Rough Sets Theory in Decision Analysis

Jiří Křupka
Institute of System Engineering and Informatics
Faculty of Economics and Administration, University of Pardubice

Abstract

This paper presents a possibility of using of the rough sets theory for a reduction of attributes (criteria, data, information) of objects in multi-attributes decision analysis in a public administration system. It is investigated one natural dimension of reducing data which is to identify equivalence classes, i.e. objects that indiscernible using the available attributes, using approximation sets.

Introduction

During last few years trends in public administration management have over come more waves with the aim to use advantages from information technology application. Modern era is demanding changes in organization management philosophy. A re-engineering has become a well known process with organization changes of a public administration structure with aim to change the classical management hierarchy especially on the level of central management. This level has been replaced by information technologies. In connection with this fact structure have been created which were able to react faster to changes and decide independently. The re-engineering may be realized by a process (a knowledge engineering) which expert knowledge is obtained, represented, refined and installed in computer-based diagnostic and decision-making systems. The term knowledge engineering has been used in the field of decision analysis [5,12]. Decision analysis[1] is an engineering discipline that addresses the pragmatics of applying decision theory to real-world problems and may be applied to scheduling, capital expansion, and research and development decisions.

The five important elements of the decision analysis are: the objectives, the alternatives, the impact, the criteria and the model [12]. In the search for the best alternative may be viewed as taking place in five stages (phases) that, in logical order, are: formulation (clarifying and constraining the problem and determining the objective), search (identifying, designing, and screening the alternatives), forecasting (predicting the future environment or operational context), modelling (building and using models to determine the impacts), and synthesis (comparing and ranking the alternatives).

Generally speaking, a decision problem [12] involves a set of objects (actions, courses of action, states, competitors, etc.) described or evaluated by a set of attributes (criteria, features, issues, etc.). Independently of further interpretation, a decision situation may be represented by a table rows which correspond to objects and columns to attributes; for each pair (object-attribute) there is known a value called descriptor. We can also say that the table represents knowledge about a decision situation. Typically, one or several decision makers (experts, agents, nature, etc.) are also involved in a decision problem. By a decision maker we understand a person or thing that works to reduce a result (observation, decision, evaluation, etc.). The attributes used to describe objects are build on some elementary features of objects. They may be nominal (also called categorical or quantitative, e.g. male or female) or cardinal

[1] The discipline of decision analysis emerged in the 1960s; it grew out of a recognition that probability and decision theory, until now applied primarily to problems of statistical estimation, also could be applied to real-world decision problems.

(also called non-nominal or quantitative, e.g. finance rations or temperature). It is possible to distinguish the following tree, most frequently decision problems: classification [4,9,13], choice and ranking.

The aim of the decision analysis is to answer the following two basic questions [2,4,13] related to a decision problem. The most general question is probability about explanation of a decision situation. Explanation means discovering important facts and dependencies in the table describing a decision situation. A more specific question is about prescription of some basing on analysis of information from the table. If this information can be interpreted as preferential model which represents a decision policy of the agent and can be used to support new decision.

There are two ways [13] of constructing a comprehensive preference model upon preferential information obtained from a decision maker involved in the decision process. The first one comes from mathematical decision analysis and consists in building a functional (on the basis a multi-attribute utility theory) or relation (on the basis an outranking relation or fuzzy relation) model. The second way comes from artificial intelligence up the comprehensive preference model via inductive knowledge acquisition (also called rule induction, inductive learning or learning from examples). The resulting model is a set of "if … then …" rules or decision tree. This way is motivated by the hypothesis that the comprehensive preference model can be inferred by studying global evaluations made by the decision maker when presented with a set of representative objects from the problem domain of interest (examples).

The information about a decision situation is usually vague because of uncertainty and imprecision coming from many sources. Vagueness may be caused by granularity of representation of the information [13]. Granularity may introduce an ambiguity (inconsistency) to explanation or prescription based on the vague information. A formal framework for dealing with granularity of information has been given by the rough set theory[2]. The rough set theory assumes a representation of the information in a table form called information system. Rows of this table correspond to object and columns to attributes. The table is just an appropriate form of description of decision situation. Rough set theory has proved to be a useful tool for analysis of a large class of multi-attribute decision problems.

