Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Information Sciences 159 (2004) 255–272

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Approaches to knowledge reduction based on variable precision rough set model

Ju-Sheng Mi [a,b,*], Wei-Zhi Wu [a,c], Wen-Xiu Zhang [a]

[a] *Institute for Information and System Sciences, Faculty of Science, Xi'an Jiaotong University, Xi'an, Shaan'xi 710049, PR China*

[b] *College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, Hebei 050016, PR China*

[c] *Information College, Zhejiang Ocean University, Zhoushan, Zhejiang 316004, PR China*

Received 29 January 2003; received in revised form 26 May 2003; accepted 15 July 2003

Abstract

This paper deals with approaches to knowledge reduction based on variable precision rough set model. The concepts of β lower distribution reduct and β upper distribution reduct based on variable precision rough sets (VPRS) are first introduced. Their equivalent definitions are then given, and the relationships among β lower and β upper distribution reducts and alternative types of knowledge reduction in inconsistent systems are investigated. It is proved that for some special thresholds, β lower distribution reduct is equivalent to the maximum distribution reduct, whereas β upper distribution reduct is equivalent to the possible reduct. The judgement theorems and discernibility matrices associated with the β lower and β upper distribution reducts are also established, from which we can obtain the approaches to knowledge reduction in VPRS.

Keywords: Rough set; Consistent set; Inconsistent system; Knowledge reduction

* Corresponding author. Present address: College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, Hebei 050016, PR China.

E-mail addresses: mijsh@263.net (J.-S. Mi), wuwz8681@sina.com (W.-Z. Wu), wxzhang@xjtu.edu.cn (W.-X. Zhang).

0020-0255/$ - see front matter
doi:10.1016/j.ins.2003.07.004

1. Introduction

Rough set theory (RST), proposed by Pawlak [11], is an extension of set theory for the study of intelligent systems characterized by inexact, uncertain or vague information. Using the concepts of lower and upper approximations in RST, the knowledge hidden in the systems may be discovered and expressed in the form of decision rules. One of the central problems of RST is classification analysis. The whole approach is inspired by the notion of inadequacy of available information to perform complete classification of objects belonging to a specified category. Quite frequently, the available information allows only for partial classification. RST can be used to this kind of classification but the classification must be completely correct or certain. In practice, however, it seems that admitting some level of uncertainty in the classification process may lead to a deeper understanding and a better utilization of properties of the data being analyzed. In [20], partial classification was taken into account by Ziarko. He extended RST by introducing a probability value β. The β value represents a bound on the conditional probability of a proportion of objects in a condition class which are classified to the same decision class. This approach operated in the context of the variable precision rough sets (VPRS) model in literature [4]. In VPRS, for any value of β and decision class, we may identify the condition classes which have the property that the largest group proportion of objects classified to the decision class is at least β, in which case each of the condition classes is classified to the decision class.

One fundamental aspect of RST involves the searching for some particular subsets of condition attributes. By such one subset the information for classification purposes provided is the same as the condition attribute set done. Such subsets are called reducts. To acquire brief decision rules from inconsistent systems, knowledge reduction is needed.

Knowledge reduction is performed in information systems by means of the notion of a reduct based on a specialization of the general notion of independence due to Marczewski [8]. In recent years, more attention has been paid to knowledge reduction in inconsistent systems in rough set research. Many types of knowledge reduction have been proposed in the area of rough sets [2,10,12,13,17,19]. Possible rules and possible reducts have been proposed as a means to deal with inconsistence in an inconsistent decision table [3,7]. Approximation rules [18] are also used as an alternative to possible rules. On the other hand, generalized decision rules and generalized decision reducts [7] provide a decision maker with more flexible selection of decision behavior. In [5], the notions of α-reduct and α-relative reduct for decision tables are defined. The α-reduct allows occurrence of additional inconsistency that is controlled by means of α parameter. In [16], Slezak presented a new concept of attribute reduction that keeps the class membership distribution unchanging for all objects in the information system. It was shown by Slezak [15] that the knowledge reduction preserving the membership distribution is equivalent to the knowledge reduction preserving the value of generalized inference measure function. A generalized knowledge reduction was also introduced in [15] that allows the value of generalized inference measure function after the attribute reduction to be different from the original one by user-specified threshold. The notion of dynamic reducts was described by Bazan [1]. Dynamic reducts are just subsets of all reducts derived both from the original decision table and from the majority of randomly chosen decision sub-tables. Although the number of dynamic reducts can be much smaller than the number of all reducts, by using fewer dynamic rules induced by dynamic reducts, as supposed to using all decision rules, however will not lower the classification capabilities. The predictive capability of generalized dynamic reducts [1] can be even better than that of dynamic reducts. Kryszkiewicz [6] investigated and compared five notions of knowledge reduction in inconsistent systems. In fact, only two of them, possible reduct and μ-decision reduct (we refer to it as distribution reduct), are essential because the others are just equivalent to one of them respectively. We can see that each of the reducts aimed at some basic requirement.

In [2,20], β-reduct based on VPRS was introduced. This type of reduct preserves the sum of objects in β lower approximations of all decision classes. But the derived decision rules from the β-reduct may be in conflict with the ones from the original system (see Example 3.1). To overcome this kind of drawback, we introduce new concepts of β lower distribution reduct and β upper distribution reduct.

In this paper, we are concerned with approaches to knowledge reduction based on variable precision rough sets model. In the next section, we give some basic notions related to information systems and rough sets. Section 3 is devoted to introducing some new types of knowledge reduction based on VPRS, namely, the β lower distribution reduct, β upper distribution reduct, and β upper possible reduct. After giving their equivalent definitions, we investigate the relationships among alternative types of knowledge reduction in inconsistent systems. It is proved that for some special thresholds, β lower distribution reduct is equivalent to the maximum distribution reduct, and whereas β upper distribution reduct is equivalent to the possible reduct. In Section 4, the judgement theorems and discernibility matrixes [14] with respect to β lower and upper distribution reducts are examined. These results provide approaches to knowledge reduction based on variable precision rough sets model, which are significant both in the theoretic and applied perspectives. A computative example is also given to illustrate our approaches. We then conclude the paper with a summary and outlook for further research in Section 5.

2. Basic notions related to rough sets

The notion of information system (IS) (sometimes called data tables, attribute-value systems, knowledge representation systems, etc.) provides a convenient tool for description of objects in terms of their attribute values. Rough sets have been introduced as a tool to deal with inexact, uncertain or vague knowledge in many branches of artificial intelligence.

