

Research Article

Parameterized Local Reduction of Decision Systems

Degang Chen, Yanyan Yang, and Xiao Zhang

Department of Mathematics and Physics, North China Electric Power University, Beijing 102206, China

Correspondence should be addressed to Degang Chen, chengdegang@263.net

Received 25 May 2012; Revised 17 September 2012; Accepted 3 October 2012

Academic Editor: Juan Manuel Peña

Copyright © 2012 Degang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One important and valuable topic in rough sets is attribute reduction of a decision system. The existing attribute reductions are designed to just keep confidence of every certain rule as they cannot identify key conditional attributes explicitly for special decision rules. In this paper, we develop the concept of θ -local reduction in order to offer a minimal description for special θ -possible decision rules. The approach of discernibility matrix is employed to investigate the structure of a θ -local reduction and compute all θ -local reductions. An example of medical diagnosis is employed to illustrate our idea of the θ -local reduction. Finally, numerical experiments are performed to show that our method proposed in this paper is feasible and valid.

1. Introduction

The concept of rough sets was originally proposed by Pawlak [1] as a mathematical approach to handle imprecision, vagueness, and uncertainty in data analysis. This theory has been demonstrated to have its usefulness and versatility in successfully solving a variety of problems [1]. The main application of rough set theory is attribute reduction in databases. Given a decision system with conditional and decision attributes, attribute reduction aims to find a subset of the original conditional attributes that contain the same information as the original one. The concept of attribute reduction can be viewed as the strongest and the most important result in rough set theory to distinguish itself from other theories.

Along the line in [1], many research works have been concentrated on computing attribute reduction and developing other types of attribute reduction under the framework of rough sets [1–23]. For example, Skowron and Rauszer [12] employed the approach of discernibility matrix to set up mathematical foundation for finding reducts. Wang [15–17] characterized attribute reduction by information entropy. Possible rules and possible reduct of all decision classes were proposed to deal with inconsistency in an inconsistent

decision table [4, 6]. In [5], in order to provide an underlying classification of knowledge reductions, five notions of knowledge reduction (possible reduct, approximation reduct, generalized decision reduct, μ -decision reduct, and μ -reduct) were investigated and compared in inconsistent systems. In fact, only two of them, possible reduct preserving upper approximations and μ -decision reduct preserving membership to all decision classes are essential because others are just equivalent to one of them, respectively. The notion of dynamic reducts was described in [2] as subsets of all reducts derived from both the original decision table and the majority of the randomly chosen decision subtables. In [10], α -reduct and α -relative reduct were proposed to allow occurrence of additional inconsistency that is controlled by means of the parameter α ($\alpha \in [0, 1]$). In [19] notions of the distribution reduct and maximum distribution reduction were proposed, and relationships among the maximum distribution reduct, the distribution reduct, and the possible reduct were discussed. In [3, 22], β -reduct was introduced to preserve the sum of objects in β -lower approximations of all decision classes ($\beta \in (0.5, 1.0]$) based on variable precision rough sets (VPRS). However, Zhou et al. [24] pointed out that the dependency function may not be monotonic when computing β -reduct and decision rules derived by the β -reduct may be in conflict with those derived from the original system. To overcome this drawback, in [9] β -lower and β -upper distribution reducts were proposed to preserve β -lower approximations and β -upper approximations of all decision classes, respectively. 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.

These attribute reductions share the following two arguments. First they are developed in terms of all decision classes and cannot explicitly identify key conditional attributes for particular decision classes, so these reductions can be viewed as global reductions. However, in many practical problems people always pay more attention to some special decision classes rather than other ones, and condition attributes and decision rules with closed connection to these special decision classes always draw much attention. For example, in decision-making of medical diagnosis, key condition attributes related to the disease always draw much attention than other ones, and it is clearly meaningful to identify such key attributes. Second, as it is well known, certain and possible rules can be extracted from a decision system, and confidence of every certain rule is 1 while confidence of every possible rule is less than 1. But most of the existing attribute reductions only offer minimal conditional attributes to keep confidence of every certain rule invariant, and possible rules with bigger confidence are ignored. However, in most practical problems, possible rules with bigger confidence are always available and applied to decision making, so it is clearly meaningful to identify key conditional attributes for possible rules with bigger confidence.