For algorithmic reasons [4], the information regarding the objects is supplied in the form of an information table, whose separate refer to distinct objects (actions), and whole columns refer to the different attributes considered. Each cell of this table indicates, therefore, an evaluation (quantitative or qualitative) of object placed in that row by means of the attribute in the corresponding column. In the case of quantitative evaluations on an attribute q, the domain of this attribute is suitably divided into and then codified, e.g. by natural numbers. This pre-processing of data, called discretization, is commonly used in machine learning.

Rough Sets Theory

A data set may be represented as tables [1,2,3]. These tables are called an information system, a decision table etc. More formally, the information system is the 4-tuple $S=(U, Q, V, f)$, where U is a finite sets of objectives (universe), $Q= \{q_1, q_2,..., q_m\}$ is a finite set of attributes, V_q is the domain of the attribute q, $V = \bigcup_{q\in Q} V_q$ and $f: U{\times}Q\rightarrow V$ is a total function such that $f(x,q) \in V_q$ for each $q \in Q$, $x \in U$, called information function. Therefore, each object x

[2] The theory of rough sets was originated by Zdzislaw Pawlak in 1970's as a result of a long term program of fundamental research on logical properties of information systems, carried out by him and a group of logicians from Polish Academy of Sciences and the University of Warsaw, Poland.

of U is described by a vector (string) $Des_Q(x)= [f(x,q_1), f(x,q_2), ..., f(x,q_m)]$, called description of x in terms of the evaluations of attributes from Q; it represents the available information about x. Obviously, $x \in U$ can be described in terms of any non-empty subset $B \subseteq Q$.

Let [2,6,7] R be a binary relation $R \subseteq X \times X$ which is reflexive (i.e. an object is in relation with itself xRx), symmetric (if xRy then yRx) and transitive (if xRy and yRz then xRz) is called an equivalence relation. The equivalence class of an element $x \in X$ consist of all objects $y \in X$ such that xRy. Let S be a information system, then with any $B \subseteq S$ there is associated an equivalence relation $IND_S(B)$:

$$IND_S(B)= \{(x,x') \in U^2 : \forall a \in B, \ \ a(x)=a(x')\}, \qquad (1)$$

where $IND_S(B)$ is called the B-indiscernibility relation. If $(x,x') \in IND_S(B)$, the object x and x' are indiscernible from each other by attributes from B. The equivalence classes of the B-indiscernibility relation are denote $[x]_B$. An equivalence relation induces a partitioning of the universe (the set of cases in our example). The partitions can be used to build new subsets of the universe. Subsets that are most often of interest have the same value of the outcome attribute.

If we put into the information system a posteriori knowledge is expressed by one distinguished attribute – decision attribute we define decision system as the kind of the information system. The form of the decision system is $S=(U, Q \cup \{d\}, V, f)$, where $d \notin Q$ is the decision attribute. The elements of Q are called conditional attributes or simply conditions. The decision attribute may take several values though binary outcomes are rather frequent. A decision system (i.e. a decision table) expresses all the knowledge about the model. This table may be unnecessarily large in part it is redundant in at least two way. The same or indiscernible objects may be represented several times, or may be superfluous.

The original rough sets idea has proved to be particularly useful in the analysis of multiattribute classification problem. However, rough sets theory can be used for problems of multicriteria decision-making, like sorting, choice or ranking [2]. The philosophy of rough sets is based on assumption that with every object of the universe there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for object description. Objects having the same description are indiscernible (similar) with respect to the constitute their description.

The primary notions of the rough sets theory are the approximation space and lower and upper approximations of a set. The approximation space is a classification of the domain of interest into disjoint categories. The classification formally represents our knowledge about the domain, i.e. the knowledge is understood here as an ability to characterize all classes of the classification, for example, in terms of features of objects belonging to the domain. Objects belonging to the same category are not distinguishable, which means that their membership status with respect to an arbitrary subset of the domain may not always be clearly definable [16]. This fact leads to the definition of a set in terms of lower and upper approximations. The lower approximation is a description of the domain objects which are known with certainty to belong to the subset of interest, whereas the upper approximation is a description of the objects which possibly belong to the subset. Any subset defined through its lower and upper approximations is called rough set. It must be emphasized that the concept of rough should not be confused with the idea of fuzzy set as they are fundamentally different, although in some sense complementary, notions.