An information system is a pair (U, A), where U is a non-empty, finite set of objects called the universe and A is a non-empty, finite set of attributes, such that $a : U \to V_a$ for any $a \in A$, where V_a is called the domain of a.

Each non-empty subset $B \subseteq A$ determines an indiscernibility relation as follows:

$$R_B = \{(x, y) \in U \times U : a(x) = a(y)\ \forall a \in B\}.$$

R_B partitions U into a family of disjoint subsets U/R_B called a quotient set of U:

$$U/R_B = \{[x]_B : x \in U\},$$

where $[x]_B$ denotes the equivalence class determined by x with respect to (w.r.t.) B, i.e.,

$$[x]_B = \{y \in U : (x, y) \in R_B\}.$$

Let $X \subseteq U$, $B \subseteq A$, one can characterize X by a pair of lower and upper approximations:

$$\underline{R_B}(X) = \{x \in U : [x]_B \subseteq X\} = \bigcup\{[x]_B : [x]_B \subseteq X\},$$

$$\overline{R_B}(X) = \{x \in U : [x]_B \cap X \neq \emptyset\} = \bigcup\{[x]_B : [x]_B \cap X \neq \emptyset\}.$$

The lower approximation $\underline{R_B}(X)$ is the set of objects that belong to X with certainty, while the upper approximation $\overline{R_B}(X)$ is the set of objects that possibly belong to X. The pair $(\underline{R_B}(X), \overline{R_B}(X))$ is referred to as the Pawlak rough set of X w.r.t. B.

For $\beta \in (0.5, 1]$, we denote

$$\underline{R_B}^{\beta}(X) = \{x \in U : P(X/[x]_B) \geqslant \beta\} = \bigcup\{[x]_B : P(X/[x]_B) \geqslant \beta\},$$

$$\overline{R_B}^{\beta}(X) = \{x \in U : P(X/[x]_B) > 1 - \beta\} = \bigcup\{[x]_B : P(X/[x]_B) > 1 - \beta\}.$$

$\underline{R_B}^{\beta}(X)$ and $\overline{R_B}^{\beta}(X)$ are called β lower approximation and β upper approximation respectively. Where $P(X/Y) = \frac{|X \cap Y|}{|Y|}$ if $|Y| > 0$, and $P(X/Y) = 1$ otherwise. $|X|$ is the cardinality of the set X.

A decision table (DT) is an information system $(U, A \cup D)$, where $A \cap D = \emptyset$. A is called the condition attribute set, while D is called the decision attribute set. If $R_A \subseteq R_D$, then we say that $(U, A \cup D)$ is consistent, otherwise it is inconsistent.

For the sake of simplicity, in the sequel, we set $D = \{d\}$, $V_d = \{1, 2, \ldots, r\}$, and $U/D = \{D_1, D_2, \ldots, D_r\}$. D_j is the decision class $D_j = \{x \in U, d(x) = j\}$.

Based on a variable precision rough set model, each $x \in \underline{R_B}^{\beta}(D_j)$ may generate a decision rule as follows:

$$\bigwedge_{c\in B}(c, c(x)) \to d = j.$$

It is easy to prove that $\underline{R_B}^{\beta}$ and $\overline{R_B}^{\beta}$ satisfy the following properties:

(1) $\underline{R_B}^{\beta}(X) = \sim \overline{R_B}^{\beta}(\sim X)\ \forall X \subseteq U$;
(2) $\underline{R_B}^{1}(X) = \underline{R_B}(X)$, $\overline{R_B}^{1}(X) = \overline{R_B}(X)\ \forall X \subseteq U$;
(3) $\underline{R_B}^{\beta}(D_i) \cap \underline{R_B}^{\beta}(D_j) = \emptyset$, $i \neq j$;
(4) $\bigcup_{j=1}^{r} \underline{R_B}^{\beta}(D_j) \subseteq \bigcup_{j=1}^{r} \overline{R_B}^{\beta}(D_j)$.
(5) The formula $\overline{R_B}^{\beta}(D_i) \cap \overline{R_B}^{\beta}(D_j) = \emptyset$ does not hold in general.

Where $\sim A$ is the complement of A. From (2), it follows that VPRS model is an extension of Pawlak rough set model.

The property (4) implies that not every object supports a decision rule in the form $\bigwedge_{c\in B}(c, c(x)) \to d = j$.

3. Notions of knowledge reduction based on VPRS

In this section, we introduce some new notions of knowledge reduction based on variable precision rough set model and show the relationships among them.

Definition 3.1. Let $(U, A \cup D)$ be a DT, $B \subseteq A$. We denote

$$\sigma_B^{\beta} = \frac{\sum\{|\underline{R_B}^{\beta}(D_j)| : j \leqslant r\}}{|U|}, \quad \lambda_B^{\beta} = \frac{\sum\{|\overline{R_B}^{\beta}(D_j)| : j \leqslant r\}}{|U|},$$

$$L_B^{\beta} = (\underline{R_B}^{\beta}(D_1), \ldots, \underline{R_B}^{\beta}(D_r)), \quad H_B^{\beta} = (\overline{R_B}^{\beta}(D_1), \ldots, \overline{R_B}^{\beta}(D_r)).$$

(1) If $\sigma_B^{\beta} = \sigma_A^{\beta}$, we say that B is a β lower approximate consistent set of $(U, A \cup D)$. If B is a β lower approximate consistent set, and no proper subset of B is β lower approximate consistent, then B is referred to as a β lower approximate reduct of $(U, A \cup D)$ (in [2,20] it is called β reduct or approximate reduct).
(2) If $\lambda_B^{\beta} = \lambda_A^{\beta}$, we say that B is a β upper approximate consistent set of $(U, A \cup D)$. If B is a β upper approximate consistent set, and no proper subset of B is β upper approximate consistent, then B is referred to as a β upper approximate reduct of $(U, A \cup D)$.
(3) If $L_B^{\beta} = L_A^{\beta}$, we say that B is a β lower distribute consistent set of $(U, A \cup D)$. If B is a β lower distribute consistent set, and no proper subset of B is β lower distribute consistent, then B is referred to as a β lower distribute reduct of $(U, A \cup D)$.

(4) If $H_B^\beta = H_A^\beta$, we say that B is a β upper distribute consistent set of $(U, A \cup D)$. If B is a β upper distribute consistent set, and no proper subset of B is β upper distribute consistent, then B is referred to as a β upper distribute reduct of $(U, A \cup D)$.

