$E_\partial(d/B) \le E_\phi(d/B) \le E_m(d/B) \tag{7}$$

where quantities $E_\partial(d/B)$ and $E_m(d/B)$ correspond to generalized and majority decision functions, considered in Examples 1 and 2, respectively.

According to the above, we can say that reasoning with generalized decision based strategy takes the minimal and reasoning with majority strategy - the maximal amount of information provided by conditional frequencies. Indeed, it meets with the intuition that by reasoning with the set of possible decision values we obtain the most vague, but the safest, answer. On the other hand, reasoning with the most frequently occurring decision values if the most risky with respect to uncertainty concerning the source of data itself, but giving the sharpest possible output.

4 Decision function based implicants

Given an interpretation of both the local and global indeterministic implications' precision, let us go back to the notion of implicant understood as a subset determining decision in a similar way to that corresponding to the whole set of conditions. In case of consistent decision tables we stated implicants as subsets which enabled to keep precision of reasoning at the initial, deterministic level. In case of inconsistencies, we would like to generalize the notion of implicant as preserving precision degrees corresponding to particular strategies of reasoning.

Definition 7. Given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ and decision function ϕ, subset $B \subseteq A$ is called a ϕ-implicant iff ϕ-based boolean implication $B \Rightarrow_{\phi_{d/B}} d$ is equivalent to $A \Rightarrow_{\phi_{d/A}} d$, i.e. iff for each $w_A \in V_A^U$ there is equality

$$\phi_{d/B}\left(w_A^{\downarrow B}\right) = \phi_{d/A}(w_A) \tag{8}$$

which assures that $w_A^{\downarrow B}$ generates the same ϕ-based decision rules with prefix "$B = w_A^{\downarrow B}$" as those with prefix "$A = w_A$". In case of identity function $\phi = id$, we call corresponding subsets of conditions μ-implicants.

For consistent decision tables, the above definition is equivalent to Definition 2. Indeed, in such a case, distribution $\mu_{d/A}(w_A)$, for each particular $w_A \in V_A^U$, corresponds to a unique vertex of $\triangle_{|V_d|-1}$ and thus it must be also the case for $\phi_{d/A}(w_A)$, what can be easily derived from assumptions Def. 4(i) and Def. 4(ii). The same must also be true for $\phi_{d/B}\left(w_A^{\downarrow B}\right)$, because of equality (8), what yields that $\mu_{d/B}\left(w_A^{\downarrow B}\right)$ must be the same vertex of $\triangle_{|V_d|-1}$ than $\mu_{d/A}(w_A)$. In particular, getting back to the objective concerning the search of minimal implicants mentioned in Section 2, the above equivalence implies the following corollary, well known for consistent case (see [3] for references).

Proposition 8. *Given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ and decision function ϕ, the problem of finding ϕ-implicant minimal in sense of cardinality is NP-hard with respect to the number of attributes in A.*

Fortunately, on the other hand, it turns out that at least for some decision functions ϕ corresponding ϕ-implicants have the same discernibility characteristics as implicants in the consistent case (compare with [4]). It enables to apply the same algorithmic methods basing on discernibility as already developed (see [3] for references) to searching for minimal ϕ-implicants.

Proposition 9. *Given decision table* $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$*, subset* $B \subseteq A$ *is a* ϕ*-implicant for decision functions* $\phi = \partial, id, m$*, iff it is a decision implicant for consistent decision table* $\mathbf{A}_\phi = (\mathbf{U}, \mathbf{A}, \phi_{\mathbf{d}/\mathbf{A}})$*, where decision* d *is replaced by multi-dimensional decision attribute* $\phi_{d/A}$*. In other words, subset* $B \subseteq A$ *is a* ∂*,* μ *or* m*-implicant iff for each pair* $w_{A,1}, w_{A,2} \in V_A^U$ *there is implication*

$$\phi_{d/A}(w_{A,1}) \neq \phi_{d/A}(w_{A,2}) \Rightarrow w_{A,1}^{\downarrow B} \neq w_{A,2}^{\downarrow B} \tag{9}$$

In fact, there is a lot of rough set applications concerning the minimal generalized decision implicants search, where the above characteristics, for $\phi = \partial$, is used. Moreover, such a discernibility representation enables to develop some tools for considering approximate ϕ-implicants, according to searching for subsets discerning *almost* all pairs of vector values $w_{A,1}, w_{A,2} \in V_A^U$ such that $\phi_{d/A}(w_{A,1}) \neq \phi_{d/A}(w_{A,2})$ or discerning only these pairs of vector values from V_A^U whose $\phi_{d/A}$-values remain *far enough* to each other in terms of some distance measure ρ (compare with [6]). Yet another possibility appears as the result of the following reformulation of the implicant's notion.

