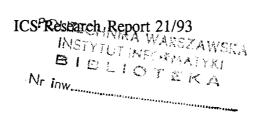
Decision Analysis Using Rough Sets

by

Zdzisław Pawlak¹, Roman Słowiński²



¹ Institute of Computer Science Warsaw University of Technology Warsaw, Poland

² Institute of Computing Science Technical University of Poznań 60-965 Poznań, Poland

and Applied Computer Science Polish Academy of Sciences 44-100 Gliwice, Poland Abstract: We show that the rough set theory is a useful tool for analysis of decision situations, in particular multi-criteria sorting problems. It deals with vagueness in the representation of a decision situation, caused by granularity of the representation. The rough set approach produces a set of decision rules involving a minimum number of most important criteria. It does not correct vagueness manifested in the representation; instead, produced rules are categorized into deterministic and non-deterministic. The set of decision rules explains a decision policy and may be used for decision support. An example illustrates the rough set analysis.

Key words: Decision, Vagueness, Rough Set Theory, Multiple Criteria, Sorting

Invited paper for XIII World Conference on operation Research IFORS 93 Lisbon, Portugal July 12-16, 1993.

Decision analysis is one of the most natural acts of human beings. It has attracted scientists for a long time who offered various mathematical tools to deal with. Mathematical decision analysis intends to bring into light those elements of a decision situation which are not evident for the actors and may influence their attitude towards the situation. More precisely, the elements revealed by the mathematical decision analysis either explain the situation or prescribe, or simply privilege, some behavior in order to increase the coherence between evolution of the decision process on the one hand and goals and value systems of the actors, on the other hand (cf. Roy 1992).

One of factors hindering revelation of the above mentioned elements is vagueness inherent to representation of a decision situation. Vagueness may be caused by granularity of the representation. Due to the granularity, the facts describing a situation are deterministic or non-deterministic. Facts are deterministic if they can be described univocally by means of "granules" of the representation, and they are non-deterministic, otherwise.

A formal framework for discovering facts from representation of a decision situation has been given by Pawlak (1982) and called a rough set theory. The rough set theory assumes the representation in a decision table form which is a special case of an information system. Rows of this table correspond to objects (actions, alternatives, candidates, patients, etc.) and columns correspond to attributes. For each pair (object, attribute) there

is known a value called descriptor. Each row of the table contains descriptors representing information about corresponding object of a given decision situation. In general, the set of attributes is partitioned into two subsets: condition attributes (criteria, tests, symptoms, etc.) and decision attributes (decisions, classifications, taxonomies, etc.).

As in decision problems the concept of criterion is often used instead of condition attribute, it should be noticed that the latter is more general than the former because the domain (scale) of a criterion has to be ordered according to decreasing or increasing preference while the domain of a condition attribute has not to be ordered. Similarly, the domain of a decision attribute may be ordered or not. The ordering property has to be in considerations concerning a special taken into account dependency among attributes - dependency in the sense concordance between criteria or experts (cf. Boryczka 1989). Apart from this specific case, the ordering property has no influence on the rough set analysis, so there will be no distinction between criteria and condition attributes unless specified.

Let us suppose a decision situation represented by a decision table where a finite set of objects is described by several condition attributes and single decision attribute. a The observation that objects may be indiscernible in descriptors is a starting point for the analysis. Indiscernibility of objects by means of condition attributes prevents generally their precise assignment to sets following from generated by the decision attribute. Given an equivalence relation

viewed as an indiscernibility relation which thus induces an approximation space made of equivalence classes, a rough set is a pair of a lower and an upper approximation of a set in terms of the classes of indiscernible objects. In other words, a rough set is a collection of objects which, in general, cannot be precisely characterized in terms of the values of the set of condition attributes, while a lower and an upper approximation of the collection can be. Using a lower and an upper approximation of a set (or family of sets - partition) one can define an accuracy and a quality of approximation. These are numbers from interval [0,1] which define how exactly one can describe the examined set of objects using available information.

Depending on the decision situation, the rough set approach brings into light different elements of this situation.