A β upper (lower) distribution consistent set is a subset of attribute set that preserves the β upper (lower) approximations of all decision classes. The decision rules derived from the β distribution consistent set are compatible with the ones derived from A, that is to say, if two decision rules derived respectively from the reduced and the original system are supported by a same object, then their decision parts must be the same. A β lower approximate consistent set preserves the sum of objects in β lower approximations of all decision classes. However, the derived decision rules may be incompatible (in conflict) with the ones from A.

Example 3.1. In [2], an inconsistent decision table is given as Table 1.

For $\beta = 0.55$, $\{c_3, c_6\}$ is a β (lower approximate) reduct, from which the derived decision rules are:

r_1: $(c_3, 1) \wedge (c_6, 1) \rightarrow (d, F)$, supported by o_1, o_2, o_4, o_5, o_7.

r_2: $(c_3, 1) \wedge (c_6, 0) \rightarrow (d, M)$, supported by o_3.

r_3: $(c_3, 0) \wedge (c_6, 0) \rightarrow (d, F)$, supported by o_6.

It is easy to see that r_1 is in conflict with the following decision rule derived from the original system:

$$(c_1, 1) \wedge (c_2, 1) \wedge (c_3, 1) \wedge (c_4, 1) \wedge (c_5, 1) \wedge (c_6, 1) \rightarrow (d, M).$$

Theorem 3.1. *Let $(U, A \cup D)$ be a DT, then a β lower distribution consistent set must be a β lower approximate consistent set, a β upper distribution consistent set must be a β upper approximate consistent set.*

Proof. It follows directly from the definitions. □

Table 1
A decision table

U	c_1	c_2	c_3	c_4	c_5	c_6	d
o_1	1	1	1	1	1	1	M
o_2	1	0	1	0	1	1	M
o_3	0	0	1	1	0	0	M
o_4	1	1	1	0	0	1	F
o_5	1	0	1	0	1	1	F
o_6	0	0	0	1	1	0	F
o_7	1	0	1	0	1	1	F

The following example illustrates that the converse of Theorem 3.1 is not always true, that is, a β lower approximate consistent set may not be a β lower distribution consistent set.

Example 3.2. Let $(U, A \cup \{d\})$ be an inconsistent DT given in Table 2.

Denote by $A = \{a_1, a_2\}$, $B = \{a_1\}$. The decision classes are

$$D_1 = \{x_1, x_2, x_4, x_9, x_{10}\}, \quad D_2 = \{x_3, x_{11}\}, \quad D_3 = \{x_5, x_7\}, \quad D_4 = \{x_6, x_8\}.$$

It can be calculated that

$$\begin{array}{ll} \underline{R_A}^{0.7}(D_1) = \{x_1, x_4, x_9, x_{10}, x_{11}\}, & \underline{R_A}^{0.7}(D_2) = \emptyset, \\ \underline{R_A}^{0.7}(D_3) = \{x_5\}, & \underline{R_A}^{0.7}(D_4) = \{x_8\}, \\ \underline{R_B}^{0.7}(D_1) = \{x_1, x_2, x_3, x_4, x_9, x_{10}, x_{11}\}, & \underline{R_B}^{0.7}(D_2) = \emptyset, \\ \underline{R_B}^{0.7}(D_3) = \emptyset, & \underline{R_B}^{0.7}(D_4) = \emptyset. \end{array}$$

Therefore $B = \{a_1\}$ is a 0.7 lower approximate consistent set but not a 0.7 lower distribution consistent set.

Similarly, a β upper approximate consistent set may not be a β upper distribution consistent set.

Definition 3.2. Let $(U, A \cup D)$ be a DT, $B \subseteq A$. For $x \in U$, we denote

$$\mu_B(x) = (P(D_1/[x]_B), \ldots, P(D_r/[x]_B)),$$
$$\gamma_B(x) = \{D_{j_i} : P(D_{j_i}/[x]_B) = \max_{j \leqslant r}\{P(D_j/[x]_B)\}\},$$
$$\delta_B(x) = \{D_j : [x]_B \cap D_j \neq \emptyset\}.$$

(1) If $\mu_B(x) = \mu_A(x)$ for all $x \in U$, we say that B is a distribution consistent set of $(U, A \cup D)$. If B is a distribution consistent set, and no proper subset of B

Table 2
An inconsistent DT

U	a_1	a_2	d
x_1	1	0	1
x_2	1	1	1
x_3	1	1	2
x_4	1	2	1
x_5	2	0	3
x_6	2	2	4
x_7	2	2	3
x_8	2	1	4
x_9	1	0	1
x_{10}	1	0	1
x_{11}	1	0	2

is distribution consistent, then B is referred to as a distribution reduct of $(U, A \cup D)$ (in [1] it is called μ reduct).

(2) If $\gamma_B(x) = \gamma_A(x)$ for all $x \in U$, we say that B is a maximum distribution consistent set of $(U, A \cup D)$. If B is a maximum distribution consistent set, and no proper subset of B is maximum distribution consistent, then B is referred to as a maximum distribution reduct of $(U, A \cup D)$.

(3) If $\delta_B(x) = \delta_A(x)$ for all $x \in U$, we say that B is a possible consistent set of $(U, A \cup D)$. If B is a possible consistent set, and no proper subset of B is possible consistent, then B is referred to as a possible reduct of $(U, A \cup D)$.

A distribution consistent set is a subset of attribute set that preserves the degree in which every object belongs to each decision class. A maximum distribution consistent set preserves all maximum decision classes. But the degree of confidence of each decision rule derived from the reduced system may not be equal to the one derived from the original system supported by the same object. A possible consistent set preserves all possible decision classes.

Theorem 3.2. *Let $(U, A \cup D)$ be a DT, then for arbitrary $\beta \in (0.5, 1]$, a distribution consistent set must be both a β lower and a β upper distribution consistent set.*

Proof. Let B be a distribution consistent set, then $\forall x \in U$, we have $\mu_B(x) = \mu_A(x)$, i.e., $P(D_j/[x]_B) = P(D_j/[x]_A)$ $\forall j \leqslant r$. Hence $x \in \underline{R_B}^{\beta}(D_j) \Longleftrightarrow x \in \underline{R_A}^{\beta}(D_j)$. Consequently $\underline{R_B}^{\beta}(D_j) = \underline{R_A}^{\beta}(D_j)$ $\forall j \leqslant r$. That is to say, B is a β lower distribution consistent set.