To improve these two arguments in the meantime, the definition of θ -local reduction is presented in this paper. First we give the concept of θ -reduction ($\theta \in (0.5, 1.0]$) to keep the confidence of those possible rules, and then we further consider θ -local reduction to offer a minimal description and extract possible decision rules with bigger confidence for special decision classes. Approach of discernibility matrix is employed to characterize the structure of θ -local reduction. It is proven that the core of θ -reduction can be expressed as the union of the cores of θ -local reductions, and the discernibility matrix of θ -reduction can be obtained by composing discernibility matrices of θ -local reductions. An example of medical diagnosis is employed to illustrate our idea of θ -local reduction, and we also perform several experiments to demonstrate the effectiveness of the idea in this paper.

The rest of this paper is structured as follows. In the next section we give some basic notions related to rough sets. In Section 3 we define θ -local reduction, and approach of discernibility matrix is employed to find θ -local reduction. In Section 4, we perform numerical experiments to demonstrate that our method proposed in this paper is feasible to process massive data. We then conclude the paper in Section 5.

2. Basic Notions Related to Rough Sets

An information system is a pair (U, A) , where $U = \{x_1, x_2, \dots, x_n\}$ is a nonempty, finite set called the universe of discourse, and $A = \{a_1, a_2, \dots, a_m\}$ is a nonempty, finite set of attributes. With every subset of attributes $B \subseteq A$, we associate a binary relation $\text{IND}(B)$, called a B -indiscernibility relation, and defined as $\text{IND}(B) = \{(x, y) \in U \times U : a(x) = a(y) \text{ for all } a \in B\}$, then $\text{IND}(B)$ is an equivalence relation and $\text{IND}(B) = \cap_{a \in B} \text{IND}(\{a\})$. By $[x]_B$, we denote the equivalence class of $\text{IND}(B)$, including x . For $X \subseteq U$, sets $\{x \in X : [x]_B \subseteq X\}$ and $\{x \in X : [x]_B \cap X \neq \emptyset\}$ are called B -lower and B -upper approximations of X in A and denoted as $\underline{B}X$ and $\overline{B}X$, respectively. If $\underline{B}X = \overline{B}X$, we say X is definable, otherwise it is indefinable.

A decision table (DT) (sometimes called a decision system) is an information system $(U, A \cup D)$, where $A \cap D = \emptyset$, $U/\text{IND}(D) = \{D_1, D_2, \dots, D_r\}$. A is a set of conditional attributes, while D is the decision attribute.

Suppose $A = \{a_1, a_2, \dots, a_m\}$, then we have $[x]_A = \cap_{i=1}^m [x]_{\text{IND}(\{a_i\})}$. If $|D_l \cap [x]_A| \neq \emptyset$, then for $x \in D_l \cap [x]_A$ we can derive decision rule as $(a_1, a_1(x)) \wedge (a_2, a_2(x)) \wedge \dots \wedge (a_m, a_m(x)) \rightarrow (D, D(x))$. This rule can be denoted in terms of sets as $[x]_A \rightarrow D_l$ and its confidence is computed as $\text{conf}([x]_A \rightarrow D_l) = |D_l \cap [x]_A| / |[x]_A|$. Following we always call the decision rule of which confidence is not less than θ as θ -possible decision rule.

3. θ -Local Reduction of Decision Systems

In this section, we first introduce the definition of θ -reduction as a global one to consider every θ -possible decision rule, and we then develop θ -local reduction as improvement of θ -reduction to address special θ -possible decision rules. Approach of discernibility matrix is employed to find θ -local reduction, and differences between θ -local reduction and β -reduct are explained.

Let $A^* = (U, A \cup D)$ be a decision system, $U/\text{IND}(A) = \{M_1, \dots, M_L\}$, $U/\text{IND}(D) = \{D_1, \dots, D_r\}$, and $D_A^\theta(x) = \{D_l : \text{conf}([x]_A \rightarrow D_l) \geq \theta, l \in \{1, \dots, r\}\}$, $x \in U$, $\theta \in (0.5, 1.0]$. If for any $l \in \{1, \dots, r\}$, $\text{conf}([x]_A \rightarrow D_l) < \theta$, then $D_A^\theta(x) = \emptyset$. Obviously, $D_A^\theta(x)$ only includes a single element unless $D_A^\theta(x) = \emptyset$. For two θ -possible decision rules $[x]_A \rightarrow D_l$ and $[y]_A \rightarrow D_l$, if $B \subseteq A$ satisfying $[x]_B = [y]_B = [x]_A \cup [y]_A$, then clearly the rule $[x]_B \rightarrow D_l$ is also a θ -possible decision rule, which implies $D_B^\theta(x) = D_A^\theta(x)$, thus we have the definition of θ -reduction.