A. In the case of a multi-criteria sorting problem, which consists in assignment of each object to an appropriate pre-defined category (for instance: acceptance, rejection or request for an additional information), the rough set analysis leads to:

- · evaluation of importance of particular criteria,
- construction of minimal subsets of independent criteria having the same discernment ability as the whole set,
- non-empty intersection of those minimal subsets gives a core of criteria which cannot be eliminated without disturbing the ability of approximating the decision,
- · elimination of redundant criteria from the decision table,
- generation of the sorting rules (deterministic or not)
 from the reduced decision table; they explain a decision
 policy and may be used for sorting new coming objects.

- B. If in the set of attributes there is no decision attribute, the decision situation corresponds to a multi-criteria description of a set of objects. The information system is equivalent in this case to the performance matrix. Rough set analysis of this matrix gives the following results:
 - detection of dependencies among criteria in the sense of their concordance.
 - · elimination of one criterion from each concordant pair of criteria, which does not affect the Pareto set of objects,
- evaluation of a grade of conflict among remaining criteria.

 C. If in the set of attributes there is no condition attribute, the decision situation may correspond to a multi-expert evaluation of a set of objects; the evaluation given by one expert corresponds to one decision attribute with an ordered domain. Similarly to situation B, the results of the rough set analysis
 - detection of dependencies among experts in the sense of their concordance,

are the following:

- elimination of one expert from each concordant pair of experts,
- evaluation of a grade of conflict among opinions of remaining experts.
- D. If in the set of attributes there are both multiple condition attributes and multiple decision attributes, the decision situation may correspond to a multi-criteria, multi-expert sorting problem. Apart from advantages listed for A, in this case, the rough set analysis enables:

 measuring of the degree of consistency of the experts with the characterization of the objects by the set of criteria.

In the next section, we recall basic concepts of the rough set theory. In the next section, we apply the rough set methodology to an exemplary sorting problem. The final section groups conclusions.

BASIC CONCEPTS OF THE ROUGH SET THEORY

Information system

By an information system we understand the 4-tuple (cf. Pawlak 1991) S=<U,Q,V, ρ >, where U is a finite set of objects, Q is a finite set of attributes, V = $\bigwedge^N V_q$ and V_q is a domain of the attribute q, and $\rho: U \times Q \longrightarrow V$ is a total function such that $\rho(x,q) \in V_q$ for every $q \in Q$, $x \in U$, called an information function.

Let $S=\langle U,Q,V,\rho \rangle$ be an information system and let $P\subseteq Q$ and $x,y\in U$. We say that x and y are indiscernible by the set of attributes P in S iff $\rho(x,q)=\rho(y,q)$ for every $q\in P$. Thus every $P\subseteq Q$ generates a binary relation on U which will be called an indiscernibility relation, denoted by IND(P). Obviously, IND(P) is an equivalence relation for any P. Equivalence classes of IND(P) are called P-elementary sets in S. The family of all equivalence classes of relation IND(P) on U is denoted by U|IND(P) or, in short, U|P.

 $Des_{\mathbf{p}}(X)$ demostes a description of P-elementary set $X \in U \mid P$ in terms of values of attributes from P, i.e.

$$Des_{p}(X) = \{(q, v): f(x,q)=v, \forall x \in X, \forall q \in P\}$$

Approximationrof sets

Let $P\subseteq Q$ and $Y\subseteq U$. The P-lower approximation of Y, denoted by $\overline{P}Y$, and the P-upper approximation of Y, denoted by $\overline{P}Y$, are defined as:

 $PY = \bigcup \{X \in U \mid P : X \subseteq Y\}$

 $\overline{P}Y = \bigcup \{X \in U \mid P : X \cap Y \neq \emptyset\}$

The P-boundary (doubtful region) of set Y is defined as

$$Bn_{\mathbf{p}}(Y) = \overline{P}Y - \underline{P}Y$$
.

Set $\underline{P}Y$ is the set of all elements of U which can be certainly classified as selements of Y, employing the set of attributes P. Set $\underline{P}Y$ is the set of elements of U which can be possibly classified as elements of Y, using the set of attributes P. The set $Bn_p(Y)$ tis the set of elements which cannot be certainly classified to Y using the set of attributes P.

