Boolean Reasoning Scheme with Some Applications in Data Mining

Andrzej Skowron and Hung Son Nguyen

Institute of Mathematics, Warsaw University,
Banacha 2, 02-097 Warsaw, Poland
Email: {skowron,son}@mimuw.edu.pl

Abstract. We present a general encoding scheme for a wide class of problems (including among others such problems like data reduction, feature selection, feature extraction, decision rules generation, pattern extraction from data or conflict resolution in multi-agent systems) and we show how to combine it with a propositional (Boolean) reasoning to develop efficient heuristics searching for (approximate) solutions of these problems. We illustrate our approach by examples, we show some experimental results and compare them with those reported in literature. We also show that association rule generation is strongly related with reduct approximation.

1 Introduction

We discuss a representation scheme for a wide class of problems including problems from such areas like decision support [14], [9], machine learning, data mining [4], or conflict resolution in multi-agent systems [10]. On the basis of the representation scheme we construct (monotone) Boolean functions with the following property: their prime implicants [3] (minimal valuations satisfying propositional formulas) directly correspond to the problem solutions (compare the George Boole idea from 1848 discussed e.g. in [3]). In all these cases the implicants close to prime implicants define approximate solutions for considered problems (compare the discussion on Challenge_9 in [12]).

The results are showing that the efficient heuristics for feature selection, feature extraction, pattern extraction from data can be developed using Boolean propositional reasoning. Moreover the experiments are showing that these heuristics can give better results concerning classification quality or/and time necessary for learning (discovery) than those derived using other methods. Our experience is showing that formulations of problems in the Boolean reasoning framework creates a promising methodology for developing very efficient heuristics for solving real-life problems in many areas. Let us also mention applications of Boolean reasoning in other areas like negotiations and conflict resolving in multi-agent systems [10].

Because of lack of space we illustrate the approach using two illustrative examples related to symbolic value grouping and association rule extraction in Data Mining (or Machine Learning) problems.

2 Basic Notions

An *information system* is a pair $\mathbb{S} = (U, A)$, where U - is a non-empty, finite set called the *universe*, A - is a non-empty, finite set of *attributes*, i.e., $a : U \rightarrow V_a$ for $a \in A$, where V_a is called the *value set* of a . Elements of U are called *situations* and interpreted as e.g. cases, states, patients, observations. The set $V = \bigcup_{a \in A} V_a$ is said to be the *domain* of A . A *decision table* is any information system of the form $\mathbb{S} = (U, A \cup \{d\})$, where $d \notin A$ is a distinguished attribute called *decision*. The elements of A are called *conditional attributes* (*conditions*).

In a given information system, in general, we are not able to distinguish all pairs of situations objects (using attributes of the system). Namely, different situations can have the same values on considered attributes. Hence, any set of attributes divides the universe U into some classes which establish a partition [9] of the set of all objects U . With any subset of attributes $B \subseteq A$ we associate a binary relation $\text{ind}(B)$, called an *indiscernibility relation*, which is defined by $\text{ind}(B) = \{(u, u') \in U \times U \text{ for every } a \in B, a(u) = a(u')\}$. The *B-discernibility* relation is defined to be the complement of $\text{ind}(B)$ in $U \times U$.

Let $\mathbb{S} = (U, A)$ be an information system, where $A = \{a_1, \dots, a_m\}$. Pairs (a, v) with $a \in A, v \in V$ are called *descriptors*. By $\mathbf{DESC}(A, V)$ we denote the set of all descriptors over A and V . Instead of (a, v) we also write $a = v$ or a_v . One can assign Boolean variable to any descriptor. The set of *terms* over A and V is the least set containing descriptors (over A and V) and closed with respect to the classical propositional connectives: $\neg$ (negation), $\vee$ (disjunction), and $\wedge$ (conjunction), i.e.,

1. Any descriptor $(a, v) \in \mathbf{DESC}(A, V)$ is term over A and V .
2. If τ, τ' are terms then $\neg\tau, (\tau \vee \tau'), (\tau \wedge \tau')$ are terms over A and V too.