Definition 3.1. $B \subseteq A$ is a θ -reduction of DT if and only if B is a minimal set such that $D_B^\theta(x) = D_A^\theta(x)$ for for all $x \in U$.

Rule $[x]_B \rightarrow D_l$ can be seemed as the reduced rule of the rule $[x]_A \rightarrow D_l$. A θ -reduction B is a set of conditional attributes that keeps confidence of every reduced rule of a θ -possible decision rule still not less than θ , since for any object x_i satisfying

$\text{conf}([x_i]_A \rightarrow D_l) \geq \theta$, ($l \in \{1, \dots, r\}$), $\text{conf}([x_i]_B \rightarrow D_l) \geq \theta$ always holds. When $\theta = 1$, a θ -reduction is a classical attribute reduction which preserves lower approximations of all decision classes, namely, preserves confidences of all certain rules.

θ -reduction is defined by considering all θ -possible decision rules; thus, it can be considered as a global one for the whole system. In many practical problems people always pay more attention to special θ -possible decision rules related to special decision classes, so we improve the θ -reduction to the θ -local reduction as following to capture key condition attributes for special decision rules.

Let $D(U, A) = \{X \subseteq U : \overline{AX} = \underline{AX}\}$, that is, $D(U, A)$ is the family of all definable sets related to A . X is the set satisfying the following conditions: (1) X is a definable set, that is, $X \in D(U, A)$; (2) every $[x]_A \subseteq X$ can derive a decision rule $[x]_A \rightarrow D_l$ such that $\text{conf}([x]_A \rightarrow D_l) \geq \theta$. We denote the family of all these X satisfying the above two conditions as $Q_\theta(U, A \cup D)$. Clearly $Q_\theta(U, A \cup D)$ is a σ -algebra, and every element in $Q_\theta(U, A \cup D)$ is the union of several elements in $\{[x]_A : \text{conf}([x]_A \rightarrow D_l) \geq \theta\}$, $\exists l \in \{1, \dots, r\}$.

Definition 3.2. Suppose $A^* = (U, A \cup D)$ is a decision system, $\mathbf{X} = \{X_1, \dots, X_N\}$ such that $\mathbf{X} \subseteq Q_\theta(U, A \cup D)$. If $\mathbf{X} \subseteq Q_\theta(U, (A - \{a\}) \cup D)$, $a \in A$, then a is θ -dispensable in A for \mathbf{X} , otherwise a is θ -indispensable in A for \mathbf{X} . The collection of all the θ -indispensable elements in A is called the θ -local core of A^* for \mathbf{X} and denoted as $\text{Core}_\mathbf{X}^\theta(A \cup D)$. We say that $B \subseteq A$ is independent in A^* for \mathbf{X} if every attribute in B is θ -indispensable in B for \mathbf{X} . A set $B \subseteq A$ is called a θ -local reduction in A^* for \mathbf{X} if B is independent in A^* for \mathbf{X} and satisfying $\mathbf{X} \subseteq Q_\theta(U, B \cup D)$, that is, B is the minimal subset of A keeping $\mathbf{X} \subseteq Q_\theta(U, B \cup D)$.

If B is a θ -local reduction for \mathbf{X} , then for every $X_i \in \mathbf{X}$ and $x \in X_i$, $\text{conf}([x]_A \rightarrow D_l) \geq \theta$ implies $\text{conf}([x]_B \rightarrow D_l) \geq \theta$, that is, a θ -local reduction in A^* for \mathbf{X} keeps confidences of reduced rules of all θ -possible decision rules determined by elements in \mathbf{X} not less than θ . Furthermore, for every $x \in X_i$ we have $D_A^\theta(x) = D_B^\theta(x)$, and $\cup\{D_A^\theta(x) : x \in X_i \in \mathbf{X}\}$ is just a group of decision classes, thus a θ -local reduction in A^* for \mathbf{X} aims to select key condition attributes for this group of decision classes rather than for all the decision classes. Let $\mathbf{X} = \{X_A^1, X_A^2, \dots, X_A^r\}$ such that $X_A^l = \cup\{[x]_A : \text{conf}([x]_A \rightarrow D_l) \geq \theta\}$, $l = 1, \dots, r$. If B is a θ -local reduction for \mathbf{X} , then $\mathbf{X} \subseteq Q_\theta(U, B \cup D)$, that is, $X_A^l = X_B^l \in Q_\theta(U, B \cup D)$ for $l = 1, \dots, r$, thus $D_B^\theta(x) = D_A^\theta(x)$ for every $x \in U$ and B is a θ -reduction. This statement implies a θ -reduction is a special case of a θ -local reduction. The θ -reduction considers all θ -possible decision rules and decision classes, while θ -local reduction are developed in terms of special θ -possible decision rules and decision classes. Specially, for $x \in U$ satisfying $\text{conf}([x]_A \rightarrow D_l) \geq \theta$, if $\mathbf{X} = \{[x]_A\}$, then a θ -local reduction for \mathbf{X} only considers one decision class D_l .