Similarly, B is a β upper distribution consistent set. $\square$

Theorem 3.3. *Let $(U, A \cup D)$ be a DT, $B \subseteq A$. We denote*

$$\beta_B = \min\{\max_{j \leqslant r} P(D_j/[x]_B) : x \in U\}, \quad \beta_0 = \min\{\beta_A, \beta_B\}.$$

If $\beta_0 > 0.5$, then

(1) *for $\beta \in (0.5, \beta_0]$, if B is a β lower distribution consistent set, then B is a maximum distribution consistent set;*

(2) *for $\beta \in (0.5, \beta_A]$, if B is a maximum distribution consistent set, then B is a β lower distribution consistent set;*

(3) *for $\beta \in (0.5, \beta_0]$, B is a maximum distribution reduct iff B is a β lower distribution reduct.*

Proof. (1) Let B be a β lower distribution consistent set, $\beta \in (0.5, \beta_0]$, then $\underline{R_B}^{\beta}(D_j) = \underline{R_A}^{\beta}(D_j)$ $\forall j \leqslant r$. For any $x \in U$, we know from $\beta_0 > 0.5$ that both $\gamma_A(x)$ and $\gamma_B(x)$ are singletons. As a result we have

$$\begin{aligned} D_j \in \gamma_A(x) &\Rightarrow P(D_j/[x]_A) \geqslant \beta_0 \Rightarrow P(D_j/[x]_A) \geqslant \beta \\ &\Rightarrow x \in \underline{R_A}^{\beta}(D_j) \Rightarrow x \in \underline{R_B}^{\beta}(D_j) \\ &\Rightarrow P(D_j/[x]_B) \geqslant \beta \Rightarrow D_j \in \gamma_B(x), \\ D_j \in \gamma_B(x) &\Rightarrow P(D_j/[x]_B) \geqslant \beta_0 \Rightarrow P(D_j/[x]_B) \geqslant \beta \\ &\Rightarrow x \in \underline{R_B}^{\beta}(D_j) \Rightarrow x \in \underline{R_A}^{\beta}(D_j) \\ &\Rightarrow P(D_j/[x]_A) \geqslant \beta \Rightarrow D_j \in \gamma_A(x). \end{aligned}$$

Therefore $\gamma_B(x) = \gamma_A(x)\ \forall x \in U$, which follows that B is a maximum distribution consistent set.

(2) Let B be a maximum distribution consistent set, then for any $x \in U$, we have $\gamma_B(x) = \gamma_A(x)$. Since $B \subseteq A$, it can be easily verified that $\mathscr{I}([x]_B) = \{[y]_A : [y]_A \subseteq [x]_B\}$ forms a partition of $[x]_B$.

For any $j \leqslant r$,

$$x \in \underline{R_A}^{\beta}(D_j) \Rightarrow P(D_j/[x]_A) \geqslant \beta > 0.5 \Rightarrow \gamma_A(x) = \{D_j\} \Rightarrow \gamma_B(x) = \{D_j\}.$$

If $[y]_A \in \mathscr{I}([x]_B)$, then $\gamma_B(x) = \gamma_B(y)$. Since $\gamma_A(y) = \gamma_B(y)$, we have $\gamma_A(y) = \gamma_B(x) = \{D_j\}$, which induces $P(D_j/[y]_A) \geqslant \beta_A > \beta$. Thus we have

$$\begin{aligned} P(D_j/[x]_B) &= \left(\sum\{|[y]_A \cap D_j| : [y]_A \in \mathscr{I}([x]_B)\}\right) \Big/ |[x]_B| \\ &= \sum\left\{P(D_j/[y]_A)\cdot\frac{|[y]_A|}{|[x]_B|} : [y]_A \in \mathscr{I}([x]_B)\right\} \\ &\geqslant \beta\cdot\sum\left\{\frac{|[y]_A|}{|[x]_B|} : [y]_A \in \mathscr{I}([x]_B)\right\} = \beta. \end{aligned}$$

Consequently $x \in \underline{R_B}^{\beta}(D_j)$, and in turn, $\underline{R_A}^{\beta}(D_j) \subseteq \underline{R_B}^{\beta}(D_j)$.

Conversely, since

$$x \in \underline{R_B}^{\beta}(D_j) \Rightarrow P(D_j/[x]_B) \geqslant \beta > 0.5 \Rightarrow \gamma_B(x) = \{D_j\} \Rightarrow \gamma_A(x) = \{D_j\},$$

we have $P(D_j/[x]_A) = \max_{i \leqslant r} P(D_i/[x]_A) \geqslant \beta_A > \beta$, as a result $x \in \underline{R_A}^{\beta}(D_j)$, and hence $\underline{R_B}^{\beta}(D_j) \subseteq \underline{R_A}^{\beta}(D_j)$.

Thus we have proved that $\underline{R_B}^{\beta}(D_j) = \underline{R_A}^{\beta}(D_j)\ \forall j \leqslant r$, i.e., B is a β lower distribution consistent set.

(3) For $\beta \in (0.5, \beta_0]$, we can conclude from (1) and (2) that B is a maximum distribution consistent set iff B is a β lower distribution consistent set. Therefore B is a maximum distribution reduct iff B is a β lower distribution reduct. $\square$

In general, if β does not satisfy the condition in Theorem 3.3, then a β lower distribution consistent set may not be a maximum distribution consistent set, likewise, a maximum distribution consistent set may not be a β lower distribution consistent set.

Table 3
A decision table

U	a_1	a_2	d
x_1	1	1	1
x_2	1	1	1
x_3	1	1	1
x_4	1	1	2
x_5	0	1	2
x_6	0	1	2
x_7	0	1	1
x_8	1	0	1
x_9	1	0	1
x_{10}	1	0	2

Example 3.3. Given an inconsistent DT (Table 3).

We can verify that $\{a_1\}$ is a maximum distribution consistent set but not a 0.7 lower distribution consistent set. Also, $\{a_2\}$ is a 0.85 lower distribution consistent set but not a maximum distribution consistent set.