With every set YSU, we can associate an accuracy of approximations of set Y by P in S, or in short, accuracy of Y, defined as: fixed

$$\alpha_{p}(Y) = \frac{\operatorname{card}(\underline{P}Y)}{\operatorname{card}(\overline{P}Y)}$$

Approximation of a partition of U

Let S be an information system, P=Q, and let $y=\{Y_1,Y_2,\ldots,Y_n\}$ be a partition of U. The origin of this partition is independent

on attributes from P; it can follow from solving a sorting problem by an expert. Subsets Y_i , $i=1,\ldots,n$, are categories of partition Y. By P-lower (P-upper) approximation of Y in S we mean sets $PY = \{PY_1, PY_2, \ldots, PY_n\}$ and $PY = \{PY_1, PY_2, \ldots, PY_n\}$, respectively. The coefficient

$$\gamma_{P}(y) = \frac{\sum_{i=1}^{n} card (PY_{i})}{card(U)}$$

is called the quality of approximation of partition Y by set of attributes P, or in short, quality of sorting. It expresses the ratio of all P-correctly sorted objects to all objects in the system.

Reduction of attributes

We say that the set of attributes $R \subseteq Q$ depends on the set of attributes $P \subseteq Q$ in S (denotation $P \longrightarrow R$) iff $IND(P) \subseteq IND(R)$. Discovering dependencies between attributes is of primary importance in the rough set approach to knowledge analysis.

Another important issue is that of attribute reduction, in such a way that the reduced set of attributes provides the same quality of sorting as the original set of attributes. The minimal subset $R\subseteq P\subseteq Q$ such that $\gamma_P(y)=\gamma_R(y)$ is called y-reduct of P (or, simply, reduct if there is no ambiguity in the understanding of y) and denoted by $RED_y(P)$. Let us notice that an information system may have more than one y-reduct. Intersection of all y-reducts is called the y-core of P, i.e. $CORE_y(P)=\bigcap RED_y(P)$. The core is a collection of the most significant attributes in the system.

Decision tables

An information system can be seen as decision table assuming that $Q=C\cup D$ and $C\cap D=\emptyset$, where C are called condition attributes, and D, decision attributes. Decision table $S=<U,C\cup D,V,\rho>$ is deterministic iff $C\to D$; otherwise it is non-deterministic. The deterministic decision table uniquely describes the decisions to be made when some conditions are satisfied. In the case of a non-deterministic table, decisions are not uniquely determined by the conditions. Instead, a subset of decisions is defined which could be taken under circumstances determined by conditions.

From the decision table a set of decision rules can be derived. Let U|IND(C) be a family of all C-elementary sets called condition classes, denoted by X_i (i=1,...,k). Let, moreover, U|IND(D) be the family of all D-elementary sets called decision classes, denoted by Y_i (j=1,...,n).

 $Des_{\mathbb{C}}(X_i) \Rightarrow Des_{\mathbb{D}}(Y_j)$ is called (C,D)-decision rule. The rules are logical statements "if ... then ..." relating descriptions of condition and decision classes. The set of decision rules for each decision class Y_j (j=1,...,n) is denoted by $\{r_{ij}\}$. Precisely,

$$\{r_{ij}\} = \{Des_{C}(X_{i}) \Rightarrow Des_{D}(Y_{j}): X_{i} \cap Y_{j} \neq \emptyset, i=1,...,k\}$$

Rule r_{ij} is deterministic iff $X_i \subseteq Y_j$, and r_{ij} is non-deterministic, otherwise.

Procedures for derivation of decision rules from decision tables were presented by Boryczka and Slowinski (1988), Slowinski and Stefanowski (1992), and by Grzymala-Busse (1992).

Types of decision situations

Using the definition of a decision table, we may characterize decision situations introduced in the first section.

- A. The multi-criteria sorting problem: card(C)>1, card(D)=1.
- B. The multi-criteria description problem: card(C)>1, card(D)=0.
- C. The multi-expert evaluation problem: card(C)=0, card(D)>1.
- D. The multi-criteria, multi-expert sorting problem: card(C)>1, card(D)>1.

MULTI-CRITERIA SORTING PROBLEM

The multi-criteria sorting problem consists in discovering decision rules, taking into account expert's (decision maker's) preferences. It is often the case that the preferences are expressed implicitly through examples of sorting decisions. A set of examples constitutes a decision table. In inductive learning, such a set is called a training sample. Decision rules are derived from the examples and then applied to new coming objects. Rough sets analysis has been applied with success to sorting problems from medicine (Fibak et al. 1986, Slowinski et al. 1988), pharmacy (Krysinski 1990), technical diagnostic (Nowicki et al. 1992) and others (the most complete reference can be found in Slowinski 1992).