Remark 3.3. In [3, 22] β -reduct was developed to keep β -dependency function. It seems to have closed connection to θ -reduction in this paper. However, they are two different concepts. First, β -reduct is proposed in the framework of VPRS, while θ -reduction are developed within the framework of classical rough set and does not need new rough set model. Second, β -reduct was introduced to preserve the sum of objects in β -lower approximations of all decision classes, and θ -reduction aims to keep θ -lower approximation of every decision class. Third, β -reduct cannot keep confidence of reduced rules of some β -possible rules not less than β as pointed in [24], but a θ -reduction can avoid this drawback by keeping confidence of every θ -possible decision rule not less than θ . At last, in VPRS possible rules with bigger confidence are due to noise, when noise is ignored, these rules are believed as

Table 1: A decision table.

U	c_1	c_2	c_3	c_4	c_5	c_6	d
x_1	1	1	1	1	1	1	M
x_2	1	0	1	0	1	1	M
x_3	0	0	1	1	0	0	M
x_4	1	1	1	0	0	1	F
x_5	1	0	1	0	1	1	F
x_6	0	0	0	1	1	0	F
x_7	1	0	1	0	1	1	F

certain ones. However, if these kinds of possible rules are not due to noise but roughness, risk will be ignored when they are applied to practical problems as certain ones. Thus, β -reduct does not have the formulism to distinguish noise and roughness. Since θ -reduction still considers all possible rules as possible ones, it can handle either noise or uncertainty at the meantime. Since a θ -reduction is a special case of a θ -local reduction, thus it is obvious that a β -reduct and a θ -local reduction are certainly different. Furthermore, a θ -local reduction is proposed to capture key attributes for special decision classes, and a β -reduct cannot do this work since it has to consider all decision classes at the meantime. Following we first give an example to indicate that β -reduct and the θ -local reduction are really different.

Example 3.4. An inconsistent decision table is given as Table 1.

Let $\beta = \theta = 0.6$, $\{\{c_4\}, \{c_1, c_3\}, \{c_3, c_6\}, \{c_2, c_5\}\}$ be the set of all 0.6-reducts, while $\{\{c_3, c_4\}, \{c_1, c_4, c_5\}, \{c_2, c_4, c_5\}, \{c_4, c_5, c_6\}, \{c_2, c_3, c_5\}, \{c_1, c_2, c_5\}, \{c_2, c_5, c_6\}\}$ is the set of all 0.6-local reduction for $X = \{X_1 = \{x_1, x_3\}, X_2 = \{x_4, x_5, x_6, x_7\}\}$. Obviously, every 0.6-reduct is not a 0.6-local reduction, and every 0.6-local reduction is not a 0.6-reduct.

Following we study the properties of the θ -local reduction. The set of all θ -local reductions in A^* for X is denoted by $\text{Red}_X^\theta(A \cup D)$, and we have the following theorem.

Theorem 3.5. $\text{Core}_X^\theta(A \cup D) = \cap \text{Red}_X^\theta(A \cup D)$.

Proof. (1) For any $a \in \text{Core}_X^\theta(A \cup D)$, $X \not\subseteq Q_\theta(U, (A - \{a\}) \cup D)$ holds. Suppose $a \notin \cap \text{Red}_X^\theta(A \cup D)$, then there exists a θ -local reduction B for X s.t. $B \subseteq A - \{a\}$, such that $X \subseteq Q_\theta(U, B \cup D) \subseteq Q_\theta(U, (A - \{a\}) \cup D) \subseteq Q_\theta(U, A \cup D)$, we get contradiction, hence $a \in \cap \text{Red}_X^\theta(A \cup D)$, namely $\text{Core}_X^\theta(A \cup D) \subseteq \cap \text{Red}_X^\theta(A \cup D)$.