The meaning $\|\tau\|_{\mathbb{S}}$ (or in short $\|\tau\|$) of τ in $\mathbb{S}$ is defined inductively as follows:

$$\begin{aligned} \|(a, v)\| &= \{u \in U : a(u) = v\} && \text{for } a \in A \text{ and } v \in V_a; \\ \|(\tau \vee \tau')\| &= \|\tau\| \cup \|\tau'\|; && \|(\tau \wedge \tau')\| = \|\tau\| \cap \|\tau'\|; \quad \|\neg\tau\| = U - \|\tau\|. \end{aligned}$$

Two terms τ and τ' are equivalent, $\tau \Leftrightarrow \tau'$, if and only if $\|\tau\| = \|\tau'\|$. In particular we have: $\neg(a = v) \Leftrightarrow \bigvee \{a = v' : v' \neq v \text{ and } v' \in V_a\}$.

The information systems (decision tables) are representations of knowledge bases discussed in Introduction: rows corresponds to consistent sets of propositional variables defined by all descriptors $a = v$ where v is the value of attribute a in a given situation and conflicting pairs, in case of information systems, are all pairs of situations which are discernible by some attributes.

Let $\mathbb{S} = (U, A)$ be an information system, where $U = \{u_1, \dots, u_n\}$, and $A = \{a_1, \dots, a_m\}$. By $\mathbb{M}(\mathbb{S})$ we denote an $n \times n$ matrix (c_{ij}) , called the *discernibility matrix* of $\mathbb{S}$, such that $c_{ij} = \{a \in A : a(u_i) \neq a(u_j)\}$ for $i, j = 1, \dots, n$.

With every discernibility matrix $\mathbb{M}(\mathbb{S})$ one can associate a *discernibility function* $f_{\mathbb{M}(\mathbb{S})}$, defined as follows. A *discernibility function* $f_{\mathbb{M}(\mathbb{S})}$ for an information system $\mathbb{S}$ is a Boolean function of m propositional variables $a_1^*, \dots, a_m^*$ (where $a_i \in A$ for $i = 1, \dots, m$) defined as the conjunction of all expressions $\bigvee c_{ij}^*$, where

$\bigvee c_{ij}^*$ is the disjunction of all elements of $c_{ij}^* = \{a^* : a \in c_{ij}\}$, where $1 \leq j < i \leq n$ and $c_{ij} \neq \emptyset$. In the sequel we write a instead of a^* .

One can show that every prime implicant of $f_{\mathbb{M}(\mathbb{S})}(a_1^*, \dots, a_k^*)$ corresponds exactly to one reduct in $\mathbb{S}$. One can see that the set $B \subset A$ is reduct if B has nonempty intersection with any nonempty set $c_{i,j}$ i.e.

$$B \text{ is reduct in } \mathbb{S} \quad \text{iff} \quad \forall_{i,j} (c_{i,j} = \emptyset) \vee (B \cap c_{i,j} \neq \emptyset)$$

One can show that prime implicants of the discernibility function correspond exactly to reducts of information systems [9], [14]. Hence, Boolean reasoning can be used for information reduction. This can be extended to feature selection and decision rule synthesis (see e.g. [2], [9]). One can show that the problem of finding a minimal (with respect to cardinality) reduct is NP-hard [14]. In general the number of reducts of a given information system can be exponential with respect to the number of attributes (more exactly, any information system $\mathbb{S}$ has at most $\binom{m}{\lfloor m/2 \rfloor}$ reducts, where $m = \text{card}(A)$). Nevertheless, existing procedures for reduct computation are efficient in many applications and for many cases one can apply some efficient heuristics (see e.g. [2]).

Moreover, in some applications (see [13]), instead of reducts we prefer to use their approximations called α -reducts, where $\alpha \in [0, 1]$ is a real parameter. The set of attributes $B \subset A$ is called α -reduct if

$$B \text{ is } \alpha\text{-reduct in } \mathbb{S} \quad \text{iff} \quad \frac{|\{c_{i,j} : B \cap c_{i,j} \neq \emptyset\}|}{|\{c_{i,j} : c_{i,j} \neq \emptyset\}|} \geq \alpha$$

One can show that for a given α , the problems of searching for shortest α -reducts and for all α -reducts are also NP-hard.

Let us note that e.g. simple greedy Johnson strategy for computing implicants close to prime implicants of the discernibility function has time complexity of order $O(k^2 n^3)$ where n is the number of objects and k is the number of attributes. Hence, for large n this heuristic will be not feasible. We will show how to construct some more efficient heuristics in case when some additional knowledge is given about problem encoded by information system or decision table.