Let S be a information system [6,14] and let $B \subseteq S$ and $X \subseteq U$. We can approximate X using only the information contained in B by constructing the B-lower and B-upper approximations of X, denoted $B_*(X)$ and $B^*(X)$ respectively [8,10,11,14], where

$$B_*(X) = \{x : [x]_B \subseteq U\} \text{ and } B^*(X)=\{x : [x]_B \cap X \neq \emptyset\}. \quad (2)$$

The objects in $B_*(X)$ can be with certainty classified as members of X on the basis of knowledge in B, while the objects in $B^*(X)$ can be only classified as possible members of X on the basis of knowledge in B. The set

$$BN_B(X) = B^*(X) - B_*(X) \quad (3)$$

is called the boundary region of X [6], and thus consists of those objects that we can't decisively classify into X on the basis of knowledge in B. The set

$$OUT_B(X) = U - B^*(X) \quad (4)$$

is called the B-outside region of X and consists of those objects which can be with certainty classified as do not belonging to X (on the basis of knowledge in B). A set is said to be rough (respectively crisp) if the boundary region in non-empty (respectively empty)[3].

One can easily show the following properties of approximation [6] (where $-X$ denote U-X):

$$B_*(X) \subseteq X \subseteq B^*(X), \quad (5)$$

$$B_*(\emptyset)= B^*(\emptyset)=\emptyset, B_*(U)=B^*(U)=U, \quad (6)$$

$$B^*(X \cup Y)= B^*(X) \cup B^*(Y), B_*(X \cap Y) = B_*(X) \cap B_*(Y), \quad (7)$$

$$X \subseteq Y \text{ implies } B_*(X) \subseteq B_*(Y) \text{ and } B^*(X) \subseteq B^*(Y), \quad (8)$$

$$B_*(X \cup Y) \supseteq B_*(X) \cup B_*(Y), B^*(X \cap Y) \subseteq B^*(X) \cap B^*(Y), \quad (9)$$

$$B_*(-X)= B_*(U-X) = -B^*(X), B^*(-X)= B^*(U-X) = -B_*(X), \quad (10)$$

$$B_*(B_*(X))= B^*(B_*(X))= B_*(X), B^*(B^*(X))= B_*(B^*(X))= B^*(X). \quad (11)$$

It is easily seen that the lower and the upper approximations of a set , are respectively, the interior and the closure of this set in topology generated by the indiscernibility relations.

One can define the following four basic classes of rough sets, i.e. four categories of vagueness [6]:

X is roughly B-definable, if $B_*(X) \neq \emptyset$ and $B^*(X) \neq U$, (12)

X is internally B-undefinable, if $B_*(X)= \emptyset$ and $B^*(X) \neq U$, (13)

X is externally B-undefinable, if $B_*(X) \neq \emptyset$ and $B^*(X)= U$, (14)

X is totally B-undefinable, if $B_*(X)= \emptyset$ and $B^*(X)= U$. (15)

The intuitive meaning of this classification is the following: X is roughly B-definable means that with the help of B we are able to decide for some elements of U that they belong to X and for some elements of U that they belong to $-X$; X is internally B-undefinable means that using B we are able to decide for some elements of U that they belong to $-X$, but we are to decide for any element of U whether it belong to X; X is externally B-undefinable means that using B we are able to decide for some elements of U that they belong to X, but we are unable to decide for any element of U whether it belongs to $-X$; X is totally B-undefinable means that using B we are unable to decide for any element of U whether it belongs to X or $-X$.

Rough set can be also characterized [6] numerically by the following coefficient

$$\alpha_B(X)= |B_*(X)| / |B^*(X)|, \quad 0 \leq \alpha_B(X) \leq 1, \quad (16)$$

[3] The letter B refers to the subset B of the attributes Q. If another subset were chosen , e.g. $F \subseteq Q$, the corresponding names of the relations would have been F-boundary region, F-lower and F-upper approximations.

called the accuracy of approximation, where $|X|$ denoted the cardinality of $X \neq \emptyset$. If $\alpha_B(X)=1$, X is crisp with respect to B (X is precise with respect to B), and otherwise, if $\alpha_B(X)<1$, X is rough with respect to B (X is vague with respect to B).