(2) For any $a \in \cap \text{Red}_X^\theta(A \cup D)$, suppose $a \notin \text{Core}_X^\theta(A \cup D)$, then $X \subseteq Q_\theta(U, (A - \{a\}) \cup D)$, therefore there exists a θ -local reduction B for X s.t. $B \subseteq A - \{a\}$, then $a \notin B$, thus $a \notin \cap \text{Red}_X^\theta(A \cup D)$, we get contradiction, hence $a \in \text{Core}_X^\theta(A \cup D)$, namely, $\cap \text{Red}_X^\theta(A \cup D) \subseteq \text{Core}_X^\theta(A \cup D)$. From (1), (2), we can prove Theorem 3.5.

According to Theorem 3.5 θ -local core can be employed as the basis of finding all θ -local reductions for X since it is included in all θ -local reductions for X .

If elements in $X = \{X_1, \dots, X_N\}$ have nonempty overlaps, then there exists a X' satisfying elements in X' have empty overlaps and $\text{Red}_X^\theta(A \cup D) = \text{Red}_{X'}^\theta(A \cup D)$. We only prove this statement when $X = \{X_1, X_2\}$. \square

Theorem 3.6. Suppose $\mathbf{X} = \{X_1, X_2\}, X_1 \cap X_2 \neq \phi$, and $\mathbf{X}' = \{X_1 \cap X_2, X_1 - X_2, X_2 - X_1\}$, then $Red_{\mathbf{X}}^{\theta}(A \cup D) = Red_{\mathbf{X}'}^{\theta}(A \cup D)$.

Proof. For any $B \subseteq A, \mathbf{X} \subset Q_{\theta}(U, B \cup D) \Leftrightarrow \mathbf{X}' \subset Q_{\theta}(U, B \cup D)$, thus we finish the proof. \square

Following we always assume elements in $\mathbf{X} = \{X_1, \dots, X_N\}$ have empty overlaps. We have the following theorem for the θ -local core.

Theorem 3.7. $\mathbf{X} = \{X_1, \dots, X_N\}$ and $D_A^{\theta}(x) = \{D_l : l \leq r\}$ for any $x \in X_i (i = 1, \dots, N), X_i \cap X_j = \phi$, then we have $Core_{\mathbf{X}}^{\theta}(A \cup D) = \cup_{i=1}^N Core_{\{X_i\}}^{\theta}(A \cup D)$.

Proof. For any $a \in Core_{\mathbf{X}}^{\theta}(A \cup D) \Leftrightarrow$ there exists X_i satisfying $X_i \notin Q_{\theta}(U, (A - \{a\}) \cup D) \Leftrightarrow a \in Core_{\{X_i\}}^{\theta}(A \cup D) \Leftrightarrow a \in \cup_{i=1}^N Core_{\{X_i\}}^{\theta}(A \cup D)$.

When $\mathbf{X} = \{X_{A'}^1, \dots, X_A^r\}, X_A^l = \cup\{[x]_A : \text{conf}([x]_A \rightarrow D_l) \geq \theta\}, l = 1, \dots, r$, a θ -local reduction for \mathbf{X} is a θ -reduction. Thus, we get the core of θ -reduction can be expressed as the union of the cores of θ -local reductions for $\{X_A^l\} (l = 1, \dots, r)$. From Theorem 3.7 we can imply elements in the θ -local core for \mathbf{X} are indispensable for certain group of decision classes. If we pay more attention to a special group of decision classes, then the θ -local reduction may offer less conditional attributes only being indispensable for them. This is the objective of θ -local reductions. Following we study the computing of θ -local reductions. \square

Definition 3.8. Let $(U, A \cup D)$ be a DT, $U/IND(A) = \{M_1, \dots, M_L\}, \mathbf{X} = \{X_1, \dots, X_N\}$, and $\mathbf{X} \subseteq Q_{\theta}(U, A \cup D)$. Denoted by $a_k(M_i)$ as the value of samples in M_i in terms of a_k . Define

$$C_{ij} = \begin{cases} \{a_k \in A : a_k(M_i) \neq a_k(M_j)\}, (M_i \cup M_j) \not\subseteq X_h, & (M_i \cup M_j) \cap X_h \neq \phi, \exists h \in \{1, \dots, N\} \\ \phi, & \text{otherwise} \end{cases} \quad (3.1)$$

then $\mathbf{M}_{\mathbf{X}}^{\theta}(U, A \cup D) = (C_{ij})_{L \times L}$ is called the θ -local discernibility matrix for \mathbf{X} .