To illustrate the rough set analysis, let us consider a simple case of selection of candidates to a school (cf. Moscarola 1978).

The candidates to the school have submitted their application packages with secondary school certificate, curriculum vitae and opinion from previous school, for consideration by an admission committee. Basing on these documents, the candidates were described using seven criteria (condition attributes). The list of these criteria together with corresponding scales, ordered from the best to the worst value, is given below:

 c_1 - score in mathematics, $\{5,4,3\}$

 c_2 - score in physics, $\{5,4,3\}$

 c_3 - score in English, $\{5,4,3\}$

 c_4 - mean score in other subjects, $\{5,4,3\}$

 c_5 - type of secondary school, $\{1,2,3\}$

 c_6 - motivation, $\{1,2,3\}$

 c_7 - opinion from previous school, $\{1,2,3\}$

Fifteen candidates having rather different application packages have been sorted by the committee after due consideration. They create the set of examples.

The decision attribute d makes a dichotomic partition y of the candidates: d=A means admission and d=R means rejection. The decision table with fifteen candidates is shown in Table 1. It is clear that $C=\{c_1,c_2,c_3,c_4,c_5,c_6,c_7\}$ and $D=\{d\}$.

Table 1 insert here

Let Y_A be the set of candidates admitted and Y_R the set of candidates rejected by the committee, $Y_A = \{x_1, x_4, x_5, x_7, x_8, x_{10}, x_{11}, x_{12}, x_{15}\}$, $Y_R = \{x_2, x_3, x_6, x_9, x_{13}, x_{14}\}$, $Y = \{Y_A, Y_R\}$. Sets Y_A and

 Y_R are D-definable sets in the decision table. There are 13 C-elementary sets: couples of indiscernible candidates $\{x_4,x_{10}\}$, $\{x_8,x_9\}$ and 11 discernible candidates. The C-lower and the C-upper approximations of sets Y_A and Y_R are equal, respectively, to:

$$\underline{CY}_{A} = \{x_{1}, x_{4}, x_{5}, x_{7}, x_{10}, x_{11}, x_{12}, x_{15}\}$$

$$\overline{CY}_{A} = \{x_{1}, x_{4}, x_{5}, x_{7}, x_{8}, x_{9}, x_{10}, x_{11}, x_{12}, x_{15}\}$$

$$\underline{Bn}_{C}(Y_{A}) = \{x_{8}, x_{9}\}$$

$$\underline{CY}_{R} = \{x_{2}, x_{3}, x_{6}, x_{13}, x_{14}\}$$

$$\overline{CY}_{R} = \{x_{2}, x_{3}, x_{6}, x_{8}, x_{9}, x_{13}, x_{14}\}$$

$$\underline{Bn}_{C}(Y_{R}) = \{x_{8}, x_{9}\}$$

The accuracy of approximation of sets Y_A and Y_R by C is equal to 0.8 and 0.71, respectively, and the quality of approximation of the decision by C is equal to 0.87.

Let us observe that the C-doubtful region of the decision is composed of two candidates: x_8 and x_9 . Indeed, they have the same value according to criteria from C but the committee has admitted x_8 and rejected x_9 . It means that the decision is inconsistent with evaluation of the candidates by criteria from C. apparently, the committee took into account an additional information from the application packages of the candidates or from an interview with them. This conclusion suggests to the committee, either adoption of an additional discriminatory criterion or, if its explicit definition would be too difficult, creation of a third category of candidates : those who should be invited to an interview.

The next step of the rough set analysis of the decision table is construction of minimal subsets of independent criteria

ensuring the same quality of sorting as the whole set C, i.e. the reducts of C. In our case, there are three such reducts:

$$RED_{y}^{1}(C) = \{c_{2}, c_{3}, c_{6}, c_{7}\}$$

$$RED_{y}^{2}(C) = \{c_{1}, c_{3}, c_{7}\}$$

$$RED_{y}^{3}(C) = \{c_{2}, c_{3}, c_{5}, c_{7}\}$$

It can be said that the committee took the fifteen sorting decisions taking into account the criteria from one of the reducts and discarded all the remaining criteria. Let us notice that criterion $\mathbf{c_4}$ has no influence at all on the decision because it is not represented in any reduct.