3 Feature Extraction by Grouping of Symbolic Values

In case of symbolic value attribute (i.e. without pre-assumed order on values of given attributes) the problem of searching for new features of the form $a \in V$ is, in a sense, from practical point of view more complicated than the for real value attributes. However, it is possible to develop efficient heuristics for this case using Boolean reasoning.

Let $\mathbb{S} = (U, A \cup \{d\})$ be a decision table. Any function $P_a : V_a \rightarrow \{1, \dots, m_a\}$ (where $m_a \leq \text{card}(V_a)$) is called a *partition* of V_{a_i} . The *rank* of P_{a_i} is the value $\text{rank}(P_i) = \text{card}(P_{a_i}(V_{a_i}))$. The family of partitions $\{P_a\}_{a \in B}$ is *consistent with B* (*B-consistent*) iff the condition $[(u, u') \notin \text{ind}(B/\{d\})]$ implies $\exists_{a \in B} [P_a(a(u)) \neq P_a(a(u'))]$ holds for any $(u, u') \in U$. It means that if two objects u, u' are discerned by B and d , then they must be discerned by partition attributes defined by $\{P_a\}_{a \in B}$. We consider the following optimization problem

PARTITION PROBLEM: SYMBOLIC VALUE PARTITION PROBLEM:

Given a decision table $\mathbb{S} = (U, A \cup \{d\})$ and a set of attributes $B \subseteq A$, search for the minimal B -consistent family of partitions (i.e. such B -consistent family $\{P_a\}_{a \in B}$ that $\sum_{a \in B} \text{rank}(P_a)$ is minimal).

To discern between pair of objects will use new binary features $a_v^{v'}$ (for $v \neq v'$) defined by $a_v^{v'}(x, y) = 1$ iff $a(x) = v \neq v' = a(y)$. One can apply the Johnson's heuristic for the new decision table with these attributes to search for minimal set of new attributes that discerns all pairs of objects from different decision classes. After extracting of these sets, for each attribute a_i we construct graph $\Gamma_a = \langle V_a, E_a \rangle$ where E_a is defined as the set of all new attributes (propositional variables) found for the attribute a . Any vertex coloring of Γ_a defines a partition of V_a . The colorability problem is solvable in polynomial time for $k = 2$, but remains NP-complete for all $k \geq 3$. But, similarly to discretization[7], one can apply some efficient heuristic searching for optimal partition.

Let us consider an example of decision table presented in Figure 1 and (a reduced form) of its discernibility matrix (Figure 1). From the Boolean function $f_{\mathbb{A}}$ with Boolean variables of the form $a_{v_1}^{v_2}$ one can find the shortest prime implicant: $\mathbf{a}_{a_2}^{a_1} \wedge \mathbf{a}_{a_3}^{a_2} \wedge \mathbf{a}_{a_4}^{a_1} \wedge \mathbf{a}_{a_4}^{a_3} \wedge \mathbf{b}_{b_4}^{b_1} \wedge \mathbf{b}_{b_4}^{b_2} \wedge \mathbf{b}_{b_3}^{b_2} \wedge \mathbf{b}_{b_3}^{b_1} \wedge \mathbf{b}_{b_5}^{b_3}$ which can be treated as graphs presented in the Figure 2. We can color vertices of those graphs as it is shown in Figure 2. The colors are corresponding to the partitions: $P_{\mathbf{a}}(a_1) = P_{\mathbf{a}}(a_3) = 1$; $P_{\mathbf{a}}(a_2) = P_{\mathbf{a}}(a_4) = 2$; $P_{\mathbf{b}}(b_1) = P_{\mathbf{b}}(b_2) = P_{\mathbf{b}}(b_5) = 1$; $P_{\mathbf{b}}(b_3) = P_{\mathbf{b}}(b_4) = 2$. At the same time one can construct the new decision table (Figure 2).