From the definition of θ -local discernibility matrix for \mathbf{X} we can easily get $\mathbf{M}_{\mathbf{X}}^{\theta}(U, A \cup D) = \cup_{i=1}^N \mathbf{M}_{\{X_i\}}^{\theta}(U, A \cup D)$, namely, $\mathbf{M}_{\mathbf{X}}^{\theta}(U, A \cup D)$ can be expressed as the union of $\mathbf{M}_{\{X_i\}}^{\theta}(U, A \cup D)$. If $\mathbf{X} = \{X_{A'}^1, \dots, X_A^r\}$, and $X_A^l = \cup\{[x]_A : \text{conf}([x]_A \rightarrow D_l) \geq \theta\}, l = 1, \dots, r$, then the discernibility matrix for \mathbf{X} can be obtained by composing discernibility matrices for $\{X_A^l\} (l = 1, \dots, r)$.

Theorem 3.9. $\mathbf{M}_{\mathbf{X}}^{\theta}(U, A \cup D) = (C_{ij})_{L \times L}$ satisfies the following properties.

- (1) It is a symmetric matrix, that is, $C_{ij} = C_{ji}$, for all $i, j \leq L$.
- (2) $C_{ij} = \phi$ if $M_i \cup M_j \subseteq X_h$ or $M_i \cup M_j \subseteq (U - X_h)$ holds, specially $C_{ii} = \phi$, for all $i, j \leq L$.

The proofs of following two theorems are straightforward.

Theorem 3.10. $Core_{\mathbf{X}}^{\theta}(A \cup D) = \{a \in A : C_{ij} = \{a\}, i, j \leq L\}$.

Theorem 3.11. $B \subseteq A$ includes a θ -local reduction for \mathbf{X} if and only if $B \cap C_{ij} \neq \phi$ for $C_{ij} \neq \phi$.

Table 2: The original decision table of patients' symptoms.

U	A				$\text{conf}([x_i]_A \rightarrow D_i)$
	Body Temperature a	Dry cough b	Headache c	Influenza d	
1	1	1	1	1	1
2	1	1	0	1	1
3	2	1	0	0	1
4	2	0	1	1	2/3
5	2	0	1	1	2/3
6	2	0	1	0	1/3
7	2	1	1	1	1
8	3	0	1	1	1/2
9	3	0	1	0	1/2
10	3	1	1	0	1

Note: in the table, body temperature a : 1 means high, 2 means slightly higher, and 3 means normal; for dry cough b , headache c , influenza d : 1 means yes and 0 means no.

By Theorem 3.10 the θ -local core is the set of single element of the θ -local discernibility matrix, thus, we can get $\text{Core}_X^\theta(A \cup D)$ from the θ -local discernibility matrix directly.

Definition 3.12. Let $(U, A \cup D)$ be a DT. A Boolean function is denoted by $f_X^\theta(U, A \cup D) = \bigwedge (\bigvee C_{ij})$, $C_{ij} \neq \emptyset$, then $f_X^\theta(U, A \cup D)$ is referred to the θ -local discernibility function for X .

Let $g_X^\theta(U, A \cup D)$ be the reduced disjunctive form of $f_X^\theta(U, A \cup D)$ by applying the distribution and absorption laws as many times as possible. Then there exist t and $A_i \subseteq A$ for $i = 1, \dots, t$ such that $g_X^\theta(U, A \cup D) = (\bigwedge A_1) \vee \dots \vee (\bigwedge A_t)$, thus, we have the following theorem.

Theorem 3.13. $\text{Red}_X^\theta(A \cup D) = \{A_1, \dots, A_t\}$.

The proof of Theorem 3.13 is similar to the one for traditional rough sets in [12].

Following we employ an example to illustrate the idea of θ -local reduction in this paper.

Example 3.14. When one suffers from a disease, certain symptoms can be observed. The doctor observes patients' symptoms and signs to implement diagnosis. In the following decision table (as Table 2 shown), ten patients' symptoms were observed and recorded. We would like to know which symptom is closely related to the influenza

$$\begin{aligned}
 U/\text{IND}(A) &= \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_5, x_6\}, \{x_7\}, \{x_8, x_9\}, \{x_{10}\}\}, \\
 U/\text{IND}(D) &= \{D_1 = \{x_1, x_2, x_4, x_5, x_7, x_8\}, D_2 = \{x_3, x_6, x_9, x_{10}\}\}.
 \end{aligned} \tag{3.2}$$