It is interesting to see the intersection of all reducts, i.e. the core of criteria:

$$CORE_{y}(C) = RED_{y}^{1}(C) \cap RED_{y}^{2}(C) \cap RED_{y}^{3}(C) = \{c_{3}, c_{7}\}$$

The core is the most essential part of set C, i.e. it cannot be eliminated without disturbing the ability of approximating the decision.

In a real case, all the reducts and the core should be submitted for consideration by the committee in view of getting its opinion about what reduct should be used to generate decision rules from the reduced decision table.

Let us suppose that the committee has chosen reduct $RED_y^2(C)$ composed of c_1, c_3, c_7 , i.e. scores in mathematics and English, and opinion from previous school. This choice could be explained in such a way that the score in mathematics (c_1) seems to the committee more important than the score in physics (c_2) plus type of secondary school (c_5) or motivation (c_6) .

Now, the decision table can be reduced to criteria represented in $RED_y^2(C)$. The decision rules generated from the reduced decision table have the following form:

rule #1: if
$$c_1=5$$
 then d=A

rule #2: if $c_3=5$ then d=A

rule #3: if $c_1=4$ and $c_7=1$ then d=A

rule #4: if $c_1=4$ and $c_3=4$ and $c_7=2$ then d=A or R

rule #5: if $c_1=3$ then d=R

rule #6: if $c_3=3$ then d=R

Five rules are deterministic and one is non-deterministic. The non-deterministic rule #4 follows from indiscernibility of candidates \mathbf{x}_8 and \mathbf{x}_9 which belong to different categories of decision. It defines a profile of candidates which should create the third category of decision, e.g. those candidates who should be invited to an interview.

The rules represent clearly the following policy of the selection committee:

Admit all candidates having score 5 in mathematics or in English. Admit also those who have score 4 in mathematics and in English but very good opinion from a previous school. In the case of score 4 in mathematics and in English but only a moderate opinion from a previous school, invite the candidate to an interview. Candidates having score 3 in mathematics or in English are to be rejected.

The considered sample of fifteen candidates can be considered as a training sample used to reveal the selection policy of the committee. This policy could be applied next to a larger set of candidates.

CONCLUDING REMARKS

The aim of this paper was to show that the rough set theory is a useful tool for analysis of decision situations, in particular multi-criteria sorting problems. This class of decision situations has a very large practical representation.

The main advantages of the rough set approach can be summarized in the following points:

- it analyses only facts hidden in the representation of a decision situation,
- it does not need any additional information like probability in statistics or grade of membership in fuzzy set theory,
- it does not correct vagueness manifested in the representation of a decision situation; instead, produced rules are categorized into deterministic and non-deterministic,
- it can be used to detect concordant criteria or experts,
- it gives reducts of independent criteria having the same discernment ability as the whole set.
- it can explain a decision policy,
- it is conceptually simple and needs simple algorithms.

REFERENCES

Boryczka, M., "Rough sets in multi-criteria decision problems",

Bulletin of the Polish Academy of Sciences, ser. Technical

Sciences (1989) 37:321-331.

Boryczka, M., Slowinski, R., "Derivation of optimal decision algorithms from decision tables using rough sets", Bulletin of the Polish Academy of Sciences, ser. Technical Sciences (1988) 36:251-260.

Fibak, J., Pawlak, Z., Slowinski, K., Slowinski, R., "Rough sets based decision algorithm for treatment of duodenal ulcer by HSV", Bulletin of the Polish Academy of Sciences, ser. Biological Sciences (1986) 34:227-246.

Grzymala-Busse Grzymala-Busse Grzymala-Busse J.W., "LERS - a system for learning from examples based on rough sets", in: Slowinski, R., ed. Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory (Dordrecht/Boston/London: Kluwer Academic Publishers 1992, 3-18).

Krysinski, J., Rough sets approach to the analysis of the structure-activity relationship of quaternary imidazolium compounds", Arzneimittel-Forschung/Drug Research (1990) 40 (II): 795-799.