$\mathbb{A}$	a	b	d
u_1	a_1	b_1	0
u_2	a_1	b_2	0
u_3	a_2	b_3	0
u_4	a_3	b_1	0
u_5	a_1	b_4	1
u_6	a_2	b_2	1
u_7	a_2	b_1	1
u_8	a_4	b_2	1
u_9	a_3	b_4	1
u_{10}	a_2	b_5	1

 $\Rightarrow$

$\mathbb{M}(\mathbb{S})$	u_1	u_2	u_3	u_4
u_5	$\mathbf{b}_{b_4}^{b_1}$	$\mathbf{b}_{b_4}^{b_2}$	$\mathbf{a}_{a_2}^{a_1}, \mathbf{b}_{b_4}^{b_3}$	$\mathbf{a}_{a_3}^{a_1}, \mathbf{b}_{b_4}^{b_1}$
u_6	$\mathbf{a}_{a_2}^{a_1}, \mathbf{b}_{b_2}^{b_1}$	$\mathbf{a}_{a_2}^{a_1}$	$\mathbf{b}_{b_3}^{b_2}$	$\mathbf{a}_{a_3}^{a_2}, \mathbf{b}_{b_2}^{b_1}$
u_7	$\mathbf{a}_{a_2}^{a_1}$	$\mathbf{a}_{a_2}^{a_1}, \mathbf{b}_{b_2}^{b_1}$	$\mathbf{b}_{b_3}^{b_1}$	$\mathbf{a}_{a_3}^{a_2}$
u_8	$\mathbf{a}_{a_4}^{a_1}, \mathbf{b}_{b_2}^{b_1}$	$\mathbf{a}_{a_4}^{a_1}$	$\mathbf{a}_{a_4}^{a_2}, \mathbf{b}_{b_3}^{b_2}$	$\mathbf{a}_{a_3}^{a_3}, \mathbf{b}_{b_2}^{b_1}$
u_9	$\mathbf{a}_{a_3}^{a_1}, \mathbf{b}_{b_4}^{b_1}$	$\mathbf{a}_{a_3}^{a_1}, \mathbf{b}_{b_4}^{b_2}$	$\mathbf{a}_{a_3}^{a_2}, \mathbf{b}_{b_4}^{b_3}$	$\mathbf{b}_{b_4}^{b_1}$
u_{10}	$\mathbf{a}_{a_2}^{a_1}, \mathbf{b}_{b_5}^{b_1}$	$\mathbf{a}_{a_2}^{a_1}, \mathbf{b}_{b_5}^{b_2}$	$\mathbf{b}_{b_5}^{b_3}$	$\mathbf{a}_{a_3}^{a_2}, \mathbf{b}_{b_5}^{b_1}$

Fig. 1. The decision table and the discernibility matrix

Fig. 2. Coloring of attribute value graphs and the reduced table.

One can extend the presented approach (see e.g. [6]) to the case when in a given decision system nominal and numeric attribute appear. The received heuristics are of very good quality.

Experiments for classification methods (see [6]) have been carried over decision systems using two techniques called “*train-and-test*” and “*n-fold-cross-validation*”. In Table 1 some results of experiments obtained by testing the proposed methods MD (using only discretization based on MD-heuristic [7] using Johnson approximation strategy) and MD-G (using discretization and symbolic value grouping) for classification quality on well known data tables from the “UC Irvine repository” are shown. The results reported in [5] are summarized in columns labeled by S-ID3 and C4.5 in Table 1). It is interesting to compare those results with regard both to the classification quality. Let us note that the heuristics MD and MD-G are also very efficient with respect to the time complexity.

Names of Tables	Classification accuracies			
	S-ID3	C4.5	MD	MD-G
Australian	78.26	85.36	83.69	84.49
Breast (L)	62.07	71.00	69.95	69.95
Diabetes	66.23	70.84	71.09	76.17
Glass	62.79	65.89	66.41	69.79
Heart	77.78	77.04	77.04	81.11
Iris	96.67	94.67	95.33	96.67
Lympho	73.33	77.01	71.93	82.02
Monk-1	81.25	75.70	100	93.05
Monk-2	69.91	65.00	99.07	99.07
Monk-3	90.28	97.20	93.51	94.00
Soybean	100	95.56	100	100
TicTacToe	84.38	84.02	97.7	97.70
Average	78.58	79.94	85.48	87.00

Table 1. The quality comparison between decision tree methods. MD: MD-heuristics; MD-G: MD-heuristics with symbolic value partition