Let $\theta = 0.6$ and $\mathbf{X} = \{X_1 = \{x_1, x_2, x_4, x_5, x_6, x_7\}, X_2 = \{x_3, x_{10}\}\}$, then $D_A^{0.6}(x) = \{D_1\}$ for any $x \in X_1$ and $D_A^{0.6}(x) = \{D_2\}$ for any $x \in X_2$. Thus, θ -local discernibility matrices for $\{X_1\}$, $\{X_2\}$ and \mathbf{X} are as follows respectively:

$$\begin{pmatrix} \phi \\ \phi & \phi \\ \{a, c\} & \{a\} & \phi \\ \phi & \phi & \{b, c\} & \phi \\ \phi & \phi & \{c\} & \phi & \phi \\ \{a, b\} & \{a, b, c\} & \phi & \{a\} & \{a, b\} & \phi \\ \{a\} & \{a, c\} & \phi & \{a, b\} & \{a\} & \phi & \phi \end{pmatrix},$$

$$\begin{pmatrix} \phi \\ \phi & \phi \\ \{a, c\} & \{a\} & \phi \\ \phi & \phi & \{b, c\} & \phi \\ \phi & \phi & \{c\} & \phi & \phi \\ \phi & \phi & \{a, b, c\} & \phi & \phi & \phi \\ \{a\} & \{a, c\} & \phi & \{a, b\} & \{a\} & \{b\} & \phi \end{pmatrix}, \quad (3.3)$$

$$\begin{pmatrix} \phi \\ \phi & \phi \\ \{a, c\} & \{a\} & \phi \\ \phi & \phi & \{b, c\} & \phi \\ \phi & \phi & \{c\} & \phi & \phi \\ \{a, b\} & \{a, b, c\} & \{a, b, c\} & \{a\} & \{a, b\} & \phi \\ \{a\} & \{a, c\} & \phi & \{a, b\} & \{a\} & \{b\} & \phi \end{pmatrix}.$$

Clearly every C_{ij} in $\mathbf{M}_X^\theta(U, A \cup D)$ is the union of the corresponding elements in $\mathbf{M}_{\{X_1\}}^{0.6}(U, A \cup D)$ and $\mathbf{M}_{\{X_2\}}^{0.6}(U, A \cup D)$. Because of $\mathbf{X} = \{X_1, X_2\} = \{X_A^1, X_A^2\}$, $X_A^l = \cup\{[x]_A : \text{conf}([x]_A \rightarrow D_l) \geq 0.6, l = 1, 2\}$, a θ -local reduction for \mathbf{X} is a θ -reduction, in other words, the discernibility matrix of θ -reduction can be obtained by composing discernibility matrices of θ -local reductions. We can easily get the corresponding θ -local reduction for $\{X_1\}$ is $\text{Red}_{\{X_1\}}^{0.6}(A \cup D) = \{\{a, c\}\}$ and $\text{Core}_{\{X_1\}}^{0.6}(A \cup D) = \{a, c\}$. Similarly, $\text{Red}_{\{X_2\}}^{0.6}(A \cup D) = \{\{a, b, c\}\}$ $\text{Core}_{\{X_2\}}^{0.6}(A \cup D) = \{a, b, c\}$ and $\text{Red}_X^{0.6}(A \cup D) = \{\{a, b, c\}\}$, $\text{Core}_X^{0.6}(A \cup D) = \{a, b, c\}$.

If we pay much more attention to influenza than others, that is, we concentrate on the decision class with the value 1. Then the 0.6-local reduction $\{a, c\}$ can be employed by doctor to judge whether a patient is with influenza. However, $\{a, b, c\}$ is a 0.6-reduction and cannot explicitly identify key conditional attributes for particular decision class. On the other hand, a θ -local reduction keeps confidence of important θ -possible decision rules extracted from special decision classes not less than θ . For instance, we have the following 0.6-possible decision rules and their reduced ones related to D_1 .

The original 0.6-possible decision rules related to D_1 are as follows:

- (i) $(a, 1) \wedge (b, 1) \wedge (c, 1) \rightarrow (d, 1)$, $\text{conf}(i) = 1$.
- (ii) $(a, 1) \wedge (b, 1) \wedge (c, 0) \rightarrow (d, 1)$, $\text{conf}(ii) = 1$.

$$(iii) (a, 2) \wedge (b, 0) \wedge (c, 1) \rightarrow (d, 1), \text{conf}(iii) = 2/3.$$

$$(iv) (a, 2) \wedge (b, 1) \wedge (c, 1) \rightarrow (d, 1), \text{conf}(iv) = 1.$$

The reduced rules are as follows:

$$(\alpha_1) (a, 1) \wedge (c, 1) \rightarrow (d, 1), \text{conf}(\alpha_1) = 1.$$

$$(\alpha_2) (a, 1) \wedge (c, 0) \rightarrow (d, 1), \text{conf}(\alpha_2) = 1.$$

$$(\alpha_3) (a, 2) \wedge (c, 1) \rightarrow (d, 1), \text{conf}(\alpha_3) = 3/4.$$

From the above we know that a θ -local reduction keeps confidence of 0.6-possible decision rules for decision class D_1 . Thus, a θ -local reduction could explicitly identify key conditional attributes for particular decision classes and keeps confidence of θ -possible decision rules in terms of these decision classes not less than θ . Therefore, θ -local reduction can be selected as an effective method to deal with massive data.