Theorem 3.4. *Let $(U, A \cup D)$ be a DT, $B \subseteq A$. We denote*

$$\alpha_B = \min\{P(D_j/[x]_B) : x \in U, D_j \cap [x]_B \neq \emptyset\}, \quad \alpha_0 = \min\{\alpha_A, \alpha_B\}.$$

Then

(1) *for $\beta \in (1 - \alpha_0, 1]$, if B is a possible consistent set, then B is a β upper distribution consistent set;*
(2) *for $\beta \in (1 - \alpha_B, 1]$, if B is a β upper distribution consistent set, then B is a possible consistent set;*
(3) *for $\beta \in (1 - \alpha_0, 1]$, B is a possible reduct iff B is a β upper distribution reduct.*

Proof. (1) Let B be a possible consistent set, then $\delta_B(x) = \delta_A(x)\ \forall x \in U$. Therefore, for any $j \leqslant r$,

$$\begin{aligned} x \in \overline{R_B}^{\beta}(D_j) &\Rightarrow P(D_j/[x]_B) > 1 - \beta \Rightarrow D_j \in \delta_B(x) \\ &\Rightarrow D_j \in \delta_A(x) \Rightarrow D_j \cap [x]_A \neq \emptyset \Rightarrow P(D_j/[x]_A) \geqslant \alpha_A \geqslant \alpha_0 > 1 - \beta \\ &\Rightarrow x \in \overline{R_A}^{\beta}(D_j). \end{aligned}$$

Conversely,

$$\begin{aligned} x \in \overline{R_A}^{\beta}(D_j) &\Rightarrow P(D_j/[x]_A) > 1 - \beta \Rightarrow D_j \in \delta_A(x) \Rightarrow D_j \in \delta_B(x) \\ &\Rightarrow D_j \cap [x]_B \neq \emptyset \Rightarrow P(D_j/[x]_B) \geqslant \alpha_B \geqslant \alpha_0 > 1 - \beta \\ &\Rightarrow x \in \overline{R_B}^{\beta}(D_j). \end{aligned}$$

Thus we can conclude that $\overline{R_B}^{\beta}(D_j) = \overline{R_A}^{\beta}(D_j)\ \forall j \leqslant r$, i.e., B is a β upper distribution consistent set.

(2) Let B be a β upper distribution consistent set, $\beta \in (1-\alpha_B, 1]$, then $\overline{R_B}^{\beta}(D_j) = \overline{R_A}^{\beta}(D_j)\ \forall j \leqslant r$.

Hence for any $x \in U$,

$$\begin{aligned} D_j \in \delta_B(x) &\Rightarrow P(D_j/[x]_B) \geqslant \alpha_B > 1-\beta \Rightarrow x \in \overline{R_B}^{\beta}(D_j) \\ &\Rightarrow x \in \overline{R_A}^{\beta}(D_j) \Rightarrow P(D_j/[x]_A) > 1-\beta \Rightarrow D_j \in \delta_A(x), \end{aligned}$$

which implies $\delta_B(x) \subseteq \delta_A(x)$. On the other hand, it is easy to see that $\delta_A(x) \subseteq \delta_B(x)$. Hence $\delta_B(x) = \delta_A(x)\ \forall x \in U$, i.e., B is a possible consistent set.

(3) For $\beta \in (1-\alpha_0, 1]$, we can see from (1) and (2) that B is a possible consistent set iff B is a β upper distribution consistent set. Therefore B is a possible reduct iff B is a β upper distribution reduct. □

In general, if β does not satisfy the condition in Theorem 3.4, then a β upper distribution consistent set may not be a possible consistent set, and a possible consistent set may not be a β upper distribution consistent set.

Example 3.4. Given a inconsistent DT (Table 4):

It can be easily verified that $\{a_1\}$ is a possible consistent set. Since $M^{0.55}_{\{a_1\}}(x_1) = \{D_2\}$, but $M^{0.55}_A(x_1) = \{D_1, D_2\}$, $\{a_1\}$ is not a 0.55 upper distribution consistent set.

We can observe that $\delta_{\{a_2\}}(x_1) = \{D_1, D_2\}$, but $\delta_A(x_1) = \{D_2\}$, so that $\{a_2\}$ is not a possible consistent set.

It can be calculated that

$$\begin{aligned} &\mu_A(x_1) = \mu_A(x_2) = (1/2, 1/2), \\ &\mu_A(x_3) = \mu_A(x_4) = \mu_A(x_5) = (1/3, 2/3), \\ &\mu_A(x_6) = (0, 1). \end{aligned}$$

Thus we have

$$\begin{aligned} &M^{0.6}_A(x_1) = M^{0.6}_A(x_2) = \{D_1, D_2\}, \\ &M^{0.6}_A(x_3) = M^{0.6}_A(x_4) = M^{0.6}_A(x_5) = M^{0.6}_A(x_6) = \{D_2\}. \end{aligned}$$

Table 4
A decision table

U	a_1	a_2	d
x_1	1	1	1
x_2	1	1	2
x_3	1	2	2
x_4	1	2	1
x_5	1	2	2
x_6	2	2	2

But

$$\mu_{\{a_2\}}(x_1) = \mu_{\{a_2\}}(x_2) = (1/2, 1/2),$$
$$\mu_{\{a_2\}}(x_3) = \mu_{\{a_2\}}(x_4) = \mu_{\{a_2\}}(x_5) = \mu_{\{a_2\}}(x_6) = (1/4, 3/4),$$

consequently,

$$M^{0.6}_{\{a_2\}}(x_1) = M^{0.6}_{\{a_2\}}(x_2) = \{D_1, D_2\},$$
$$M^{0.6}_{\{a_2\}}(x_3) = M^{0.6}_{\{a_2\}}(x_4) = M^{0.6}_{\{a_2\}}(x_5) = M^{0.6}_{\{a_2\}}(x_6) = \{D_2\}.$$

Thus we have concluded that $\{a_2\}$ is a 0.6 upper distribution consistent set.

Using the labels in Theorem 3.4, it can be easily calculated that $\alpha_A = 1/3$, $\alpha_{\{a_1\}} = 2/5$, and $\alpha_0 = 1/3$. Therefore, for any $\beta \in (2/3, 1]$, we conclude that $\{a_1\}$ is a β upper distribution consistent set, and of course, $\{a_1\}$ is a β upper distribution reduct.