4 Association Rule Generation

Given an information table $\mathbb{A} = (U, A)$. By *descriptors* we mean terms of form $(a = v)$, where $a \in A$ is an attribute and $v \in V_a$ is a value in the domain of a (see [8]). The notion of descriptor can be generalized by using terms of form $(a \in S)$, where $S \subseteq V_a$ is a set of values. By *template* we mean the conjunction of descriptors, i.e. $\mathbf{T} = D_1 \wedge D_2 \wedge \dots \wedge D_m$, where $D_1, \dots, D_m$ are either simple or generalized descriptors. We denote by $length(\mathbf{T})$ the number of descriptors being in $\mathbf{T}$. An object $u \in U$ is satisfying the template $\mathbf{T} = (a_{i_1} = v_1) \wedge \dots \wedge (a_{i_m} = v_m)$ if and only if $\forall_j a_{i_j}(u) = v_j$. Hence the template $\mathbf{T}$ describes the set of objects having the common property: “*the values of attributes $a_{j_1}, \dots, a_{j_m}$ on these objects are equal to $v_1, \dots, v_m$, respectively*”. The *support* of $\mathbf{T}$ is defined by $support(\mathbf{T}) = |\{u \in U : u \text{ satisfies } \mathbf{T}\}|$. The long templates with large support are preferred in many Data Mining tasks.

Problems of finding optimal large templates (for many optimization functions) are known as being NP-hard with respect to the number of attributes

involved into descriptors (see e.g. [8]). Nevertheless, the large templates can be found quite efficiently by *Apriori* and *Apriori Tid* algorithms (see [1,15]). A number of other methods for large template generation has been proposed e.g. in [8].

Association rules and their generations can be defined in many ways (see [1]). Here, according to the presented notation, association rules can be defined as implications of the form $(\mathbf{P} \Rightarrow \mathbf{Q})$, where $\mathbf{P}$ and $\mathbf{Q}$ are different simple templates, i.e. formulas of the form

$$(a_{i_1} = v_{i_1}) \wedge \dots \wedge (a_{i_k} = v_{i_k}) \Rightarrow (a_{j_1} = v_{j_1}) \wedge \dots \wedge (a_{j_l} = v_{j_l}) \quad (1)$$

These implication can be called *generalized association rules*, because association rules are originally defined by formulas $\mathbf{P} \Rightarrow \mathbf{Q}$ where $\mathbf{P}$ and $\mathbf{Q}$ are the sets of items (i.e. goods or articles in stock market) e.g. $\{A, B\} \Rightarrow \{C, D, E\}$ (see [1]). One can see that this form can be obtained from 1 by replacing values on descriptors by 1 i.e.: $(A = 1) \wedge (B = 1) \Rightarrow (C = 1) \wedge (D = 1) \wedge (E = 1)$.

Usually, for a given information table $\mathbb{A}$, the quality of the association rule $\mathcal{R} = \mathbf{P} \Rightarrow \mathbf{Q}$ can be evaluated by two coefficients called *support* and *confidence* with respect to $\mathbb{A}$. The support of the rule $\mathcal{R}$ is defined by the number of objects from $\mathbb{A}$ satisfying the condition $(\mathbf{P} \wedge \mathbf{Q})$ i.e. $\text{support}(\mathcal{R}) = \text{support}(\mathbf{P} \wedge \mathbf{Q})$. The second coefficient – confidence of $\mathcal{R}$ – is the ratio between the support of $(\mathbf{P} \wedge \mathbf{Q})$ and the support of $\mathbf{P}$ i.e. $\text{confidence}(\mathcal{R}) = \frac{\text{support}(\mathbf{P} \wedge \mathbf{Q})}{\text{support}(\mathbf{P})}$. The following problem has been investigated by many authors (see e.g. [1,15]): *For a given information table $\mathbb{A}$, an integer s , and a real number $c \in [0, 1]$, find as many as possible association rules $\mathcal{R} = (\mathbf{P} \Rightarrow \mathbf{Q})$ such that $\text{support}(\mathcal{R}) \geq s$ and $\text{confidence}(\mathcal{R}) \geq c$.* All existing association rule generation methods consists of two main steps:

1. Generate as many as possible templates $\mathbf{T} = D_1 \wedge D_2 \dots \wedge D_k$ such that $\text{support}(\mathbf{T}) \geq s$ and $\text{support}(\mathbf{T} \wedge D) < s$ for any descriptor D (i.e. maximal templates among those which are supported by more than s objects).
2. For any template $\mathbf{T}$, search for a partition $\mathbf{T} = \mathbf{P} \wedge \mathbf{Q}$ such that $\text{support}(\mathbf{P}) < \frac{\text{support}(\mathbf{T})}{c}$ and $\mathbf{P}$ is the smallest template satisfying this condition.

In this paper we show that the second steps can be solved using rough set methods and Boolean reasoning approach.