We also define a quality of the approximation of X by means of the attributes as [4]

$$\gamma_B(X) = |B_*(X)| \; / \; |X| \; . \tag{17}$$

The quality $\gamma_B(X)$ represents the relative frequency of the objects correctly classified using the attributes from B. Moreover, we have $0 \leq \alpha_B(X) \leq \gamma_B(X) \leq 1$, $\gamma_B(X) = 0$ if $\alpha_B(X) = 0$ and $\gamma_B(X) = 1$ if $\alpha_B(X) = 1$.

The definition of approximations of a subset $X \subseteq U$ can be extended to a classification, i.e. a partition $Y = \{Y_1, Y_2, \ldots, Y_n\}$ of U. Subsets Y_i , i= 1, 2, …, n, are disjunctive classes of Y. By B-lower and B-upper approximation of Y in S we means sets $B_*(X)Y = \{B_*(X)Y_1, B_*(X)Y_2, \ldots, B_*(X)Y_n\}$ and $B^*(X)Y = \{B^*(X)Y_1, B^*(X)Y_2, \ldots, B^*(X)Y_n\}$ respectively. The coefficient

$$\gamma_B(Y) = \Sigma^n_{i=1} |B_*(X)Y_i| \; / \; |U| \tag{18}$$

is called [4] quality of the approximation of classification Y by set of attributes B, or in short quality of classification. It expresses the ratio of all B-correctly classified objects to all object in the system.

The value of $\mu^B_X(x)$ (rough membership function) may be interpreted analogously to conditional probability and may be understood as the degree of certainty (credibility) to which x belong to X. Between the rough membership function and the approximation of X the following relation ships are valid [2]:

$$B_*(X) = \{x \in U : \mu^B_X(x) = 1\}, \tag{19}$$

$$B^*(X) = \{x \in U : \mu^B_X(x) > 0\}, \tag{20}$$

$$BN_B(X) = \{x \in U : 0 < \mu^B_X(x) < 1\}, \tag{21}$$

$$B_*(-X) = B_*(U-X) = \{x \in U : \mu^B_X(x) = 0\}. \tag{22}$$

In the rough sets theory there is, therefore, a close link between vagueness (granularity) connected with rough approximation of sets and uncertainty connected with rough membership of object to sets.

Conclusions

A decision analysis is a important part of a judgement or decision in the contemporary public administration where decision makers have to make decisions on the basis of huge data. These data are complicated and uncertainty in generally. To be successful that means to choose the best decision, the decision makers have worked with a modern soft computing methods, for example: fuzzy systems, neural networks, genetic algorithms and rough logic[4].

The main specific problems addressed by the theory of rough sets are: representation of uncertain or imprecise knowledge; empirical learning and knowledge acquisition from experience; knowledge analysis in discovery in data; analysis of conflicts; evaluation of the quality of the available information with respect to its consistency and the presence or absence of repetitive data patterns; identification and evaluation of data dependencies;

---

[4] p.5, Fig.1.2. Paradigm shifts from probabilistic uncertainty to general uncertainty: complex reduction. It was published in Chen S. H., *Computation Intelligence in Economics and Finance*. Heidelberg: Springer, 2004, 480pp.

approximate pattern classification; reasoning with uncertainty; information-preserving data or attributes reduction [16].

The other dimension in reduction is to keep only those attributes that preserve the indiscernibility relation and, consequently, set approximation. The rejected attributes are redundant since their removal can't worsen the classification. There is usually several such subsets of attributes and those which are minimal are called reducts. Computing equivalence classes is straightforward. This means that computing reducts it is a non-trivial task that can't be solved by a simple minded increase of computational resources [10,15].

The fact that we work with uncertainty, fuzzy, incomplete and differently structured data and information causes the lack of the presented groups of methods. It is possible to avoid it by applying the theory of rough sets may be used for reduction of attributes in formulation of multi-attribute decision-making during the creation of the classification model of the economic processes (e.g. gross domestic product, an unemployment rate, an agricultural index etc.).

