THE ROUGH SETS THEORY BASED EXPERT SYSTEMS

Vladimir Brtka, Ph.D*
brtkav@gmail.com
Eleonora Brtka, M.Sc
brtka@sbb.rs
Visnja Ognjenovic, M.Sc
visnjaognjenovic@gmail.com
University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Department of Computer Science, Republic of Serbia

ABSTRACT

The paper deals with the decision rules synthesis in the domain of expert systems and knowledge based systems. These systems incorporate expert knowledge which is often expressed in the If Then form. As it is very hard for experts to formally articulate their knowledge, automated decision rule composing algorithms have been used. The rule composing algorithms are often based on the rough sets theory. Originally, rules are composed from data table by equivalence relation, while in this paper we investigate the rules based on dominance relation. The main goal of this paper is to single out possible benefits and advantages of dominance relation based rules over equivalence relation based rules.

Key words: expert systems, knowledge based systems, the rough sets theory, decision rules

INTRODUCTION

The history of the use of computers in engineering problems goes together with the developments in computer hardware and software technology. Engineer utilize principles of science and mathematics to develop new technologies which are then used to create products, structures, machines, processes or even entire systems. It is well known that different tasks in engineering problem solving require different computational tools. Inference from a set of facts, which simulate intelligent decision making, plays a important role in some problem-solving tasks, which involve creativity, while creativity implies the ability to produce novel solutions which are better than previous solutions. These computational tools should be able to use expert knowledge of the problem domain for decision making. The software tools that involve expert knowledge and inference mechanism are called expert systems (ES). ES deals with knowledge processing and complex decision-making problems. The “bottleneck” of the expert systems is a problem of formal articulation of the expert knowledge. It is very hard for the expert to formally express knowledge. Broader based systems which use different knowledge sources are called Knowledge Based Systems (KBS). This paper investigates the If Then form based decision rules, defined or expressed by experts. These rules are readable and easy to understand (Luger, Stubblefield, 1993, Jones, 2008). It is well known that rule synthesis is possible by usage of the rough sets theory. Originally, the rough sets theory based rule induction used equivalence relation, while in this paper we investigate dominance relation. The main goal of this paper is to point out the differences between equivalence relation and dominance relation based rules.

THE ROUGH SETS THEORY

The rough sets theory was developed in the early 1980s (Pawlak et al.). This new approach proved to be very useful for the data analysis in various domains. It is of importance to artificial intelligence (AI) and cognitive sciences in the domains of machine learning, knowledge acquisition, data mining, decision analysis, expert systems, decision support systems, pattern recognition and inductive reasoning. Many
results show that methods based on rough set model are most appropriate, especially in domains of medicine (Ohrn, Brtka et al.) and computer sciences (Dobrilovic et al.). In many cases, data is represented in the form of a flat table with rows containing objects and columns containing attributes. Each combination of an object and an attribute can be characterized by a value, usually numerical, corresponding to a state. The rough set theory may be used to extract knowledge hidden in data and express it in the form of rules. The If Then rules are very useful because they can be inspected and interpreted directly, and the results of decisions are easy to explain. In the rough sets theory there is, usually one, attribute called decision attribute while the other attributes are called condition attributes. Information systems of this kind are called decision systems (Pawlak, Skowron, Greco, Benedetto and Slowinski).

The Indiscernibility Relation

The indiscernibility relation is the mathematical basis of the rough sets theory. Every object of the universe is described by certain amount of information expressed by means of some attributes used for object description. Objects characterized by the same information are indiscernible in view of the available information about them. As in (Greco, Benedetto, Slowinski, 1998), let $U$ be a universe (finite set of objects), $Q = \{q_1, q_2, \ldots, q_m\}$ is a finite set of attributes, $V_q$ is the domain of attribute $q$ and $V = \bigcup_{q \in Q} V_q$. An information system is the 4-tuple $S = (U, Q, V, f)$ where $f = U \times Q \rightarrow V$ is a total function such that $f(x, q) \in V_q$ for each $q \in Q, x \in U$, called information function. Each object of universe is described by a vector: $\text{Des}_q(x) = [f(x, q_1), f(x, q_2), \ldots, f(x, q_m)]$, where $x \in U$. To every non-empty subset of attributes $P$ is associated an indiscernibility relation on $U$, denoted by $I_P$:

$$I_P = \{(x, y) \in U \times U : f(x, q) = f(y, q), \forall q \in P\}$$

(1)

This relation is an equivalence relation (reflexive, symmetric and transitive). The family of all the equivalence classes of the $I_P$ is denoted by $U/I_P$ and class containing an element $x$ by $I_P(x)$.

Set Approximations and Data Reduction

Formally, let $X$ be a non-empty set of $U$ and $\emptyset \neq P \subseteq Q$. Set $X$ is approximated by means of $P$–lower (2) and $P$–upper (3) approximations of $X$:

$$P(X) = \{x \in U : I_P(x) \subseteq X\}$$

(2)

$$\overline{P}(X) = \bigcup_{x \in X} I_P(x)$$

(3)

The $P$–boundary of $X$ is denoted by $Bn(X)$:

$$Bn(X) = \overline{P}(X) - P(X)$$

(4)

Following relation holds: $P(X) \subseteq X \subseteq \overline{P}(X)$. So, if an object $x$ belongs to lower approximation of $X$, it is certainly an element of $X$, but if $x$ belongs to upper approximation of $X$, it may belong to the set $X$. $P$–boundary set of $X$ constitutes “doubtful region” – nothing can be said with certainty about the belonging of its elements to the set $X$.