Theorem 3.5. *Let $(U, A \cup D)$ be a DT, then* 1 *upper distribution consistent set is* 1 *lower distribution consistent set.*

Proof. Let B be an 1 upper distribution consistent set, then $\overline{R_B}^1(D_j) = \overline{R_A}^1(D_j)$ $\forall j \leqslant r$. It follows that, for any $x \in U$,

$$P(D_j/[x]_B) > 0 \iff P(D_j/[x]_A) > 0,$$

that is to say,

$$[x]_B \cap D_j \neq \emptyset \iff [x]_A \cap D_j \neq \emptyset.$$

Since $\{D_j : j \leqslant r\}$ forms a partition of U, we conclude that $[x]_B \subseteq D_j \iff [x]_A \subseteq D_j$, which implies $\underline{R_B}^1(D_j) = \underline{R_A}^1(D_j)$. Thus we have shown that B is an 1 lower distribution consistent set. $\square$

4. Approaches to knowledge reduction based on VPRS

This section provides approaches to β upper distribution reduct and β lower distribution reduct based on variable precision rough set model. Let us first give some equivalence descriptions for each consistent set.

Theorem 4.1. *Let $(U, A \cup D)$ be a DT, $B \subseteq A$. We denote*

$$M^{\beta}_B(x) = \{D_j : x \in \overline{R_B}^{\beta}(D_j)\}, \quad x \in U,$$
$$G^{\beta}_B(x) = \{D_j : x \in \underline{R_B}^{\beta}(D_j)\}, \quad x \in U.$$

Then

(1) *B is a β upper distribution consistent set iff* $M_B^\beta(x) = M_A^\beta(y)\ \forall x \in U$.
(2) *B is a β lower distribution consistent set iff* $G_B^\beta(x) = G_A^\beta(y)\ \forall x \in U$.

Proof. (1) Combining the facts that $x \in \overline{R_A}^\beta(D_j) \Longleftrightarrow D_j \in M_A^\beta(x)$, and $x \in \overline{R_B}^\beta(D_j) \Longleftrightarrow D_j \in M_B^\beta(x)$, we can easily conclude (1).

(2) It is similar to the proof of (1). □

Theorem 4.2 (Judgement theorem of knowledge reduction). *Let* $(U, A \cup D)$ *be a DT*, $B \subseteq A$, *then*

(1) *B is a β upper distribution consistent set iff if* $x, y \in U$ *satisfy* $M_A^\beta(x) \neq M_A^\beta(y)$, *then* $[x]_B \cap [y]_B = \emptyset$.
(2) *B is a β lower distribution consistent set iff if* $x, y \in U$ *satisfy* $G_A^\beta(x) \neq G_A^\beta(y)$, *then* $[x]_B \cap [y]_B = \emptyset$.

Proof. Since $B \subseteq A$, it is easy to verify that $\mathscr{J}([x]_B) = \{[y]_A : [y]_A \subseteq [x]_B\}$ forms a partition of $[x]_B$.

(1) "⇒" If $[x]_B \cap [y]_B \neq \emptyset$, $x, y \in U$, then $[x]_B = [y]_B$, therefore $M_B^\beta(x) = M_B^\beta(y)$. Since B is a β upper distribution consistent set, we have $M_B^\beta(x) = M_A^\beta(x)$, $M_B^\beta(y) = M_A^\beta(y)$. Thus $M_A^\beta(x) = M_A^\beta(y)$.

"⇐" For any $x \in U$, if $[y]_A \subseteq [x]_B$, then $[x]_B \cap [y]_B \neq \emptyset$. By the assumption we obtain $M_A^\beta(x) = M_A^\beta(y)$.

For any $j \leqslant r$, if $x \in \overline{R_B}^\beta(D_j)$, then by the definition we have $[x]_B \subseteq \overline{R_B}^\beta(D_j)$. Since $[x]_B = \bigcup\{[y]_A : [y]_A \in \mathscr{J}([x]_B)\}$, we obtain that $[y_0]_A \subseteq \overline{R_B}^\beta(D_j)$ for all $[y_0]_A \in \mathscr{J}([x]_B)$, that is, $D_j \in M_A^\beta(y_0)$. Therefore $D_j \in M_A^\beta(x)$, and in tune, $x \in \overline{R_A}^\beta(D_j)$.

On the other hand, if $x \in \overline{R_A}^\beta(D_j)$, then $D_j \in M_A^\beta(x)$. Hence for all $[y]_A \in \mathscr{J}([x]_B)$, we have $D_j \in M_A^\beta(y)$, that is to say, $P(D_j/[y]_A) > 1 - \beta$. Therefore we have that

$$\begin{aligned} P(D_j/[x]_B) &= \left(\sum\{|[y]_A \cap D_j| : [y]_A \in \mathscr{J}([x]_B)\}\right)\Big/|[x]_B| \\ &= \sum\left\{P(D_j/[y]_A) \cdot \frac{|[y]_A|}{|[x]_B|} : [y]_A \in \mathscr{J}([x]_B)\right\} \\ &> (1-\beta)\sum\left\{\frac{|[y]_A|}{|[x]_B|} : [y]_A \in \mathscr{J}([x]_B)\right\} = 1 - \beta. \end{aligned}$$

As a result $x \in \overline{R_B}^\beta(D_j)$.

Thus we conclude that $\overline{R_B}^\beta(D_j) = \overline{R_A}^\beta(D_j)$ for all $j \leqslant r$, i.e., B is a β upper distribution consistent set.

(2) It is similar to the proof of (1). □

Theorem 4.2 provides approaches to judge whether a subset of attributes is β lower (β upper respectively) distribute consistent or not. In [9,21], Ziarko and Shan presented a method for computing reducts. Now we provide approaches to knowledge reduction based on variable precision rough set model. To this end, we first give the following notions.

Definition 4.1. Let $(U, A \cup D)$ be a DT, $U/R_A = \{C_1, \dots, C_m\}$. We denote

$$D_1^{*\beta} = \{([x]_A, [y]_A) : \quad M_A^\beta(x) \neq M_A^\beta(y)\},$$
$$D_2^{*\beta} = \{([x]_A, [y]_A) : \quad G_A^\beta(x) \neq G_A^\beta(y)\}.$$

Denoted by $a_k(C_i)$ the value of a_k w.r.t. the objects in C_i. Define

$$D_l^\beta(C_i, C_j) = \begin{cases} \{a_k \in A : a_k(C_i) \neq a_k(C_j)\}, & (C_i, C_j) \in D_l^{*\beta}, \\ A, & (C_i, C_j) \notin D_l^{*\beta}, \end{cases} \quad (l = 1, 2).$$

Then $D_l^\beta(C_i, C_j)$, $(l = 1, 2)$ are referred to as β upper distribution and β lower distribution discernibility attribute sets respectively. And $D_l^\beta = (D_l^\beta(C_i, C_j),\ i, j \leqslant m)$, $(l = 1, 2)$ are referred to as β upper distribution and β lower distribution discernibility matrices respectively.