References

[1] Bazan J. G., A Comparison of Dynamic and Non-Dynamic Rough Set Methods for Extracting Laws from Decision Tables. In: Polkowski, L., Skowron, A., *Rough Sets in Knowledge Discovery 1. Methodology and Applications. Part 3: Methods and Applications*. Heidelberg: Physica –Verlag. 1998, s.321-365.

[2] Bazan J. G. etc., *Rough Set Algorithms in Classification Problem*, 1999, s.1-40. http://www.mimuw.edu.pl/~jakubw/bib/RSMethods.pdf

[3] Dűntsch I., Gediga G., *Rough Set Data Analysis*, 1999, s.1-25. http://www.infj.ulst.ac.uk/~cccz23/papers/rsda.pdf

[4] Greco S., Matarazzo B., Słowiński R., The Use of Rough Sets and Fuzzy Sets in MCDM. In: Gal T., et al., *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications*. Dordrecht: Kluwer Academic Publishers, 1999, s.14:1-59.

[5] Horvitz E. J., Breese J. S., Henrion M., Decision Theory in Expert Systems and Artificial Intelligence. *International Journal of Approximate Reasoning- Special Issue*, 1988, s.2:247-302.

[6] Komorowski J., Polkovski L., Skowron A., Rough Sets: A Tutorial, 1999, 112s. http://esslli.let.uu.nl/Courses/skowron/skowron.ps

[7] Kowalczyk W., Rough Data Modelling: A New Technique for Analyzing Data. In: Polkowski, L., Skowron, A., *Rough Sets in Knowledge Discovery 1. Methodology and Applications. Part 3: Methods and Applications*. Heidelberg: Physica –Verlag. 1998, s.400-421.

[8] Lin T. Y., Introduction to the Special Issue on Rough Sets. *International Journal of Approximate Reasoning- Special Issue*, 15, 4, 1996, s.287-288

[9] Machova K., Paralič, J., Basic Principle of Cognitive Algorithms Design. Proc. of *The International Conference Computational Cybernetics ICCC2003*, Siófok, Hungary, August 29-30, 2003, s.245-247.

[10] Øhrn A., Komorowski J., Skowron A., Synak P., The Design and Implementation of a Knowledge Discovery Toolkit Based on Rough Sets – The ROSETTA System. In: Polkowski L., Skowron A., *Rough Sets in Knowledge Discovery 1. Methodology and Applications. Part 3: Methods and Applications*. Heidelberg: Physica –Verlag. 1998, s.376-399

[11] Polkowski L., Skowron A., Rough Mereology - A New Paradigm for Approximate Reasoning. *International Journal of Approximate Reasoning- Special Issue*, 15, 4, 1996, s.333-365.

[12] Quade E. S., *Analysis for Public Decision*. Third edition. New Jersey: Prentice Hall, 1989, 409s.

[13] Słowiński R., Rough Set Analysis of Multi-Attribute Decision Problems. In: Ziarko W. P., *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Proc. of the International Workshop on Rough Sets and Knowledge Discovery RSKD´93, Banff, Alberta, Canada, 12-15 Oct.1993. Heidelberg: Springer–Verlag. 1994, s.136-142.

[14] Yao Y. Y., Two View of the Theory of Rough Sets in Finite Universes. *International Journal of Approximate Reasoning- Special Issue*, 15, 4, 1996, s.291-317.

[15] Yao Y.Y., Generalized Rough Set Models. In: Polkowski, L., Skowron, A., *Rough Sets in Knowledge Discovery 1. Methodology and Applications. Part 3: Methods and Applications*. Heidelberg: Physica – Verlag. 1998, s.286-320.

[16] Ziarko W., What is Rough Sets Theory?. Electronic Bulletin of the Rough Set Community, 1993 http://cs.uregina.ca/~roughset/rs.intro.txt

Kontaktní adresa:
doc. Ing. Jiří Křupka, Ph.D.
Ústav systémového inženýrství a informatiky
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 95
532 10 Pardubice
e-mail: Jiri.Krupka@upce.cz
tel.: 466036515

Recenzoval: prof. Ing. Vladimír Olej, CSc., Ústav systémového inženýrství a informatiky Fakulta ekonomicko-správní, Univerzita Pardubice