One approach to data reduction is to identify equivalence classes. Savings are to be made since only one element of the equivalence class is needed to represent the entire class. Another important issue is data reduction. Main problem is how to keep only those attributes that preserve the indiscernibility relation (1) and consequently, set approximation. The rejected attributes are redundant (superfluous) since their removal cannot worsen the classification. Let $\emptyset \neq P \subseteq Q$ and $a \in P$. Attribute $a$ is superfluous in $P$ if $I_P = I_{P_{\{a\}}}$.

Usually, there are several subsets of such attributes and those that are minimal are called reducts.

Decision Rules Synthesis

Each object $x$ that belongs to a decision system determines one decision rule:

430
\[ a = a(x) \Rightarrow d = d(x) \text{, here } a(x) \text{ stands for the value of attribute } a \text{ of an object } x. \] The expression \( a = a(x) \) is called descriptor. If there is one decision attribute \( d \) we have:

\[ \bigwedge_{a \in C} a = a(x) \Rightarrow d = d(x). \]

Now, it is possible to investigate rules of the form: IF \( \alpha \) THEN \( \beta \). Here \( \alpha \) (rule’s antecedent) denotes a conjunction (AND logical operator) of descriptors that only involve attributes of some reduct and \( \beta \) (rule’s consequent) denote a descriptor \( d=d(x) \), where \( d \) is decision attribute. Once the reducts have been computed the rules are composed by overlaying every reduct over the decision table and simply reading the values. Rules supported by lower approximation are called exact, rules supported by boundary region are called inexact because they involve OR logical operator in the Then part.

### SET APPROXIMATIONS BY MEANS OF DOMINANCE RELATIONS

As in (Greco, Benedetto, Slowinski, 1998) let \( S_a \) be an outranking relation on universe \( U \), so that for condition attribute \( a \in C \) and objects \( x, y \in U \) we have \( xS_a y \), which means: “\( x \) is at least as good as \( y \) with respect to attribute (criterion) \( a \)”. Let \( P \subseteq C \) : object \( x \) dominates object \( y \) (denotation \( xD_P y \)), if \( xS_a y \) stands for every \( a \in P \). Now, it is possible to define sets:

\[
D^+_p(x) = \{ y \in U : yD_p x \} \quad (5)
\]

\[
D^-_p(x) = \{ y \in U : xD_p y \} \quad (6)
\]

Let \( Cl = \{ Cl_t, t \in T \}, \ T = \{1, \ldots, n\} \) be a set of classes of \( U \), which means that each element of \( U \) belongs to one and only one class. For \( x, y \in U \) and \( r, s \in T \) we have:

\[
(x \in Cl_r, y \in Cl_s, r > s) \Rightarrow (xSy \land \neg(ySx)),
\]

It is possible to define a set:

\[
Cl^2_t = \bigcup_{s \geq t} Cl_s \quad (7)
\]

Furthermore, it is possible to define \( P \)-lower (8) and \( P \)-upper (9) approximations of \( X \) :

\[
P(Cl^2_t) = \{ x \in U : D^+_p(x) \subseteq Cl^2_t \} \quad (8)
\]

\[
\overline{P}(Cl^2_t) = \bigcup_{x \in Cl^2_t} D^-_p(x) \quad (9)
\]

By analogy, for:

\[
Cl^2_t = \bigcup_{s \leq t} Cl_s \quad (10)
\]

we have:

\[
P(Cl^2_t) = \{ x \in U : D^-_p(x) \subseteq Cl^2_t \} \quad (11)
\]

\[
\overline{P}(Cl^2_t) = \bigcup_{x \in Cl^2_t} D^+_p(x) \quad (12)
\]

It is quite obvious that rules of the If Then form can be induced based on set approximations by means of dominance relation. The question is: What is an exact form of these rules and what is the difference to If Then rules induced by indiscernibility relation (1)?

### CONCLUSION

By approximations (8), (9), (11) and (12) which were obtained by dominance relations (5) and (6) it is possible to induce following If Then rules supported by:

1. \( P \)-lower approximations of the classes (7),

\[
\text{If } f(x, a_1) \geq v_{a_1} \text{ and } \ldots \text{ and } f(x, a_n) \geq v_{a_n} \text{ Then } x \in Cl^2_t
\]
2. P–lower approximations of the classes (10),

\[
\text{If } f(x, a_i) \leq v_{a_i} \text{ and } \ldots \text{ and } f(x, a_n) \leq v_{a_n} \text{ Then } x \in Cl^2_i
\]

3. P–boundaries of the classes (7) and (10),

\[
\text{If } f(x, a_i) \geq v_{a_i} \text{ and } \ldots \text{ and } f(x, a_k) \geq v_{a_k} \text{ and }
\]

\[
f(x, a_{k+1}) \leq v_{a_{k+1}} \text{ and } \ldots \text{ and } f(x, a_n) \leq v_{a_n} \text{ Then } x \in Cl^2_i \text{ or } x \in Cl^2_j
\]

This approach in general, gives more synthetic representation of knowledge contained in the decision table than the set of rules induced by classical approach based on equivalence (indiscernibility) relation (Greco et al.).

Possible benefits and advantages of the dominance based rules over equivalence based rules are:

- the minimal set of rules contains a smaller number of rules,
- smaller number of used condition attributes in the If parts of the rules,
- easier to understand and interpret,
- possibility to express more general knowledge.

The drawback is increased computational time needed to generate rules. Future work will include a further development of the software application for rule synthesis based on dominance relation. This application is a part of web based system for data analysis.

ACKNOWLEDGMENT

This research is financially supported by Ministry of Science and Technological Development, Republic of Serbia, under the project number TR32044 “The development of software tools for business process analysis and improvement”.

REFERENCES