Theorem 4.3. *Discernibility matrices* $D_l^\beta = (D_l^\beta(C_i, C_j), i, j \leqslant m)$, $(l = 1, 2)$ *satisfy the following properties*:

(1) *They are all symmetric matrices, i.e.*, $D_l^\beta(C_i, C_j) = D_l^\beta(C_j, C_i)\ \forall i, j \leqslant m$;
(2) *Elements in the main diagonals are all* A, *i.e.*, $D_l^\beta(C_i, C_i) = A\ \forall i \leqslant m$;
(3) $D_l^\beta(C_i, C_j) \subseteq D_l^\beta(C_i, C_s) \cup D_l(C_s, C_j)\ \forall i, s, j \leqslant m$.

Proof. It need only to prove (3). Since

$$\begin{aligned} a_k \notin D_l^\beta(C_i, C_s) \cup D_l^\beta(C_s, C_j) &\Rightarrow a_k \notin D_l^\beta(C_i, C_s) \quad \text{and} \quad a_k \notin D_l^\beta(C_s, C_j) \\ &\Rightarrow a_k(C_i) = a_k(C_s) \quad \text{and} \quad a_k(C_s) = a_k(C_j) \\ &\Rightarrow a_k(C_i) = a_k(C_j) \\ &\Rightarrow a_k \notin D_l^\beta(C_i, C_j), \end{aligned}$$

we conclude that (3) holds. $\square$

Theorem 4.4. *Let* $(U, A \cup D)$ *be a DT*, $B \subseteq A$, *then*

(1) B *is a* β *upper distribution consistent set iff* $B \cap D_1^\beta(C_i, C_j) \neq \emptyset$ *for all* $(C_i, C_j) \in D_1^{*\beta}$.
(2) B *is a* β *lower distribution consistent set iff* $B \cap D_2^\beta(C_i, C_j) \neq \emptyset$ *for all* $(C_i, C_j) \in D_2^{*\beta}$.

Proof. (1) Suppose B is a β upper distribution consistent set. For any $(C_i, C_j) \in D_1^{*\beta}$, we can find $x, y \in U$ such that $C_i = [x]_A$, $C_j = [y]_A$, then $M_A^\beta(x) \neq M_A^\beta(y)$. We obtain from Theorem 4.2 that $[x]_B \cap [y]_B \neq \emptyset$. Then there exists $a_k \in B$ such that $a_k(x) \neq a_k(y)$, i.e., $a_k(C_i) \neq a_k(C_j)$. Hence $a_k \in D_1^\beta(C_i, C_j)$, and in turn, $B \cap D_1^\beta(C_i, C_j) \neq \emptyset$.

Conversely, if there exists $(C_i, C_j) \in D_l^{*\beta}$ such that $B \cap D_1^\beta(C_i, C_j) = \emptyset$, we can select $x, y \in U$ satisfying $C_i = [x]_A$, $C_j = [y]_A$, then $M_A^\beta(x) \neq M_A^\beta(y)$. For any $a_k \in B$, since $a_k \notin D_1^\beta(C_i, C_j)$, we have $a_k(C_i) = a_k(C_j)$. Consequently $a_k(x) = a_k(y)$, which implies $[x]_B = [y]_B$. By Theorem 4.2 we conclude that B is not a β upper distribution consistent set. Thus we complete the proof of (1).

The proof of (2) is similar to the proof of (1). □

Definition 4.2. Let $(U, A \cup D)$ be a DT, $D_l^\beta = (D_l^\beta(C_i, C_j), i, j \leqslant m)$, $(l = 1, 2)$, the β upper distribution and the β lower distribution discernibility matrices respectively. Denoted by

$$\begin{aligned} M_l^\beta &= \bigwedge \left\{ \bigvee \{a_k : a_k \in D_l^\beta(C_i, C_j)\} : i, j \leqslant m \right\} \\ &= \bigwedge \left\{ \bigvee \{a_k : a_k \in D_l^\beta(C_i, C_j)\} : (C_i, C_j) \in D_l^{*\beta} \right\} \quad (l = 1, 2). \end{aligned}$$

Then M_l^β, $(l = 1, 2)$, are referred to the β upper distribution and the β lower distribution discernibility functions respectively.

Theorem 4.5. *Let $(U, A \cup D)$ be a DT. The minimal disjunctive normal form of each discernibility function M_l^β $(l = 1, 2)$, is*

$$M_l^\beta = \bigvee_{k=1}^{t} \left(\bigwedge_{s=1}^{q_k} a_{i_s} \right) \quad (l = 1, 2).$$

Denoted by $B_{lk} = \{a_{i_s} : s = 1, \ldots, q_k\}$, then $\{B_{lk} : k = 1, \ldots, t\}$, $(l = 1, 2)$, are just the set of all β upper and β lower distribution reducts respectively.

Proof. It follows directly from Theorem 4.4 and the definition of minimal disjunctive normal forms of the discernibility functions. □

Theorem 4.5 provides practical approaches to the above two kinds of knowledge reduction based on variable precision rough set model. Now we present an example to illustrate this approach.

Example 4.1. Given an inconsistent DT (Table 5):

The condition classes of objects are

$$C_1 = \{x_1\}, \quad C_2 = \{x_2\}, \quad C_3 = \{x_3, x_5, x_6\}, \quad C_4 = \{x_4\}.$$

Table 5
An inconsistent DT

U	a_1	a_2	a_3	a_4	d
x_1	1	0	0	0	1
x_2	0	1	1	1	2
x_3	0	1	0	0	2
x_4	0	1	1	0	2
x_5	0	1	0	0	1
x_6	0	1	0	0	1

The decision classes of objects are

$$D_1 = \{x_1, x_5, x_6\}, \quad D_2 = \{x_2, x_3, x_4\}.$$

It can be easily calculated that

$$\mu_A(x_1) = (1, 0), \quad \mu_A(x_2) = (0, 1), \quad \mu_A(x_4) = (0, 1),$$
$$\mu_A(x_3) = \mu_A(x_5) = \mu_A(x_6) = (2/3, 1/3).$$

Hence

$$M_A^{0.7}(x_1) = \{D_1\}, \quad M_A^{0.7}(x_2) = \{D_2\}, \quad M_A^{0.7}(x_4) = \{D_2\},$$
$$M_A^{0.7}(x_3) = M_A^{0.7}(x_5) = M_A^{0.7}(x_6) = \{D_1, D_2\}.$$