4.1 Boolean Reasoning Approach for Association Rule Generation

Let us assume that the template $\mathbf{T} = D_1 \wedge D_2 \wedge \dots \wedge D_m$, which is supported by at least s objects, has been found. For a given confidence threshold $c \in (0; 1)$ the decomposition $\mathbf{T} = \mathbf{P} \wedge \mathbf{Q}$ is called *c-irreducible* if $\text{confidence}(\mathbf{P} \Rightarrow \mathbf{Q}) \geq c$ and for any decomposition $\mathbf{T} = \mathbf{P}' \wedge \mathbf{Q}'$ such that $\mathbf{P}'$ is a sub-template of $\mathbf{P}$, $\text{confidence}(\mathbf{P}' \Rightarrow \mathbf{Q}') < c$. One can prove

Theorem 1. *Let $c \in [0; 1]$. The problem of searching for the shortest association rule from the template $\mathbf{T}$ for a given table $\mathbb{S}$ with confidence limited by c (Optimal c -Association Rules Problem) is NP-hard.*

For solving the presented problem, we show that the problem of searching for optimal association rules from the given template is equivalent to the problem of searching for local α -reducts for a decision table, which is well known problem in rough set theory. We construct the new decision table $\mathbb{S}|_{\mathbf{T}} = (U, A|_{\mathbf{T}} \cup d)$ from the original information table $\mathbb{S}$ and the template $\mathbf{T}$ as follows:

1. $A|_{\mathbf{T}} = \{a_{D_1}, a_{D_2}, \dots, a_{D_m}\}$ is a set of attributes corresponding to the descriptors of $\mathbf{T}$ such that $a_{D_i}(u) = \begin{cases} 1 & \text{if the object } u \text{ satisfies } D_i, \\ 0 & \text{otherwise.} \end{cases}$
2. the decision attribute d determines if the object satisfies template $\mathbf{T}$ i.e. $d(u) = \begin{cases} 1 & \text{if the object } u \text{ satisfies } \mathbf{T}, \\ 0 & \text{otherwise.} \end{cases}$

The following theorems describe the relationship between association rules problem and reduct searching problem.

Theorem 2. *For a given information table $\mathbb{S} = (U, A)$, the template $\mathbf{T}$, the set of descriptors $\mathbf{P}$ the implication $(\bigwedge_{D_i \in \mathbf{P}} D_i \Rightarrow \bigwedge_{D_j \notin \mathbf{P}} D_j)$ is*

1. 100%-irreducible association rule from $\mathbf{T}$ if and only if $\mathbf{P}$ is reduct in $\mathbb{S}|_{\mathbf{T}}$.
2. c -irreducible association rule from $\mathbf{T}$ if and only if $\mathbf{P}$ is α -reduct in $\mathbb{S}|_{\mathbf{T}}$, where $\alpha = 1 - (\frac{1}{c} - 1) / (\frac{n}{s} - 1)$, n is the total number of objects from U and $s = \text{support}(\mathbf{T})$.

Searching for minimal α -reducts is well known problem in Rough Sets theory. One can show, that the problem of searching for all α -reducts as well as the problem of searching for shortest α -reducts is NP-hard. Great effort has been done to solve those problems. In the next papers we present the rough set based algorithms for association rule generation for large data table using SQL queries.

4.2 The Example

The following example illustrates the main idea of our method. Let us consider the following information table $\mathbb{A}$ with 18 objects and 9 attributes. Assume that the template $\mathbf{T} = (a_1 = 0) \wedge (a_3 = 2) \wedge (a_4 = 1) \wedge (a_6 = 0) \wedge (a_8 = 1)$ has been extracted from the information table $\mathbb{A}$. One can see that $\text{support}(\mathbf{T}) = 10$ and $\text{length}(\mathbf{T}) = 5$. The new constructed decision table $\mathbb{A}|_{\mathbf{T}}$ is presented in Table 2. The discernibility function for $\mathbb{A}|_{\mathbf{T}}$ can be described as follows

$$\begin{aligned} f(D_1, D_2, D_3, D_4, D_5) = & (D_2 \vee D_4 \vee D_5) \wedge (D_1 \vee D_3 \vee D_4) \wedge (D_2 \vee D_3 \vee D_4) \\ & \wedge (D_1 \vee D_2 \vee D_3 \vee D_4) \wedge (D_1 \vee D_3 \vee D_5) \\ & \wedge (D_2 \vee D_3 \vee D_5) \wedge (D_3 \vee D_4 \vee D_5) \wedge (D_1 \vee D_5) \end{aligned}$$