Moscarola, J., "Multi-criteria decision aid - two applications in education management", in: S.Zionts, ed., Multiple Criteria

Problem Solving (Berlin: Springer-Verlag, Lecture Notes in Economics and Mathematical Systems, vol.155, 1978, 402-423).

Nowicki, R., Slowinski, R., Stefanowski, J., "Rough sets analysis of diagnostic capacity of vibroacoustic symptoms", Journal of Computers & Mathematics with Applications (1992) 24:109-123.

Pawlak, Z., "Rough Sets", International Journal of Information & Computer Sciences (1982) 11:341-356.

Pawlak, Z., Rough Sets. Theoretical Aspects of Reasoning about

Data (Dordrecht/Boston/London: Kluwer Academic Publishers 1991).

Roy, B., "Decision science or decision aid science", European Journal of Operational Research (1992) Special Issue on Model Validation in Operations Research (to appear).

Slowinski, R., ed. Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory (Dordrecht/Boston/London: Kluwer Academic Publishers 1992).

Slowinski, K., Slowinski, R., Stefanowski, J., "Rough sets approach to analysis of data from peritoneal lavage in acute pancreatitis", Medical Information (1988) 13:143-159.

Slowinski, R., Stefanowski, J., "'RoughDAS' and 'RoughClass' software implementations of the rough sets approach", in: Slowinski, R., ed. Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory (Dordrecht/Boston/London: Kluwer Academic Publishers 1992, 445-456).

Table 1. Decision table composed of sorting examples

Criterion Candidate	c ₁	c ₂	с ₃	c ₄	c ₅	c ₆	c ₇	Decision d
* ₁	4	4	4	4	2	2	1	A
x ₂	3	3	4	3	2	1	1	R
x ₃	3	4	3	3	1	2	2	R
× ₄	5	3	5	4	2	1	2	A
* ₅	4	4	5	4	2	2	1	A
x 6	3	4	3	3	2	1	3	R
x ₇	4	4	5	4	2	2	2	A
× ₈	4	4	4	4	2	2	2	A
x ₉	4	4	4	4	2	2	2	R
x ₁₀	5	3	5	4	2	1	2	A
x ₁₁	5	4	4	4	1	1	2	A
x ₁₂	5	3	4	4	2	2	2	A
х ₁₃	4	3	3	3	3	2	2	R
× ₁₄	3	3	4	3	2	3	3	R
^x 15	4	5	5	4	2	1	1	A

Recently published Research Reports of the Institute of Computer Science, W.U.T.

8/93	Michał Ostrowski, Amorficzny model danych, February 1993.
9/93	Lech T. Polkowski, Maria Semeniuk-Polkowska, Model of Anaphora Resolution Based on Distributional-Algebraic Approach in Analytical Models Theory of Algebraic Linguistics, February 1993.
10/93	Wacław Struk, Koncepcja nawigacji hipertekstowej w wielomedialnych bazach danych, luty 1993.
11/93	Lech T. Polkowski, Maria Semeniuk-Polkowska. A Formal Approach to Anaphora within Montague Grammar. Anaphoric Sets and Stable Situations, February 1993.
12/93	J. A. Pomykała, J.M. Pomykała, On Regularity of Hypergraph Sequences, February 1993.
13/93	Jarosław Chudziak, Henryk Rybiński, James Vorbach, On Mapping Semantic Data Model Schemata into First-Order Logic, February 1993.
14/93	Maciej Modrzejewski, Feature Selection Using Rough Sets Theory, March 1993.
15/93	Moumine Ballo, Modelowanie terenu z wykorzystaniem map topograficznych, marzec 1993.
16/93	Moumine Ballo, Nowy algorytm wyznaczania części wspólnej dwóch wielokątów prostych, marzec 1993.
17/93	Piotr Sapiecha, Boltzmann Machine for Maximal Queens Problem, March 1993.
18/93	Andrzej Skowron, Zbigniew Suraj, A Rough Set Approach to Real-Time State, Identification for Decision Making, March 1993.
19/93	Mieczysław Muraszkiewicz, Michał Ostrowski, Henryk Rybiński, Handling Amorphous Data Structures in Unified Information Systems, March 1993.
20/93	Zdzisław Pawlak, Rough Sets Present State and the Future, March 1993.