Thus we have

$$D_1^{*0.7} = \{(C_1, C_2), (C_1, C_3), (C_1, C_4), (C_2, C_3), (C_3, C_4)\}.$$

Since

$$D_1^{0.7}(C_1, C_2) = \{a_1, a_2, a_3, a_4\}, \quad D_1^{0.7}(C_1, C_3) = \{a_1, a_2\}, \quad D_1^{0.7}(C_1, C_4) = \{a_1, a_2, a_3\}, \quad D_1^{0.7}(C_2, C_3) = \{a_3, a_4\}, \quad D_1^{0.7}(C_3, C_4) = \{a_3\},$$

then

$$M_1^{0.7} = (a_1 \vee a_2 \vee a_3 \vee a_4) \wedge (a_1 \vee a_2) \wedge (a_1 \vee a_2 \vee a_3) \wedge (a_3 \vee a_4) \wedge a_3$$
$$= (a_1 \wedge a_3) \vee (a_2 \wedge a_3).$$

By Theorem 4.5 we conclude that both $\{a_1, a_3\}$ and $\{a_2, a_3\}$ are the 0.7 upper distribution reducts of the DT.

It can be calculated that

$$G_A^{0.6}(x_1) = \{D_1\}, \quad G_A^{0.6}(x_2) = \{D_2\},$$
$$G_A^{0.6}(x_3) = G_A^{0.6}(x_5) = G_A^{0.6}(x_6) = \{D_1\},$$
$$G_A^{0.6}(x_4) = \{D_2\}.$$

Then we have

$$D_2^{*0.6} = \{(C_1, C_2), (C_1, C_4), (C_2, C_3), (C_3, C_4)\}.$$

Therefore

$$M_2^{0.6} = (a_1 \vee a_2 \vee a_3 \vee a_4) \wedge (a_1 \vee a_2 \vee a_3) \wedge (a_3 \vee a_4) \wedge a_3 = a_3.$$

Thus we derive that $\{a_3\}$ is the unique 0.6 lower distribution reduct.

5. Conclusion

Many types of attribute reduction in inconsistent decision tables have been proposed based on RST. In this paper, we have studied knowledge reduction in variable precision rough set model. Two new types of attribute reduction based on VPRS, β lower distribution reduct and β upper distribution reduct, are introduced. They preserve all decision classes in some available classification level, and eliminate the drawback of β reduct which may change decision results of some objects. It has been proved that for certain parameter β, the β lower distribution reduct is equivalent to the maximum distribution reduct, whereas the β upper distribution reduct is equivalent to the possible reduct. The relationships among different kinds of knowledge reduction in inconsistent systems are discussed as well. We have obtained new approaches to knowledge reduction based on variable precision rough set model by providing the discernibility matrices.

Though the information systems we discussed here are inconsistent, they are all complete. Since incomplete information systems are more complicated than complete information systems, further research of knowledge reduction for different requirements in incomplete information systems is needed. In further research, we will develop the proposed approaches to more generalized and more complicated information systems such as incomplete information systems and fuzzy information systems.

Acknowledgements

The authors would like to thank the anonymous referees for their very constructive comments. This work is supported by the Nature Science Foundation of China (10271039) and the National 973 Program of China (2002CB312200).

References

[1] J. Bazan, A comparision of dynamic and non-dynamic rough set methods for extracting laws from decison tables, in: L. Polkowski, A. Skowron (Eds.), Rough Sets in Knowledge Discovery, Phisica-Verlag, Heidelberg, 1998, pp. 321–365.

[2] M. Beynon, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research 134 (2001) 592–605.

[3] J. Grzymala-Busse, X. Zuo, Classification strategies using certain and possible rules in: LNAI 1424, in: RSCTC'98, Springer, 1998, pp. 37–44.

[4] J.D. Katzberg, W. Ziarko, Variable precision rough sets with asymmetric bounds, in: W. Ziarko (Ed.), Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer, 1994, pp. 167–177.

[5] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: tutorial, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hybridization, A New Trend in Decision Making, Springer, 1999, pp. 3–98.

[6] M. Kryszkiewicz, Comparative study of alternative type of knowledge reduction in inconsistent systems, International Journal of Intelligent Systems 16 (2001) 105–120.

[7] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1998) 39–49.

[8] E. Marczewski, A general scheme of independence in mathematics, Bulletin de L Academie Polonaise des Sciences—Serie des Sciences Mathematiques Astronomiques et Physiques 6 (1958) 731–736.

[9] N. Shan, W. Ziarko, Data-based acquisition and incremental modification of classification rules, Computational Intelligence: An International Journal 11 (2) (1995) 357–370.

[10] H.S. Nguyen, D. Slezak, Approximation reducts and association rules correspondence and complexity results, in: N. Zhong, A. Skowron, S. Oshuga (Eds.), Proceedings of RSFDGrC'99, Yamaguchi, Japan, LNAI 1711, 1999, pp. 137–145.

[11] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (5) (1982) 341–356.

[12] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Boston, 1991.

[13] M. Quafatou, α-RST: a generalization of rough set theory, Information Sciences 124 (2000) 301–316.

[14] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R. Slowinski (Ed.), Intelligent Decision Support—Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, 1992, pp. 331–362.

[15] D. Slezak, Approximate reducts in decision tables, in: Proceedings of IPMU'96, Granada, Spain, vol. 3, 1996, pp. 1159–1164.

[16] D. Slezak, Searching for dynamic reducts in inconsistent decision tables, in: Proceedings of IPMU'98, Paris, France, vol. 2, 1998, pp. 1362–1369.

[17] R. Slowinski, C. Zopounidis, A.I. Dimitras, Prediction of company acquisition in Greece by means of the rough set approach, European Journal of Operational Research 100 (1997) 1–15.

[18] J. Stefanowski, On rough set based approaches to induction of decision rules, in: J. Polkowski, A. Skowron (Eds.), Rough Sets in Knowledge Discovery, vol. 1, Physica-Verlag, Heidelberg, 1998, pp. 500–529.

[19] W.-X. Zhang, W.-Z. Wu, J.-Y. Liang, D.-Y. Li, Theory and method of rough sets, Science Press, Beijing, 2001.

[20] W. Ziarko, Variable precision rough set model, Journal of Computer and System Sciences 46 (1) (1993) 39–59.

[21] W. Ziarko, N. Shan, A method for computing all maximally general rules in attribute-value systems, Computational Intelligence: an International Journal 12 (2) (1996) 223–234.