Definition 10. Given decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$ and decision function ϕ, subset $B \subseteq A$ is called an average ϕ-implicant iff the average precision of ϕ-based boolean implication $B \Rightarrow_{\phi_{d/B}} d$ is not less than the average precision of $A \Rightarrow_{\phi_{d/A}} d$, i.e. iff there is inequality

$$E_\phi(d/B) \geq E_\phi(d/A) \tag{10}$$

The above definition has an important advantage with respect to Definition 7. Namely, so far we were dealing with precision degrees of particular decision rules, without possibility of talking about the global increase or decrease of precision for different subsets of conditions. Thus, ϕ-implicants could be considered only in terms of the preservence of local information. Now the situation changes, since we begin to handle the concrete numerical formula referring to global precision of ϕ-based boolean implications. Thus, inequality (10) does make sense, although one might require that such global degrees of precision should satisfy some additional properties.

Definition 11. Decision function ϕ satisfying conditions of Definition 4 is called monotonic iff for each decision table $\mathbf{A} = (\mathbf{U}, \mathbf{A}, \mathbf{d})$, for each subset $B \subseteq A$ there is inequality $E_\phi(d/B) \leq E_\phi(d/A)$, where equality holds iff for each $w_A \in V_A^U$ equality (8) holds.

In terms of the above definition, the notions of ϕ-implicant and average ϕ-implicant are equivalent for any monotonic function ϕ. This is very important in view of possibility of encoding the whole information concerning the preservence of the decision rules behavior in one numerical quantity. From this point of view, the following fact is very helpful.

Proposition 12. *Decision functions* $\phi = \partial, id, m$ *are monotonic in sense of Definition 11.*

In particular, the above result enables to search for minimal approximate implicants which determine decision in *almost* same degree as the whole set of conditions in a reasonable time. By "*almost*" we understand the replacement of inequality (10) by $E_\phi(d/B) + \varepsilon \geq E_\phi(d/A)$ for a given approximation degree $\varepsilon \in [0, E_\phi(d/A))$. Since the problem of finding minimal average ϕ-implicant is NP-hard (for the same reason as in case of ϕ-implicants before) and we are not likely to expect better situation for its approximations, it is crucial to learn how to search for minimal ε-approximate average ϕ-implicants indeed. Here, for decision functions $\phi = \partial, id, m$, Proposition 12 enables to adapt another group of already developed algorithmic tools for searching for minimal deterministic implicants. These tools do not refer to discernibility characteristics straightforwardly, taking the advantage of the fact that for any implicant we can check in a linear time with respect to its cardinality, whether it is minimal in sense of inclusion. Indeed, monotonicity of $\phi = \partial, id, m$ assures that it is enough to know for a given ε-approximate average ϕ-implicant $B \subseteq A$ that for each $a \in B$ subset $B \setminus \{a\}$ does not satisfy inequality $E_\phi(d/B \setminus \{a\}) + \varepsilon \geq E_\phi(d/A)$, to claim the same about all subsets of B. This property enables to use techniques developed e.g. in [7] to find in a real time minimal ε-approximate average ϕ-implicants for monotonic functions $\phi = \partial, id, m$, under any fixed approximation degree.

5 Conclusions

We considered the class of decision functions as modeling different strategies of uncertain reasoning with data. As a conclusion, we obtained a wide class of approximate decision function based implicants providing intuitive and flexible tools for extracting information about dependencies from data.

Acknowledgements This work was supported by KBN Research Grant 8T11C01011 and ESPRIT project 20288 CRIT-2. The author does acknowledge the help of Polish-Japanese Institute of Information Technology.

References

1. Pawlak, Z.: Rough sets - Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Dordrecht (1991)
2. Pawlak, Z., Skowron, A.: Rough membership functions. In: R.R. Yaeger, M. Fedrizzi, and J. Kacprzyk (eds.), Advances in the Dempster Shafer Theory of Evidence, John Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore (1994) 251–271
3. Polkowski, L., Skowron, A. (eds.): Rough Sets in Knowledge Discovery. Physica-Verlag, Heidelberg (1998) 1–576 (Part 1), 1-601 (part 2)
4. Skowron, A., Rauszer, C.: The discernibility matrices and functions in information systems. In: R. Słowiński (ed.), Intelligent Decision Support. Handbook of Applications and Advances of the Rough Set Theory, Kluwer Academic Publishers, Dordrecht (1992) 311–362
5. Ślęzak, D.: Approximate reducts in decision tables. In: Proceedings of the Sixth International Conference, Information Procesing and Management of Uncertainty in Knowledge-Based Systems (IPMU'96), July 1–5, Granada, Spain (1996) 3 1159–1164
6. Ślęzak, D.: Searching for frequential reducts in decision tables with uncertain objects. In: L. Polkowski, A. Skowron (eds.), Proceedings of the First International Conference on Rough Sets and Current Trends in Computing (RSCTC'98), June 22–26, Warsaw, Poland, Springer-Verlag, Berlin Heidelberg (1998) 52–59
7. Wróblewski, J.: Genetic algorithms in decomposition and classification problem. In: L. Polkowski and A. Skowron (eds.), Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems, Physica-Verlag, Heidelberg (1998) 471–487