After its simplification we obtain six reducts: $f(D_1, D_2, D_3, D_4, D_5) = (D_3 \wedge D_5) \vee (D_4 \wedge D_5) \vee (D_1 \wedge D_2 \wedge D_3) \vee (D_1 \wedge D_2 \wedge D_4) \vee (D_1 \wedge D_2 \wedge D_5) \vee (D_1 \wedge D_3 \wedge D_4)$ for the decision table $\mathbb{A}|_{\mathbf{T}}$. Thus, we have found from $\mathbf{T}$ six association rules with (100%)-confidence. For $c = 90\%$, we would like to find α -reducts for the decision

extension of our knowledge bases by adding a new component corresponding to concordance (indiscernible) pairs of situations and to require to preserve some constraints described by this component. We also plan to extend approach using rough mereology [10].

Acknowledgement: This work was partially supported by the Research Program of the European Union - ESPRIT-CRIT2 No. 20288

References

1. Agrawal R., Mannila H., Srikant R., Toivonen H., Verkamo A.I., 1996. Fast discovery of association rules. In V.M. Fayad, G.Piatetsky Shapiro, P. Smyth, R. Uthurusamy (eds): *Advanced in Knowledge Discovery and Data Mining*, AAAI/MIT Press, pp. 307-328.
2. J. Bazan. A comparison of dynamic non-dynamic rough set methods for extracting laws from decision tables. In: L. Polkowski and A. Skowron (Eds.), *Rough Sets in Knowledge Discovery 1: Methodology and Applications*, Physica-Verlag, Heidelberg, 1998, 321-365.
3. E.M. Brown. *Boolean Reasoning*, Kluwer Academic Publishers, Dordrecht, 1990.
4. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds.). *Advances in Knowledge Discovery and Data Mining*. MIT/AAAI Press, Menlo Park, 1996.
5. J. Friedman, R. Kohavi, Y. Yun. Lazy decision trees, *Proc. AAAI-96*, 717-724.
6. H.S. Nguyen and S.H. Nguyen. Pattern extraction from data, *Fundamenta Informaticae* **34**, 1998, pp. 129-144.
7. H.S. Nguyen and A. Skowron. Boolean reasoning for feature extraction problems, *Proc. ISMIS'97*, LNAI 1325, Springer-verlag, Berlin, 117-126.
8. Nguyen S. Hoa, A. Skowron, P. Synak. Discovery of data pattern with applications to Decomposition and classification problems. In L. Polkowski, A. Skowron (eds.): *Rough Sets in Knowledge Discovery 2*. Physica-Verlag, Heidelberg 1998, pp. 55-97.
9. Z. Pawlak. *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht, 1991.
10. L. Polkowski and A. Skowron. Rough sets: A perspective. In: L. Polkowski and A. Skowron (Eds.). *Rough Sets in Knowledge Discovery 1: Methodology and Applications*. Physica-Verlag, Heidelberg, 1998, 31-56.
11. J.R. Quinlan. *C4.5. Programs for machine learning*, Morgan Kaufmann, San Mateo, CA, 1993.
12. B. Selman, H. Kautz and D. McAllester. Ten Challenges in Propositional Reasoning and Search, *Proc. IJCAI'97*, Japan.
13. Skowron A. Synthesis of adaptive decision systems from experimental data. In A. Aamodt, J. Komorowski (eds), *Proc. of the 5th Scandinavian Conference on AI (SCAI'95)*, IOS Press, May 1995, Trondheim, Norway, 220-238.
14. A. Skowron and C. Rauszer. The discernibility matrices and functions in information systems, in: R. Słowiński (Ed.), *Intelligent decision support: Handbook of applications and advances of the rough sets theory*, Kluwer Academic Publishers, Dordrecht, 1992, 331-362.
15. Mohammed Javeed Zaki, Srinivasan Parthasarathy, Mitsunori Ogihara, Wei Li. New Parallel Algorithms for Fast Discovery of Association Rules. In *Data Mining and Knowledge Discovery : An International Journal*, special issue on Scalable High-Performance Computing for KDD, Vol. 1, No. 4, Dec. 1997, pp 343-373.