4. Algorithm to Find One θ -Local Reduction and Numerical Experiments

In this section, we develop an algorithm to find a θ -local reduction. Then we perform numerical experiments for massive data sets to demonstrate that we can reduce the number of condition attributes and keep classification accuracies of raw data with θ -local reduction, which initially implies that the method proposed in this paper is feasible to process massive data.

4.1. Algorithm to Find One θ -Local Reduction

In the subsection, we develop an algorithm (Heuristic) to find one θ -local reduction by the approach of discernibility matrix proposed in Section 3.

Algorithm 4.1. To find one θ -local reduction for X_A^l of a certain decision class the following should be carried out:

Input: U, A, D, θ, X_A^l .

Output: One θ -local reduction $\text{Red}_{X_A^l}^\theta(A \cup D)$.

Initialize: $\text{Red}_{X_A^l}^\theta(A \cup D) = \emptyset$.

Step 1: Compute C_{ij} by Definition 3.8.

Step 2: Compute $\text{Core}_{X_A^l}^\theta(A \cup D) = \{a: C_{ij} = \{a\}\}$; and delete those C_{ij} with nonempty overlap with $\text{Core}_{X_A^l}^\theta(A \cup D)$.

Step 3: Let $\text{Red}_{X_A^l}^\theta(A \cup D) = \text{Core}_{X_A^l}^\theta(A \cup D)$.

Step 4: Add the element a whose frequency of occurrence is maximum in all C_{ij} into $\text{Red}_{X_A^l}^\theta(A \cup D)$; and delete those C_{ij} with nonempty overlap with $\text{Red}_{X_A^l}^\theta(A \cup D)$.

Step 5: If there still exist some $C_{ij} \neq \emptyset$, go to Step 4; otherwise, go to Step 6.

Table 3: Data sets description.

Datasets	Sample	Data type	Condition attributes	Decision classes
Breast Tissue	106	Real number	9	4
Credit Approval	653	Mix number	15	2
Ionosphere	351	Real number	34	2
Spect	267	Symbolic number	22	2
Wdbc	569	Real number	31	2
Wine	178	Real number	13	3

Step 6: If $\text{Red}_{X_A}^\theta(A \cup D)$ is not independent, delete the redundant elements in $\text{Red}_{X_A}^\theta(A \cup D)$.

Step 7: Output $\text{Red}_{X_A}^\theta(A \cup D)$.

The computational complexity of this algorithm is $O(|U|^2 \times |A|)$. Here $|U|$ is the size of universe, $|A|$ is the number of condition attributes.

4.2. Numerical Experiments

In this subsection, we perform experiments to demonstrate that with θ -local reduction and θ -reduction, condition attributes of a massive data set can be reduced with a satisfied parameter θ . We also employ support vector machine (SVM) as a classifier to compare the classification accuracies of reduced and raw data sets. The experiments are set up as follows.

4.2.1. Experimental Setup

Dataset

Six datasets from University of California, Irvine (UCI) Machine Learning Repository [25] are used (see Table 3).

Classifier

SVM in SVM-KM MATLAB Toolbox is employed as the classifier.

Dataset Split

In the process of classification, 10-fold cross-validation is applied on the six datasets.

Dataset Discretization

The fuzzy C-mean method proposed in [26] is used to discretize real valued condition attributes.

Table 4: The comparisons on selected attributes between θ -local reduction and θ -reduction.

	Class 1	Class 2	Class3	Class 4	Parameter
Breast tissue					0.6667
θ -local reduction	1, 6	1, 2, 3	4, 9	4, 6, 7, 9	
θ -reduction	1, 6, 2, 3, 7, 8, 9	1, 2, 3, 6, 7, 8, 9	4, 9, 2, 3, 6, 7	4, 6, 7, 9, 2, 3	
Wdbc					0.8333
θ -local reduction	1, 2, 4, 6, 7, 9, 11, 12, 14, 15, 22, 25, 27, 28, 29	1, 2, 7, 9, 10, 11, 12, 15, 19, 22, 23, 25, 27, 28, 29			
θ -reduction	1, 2, 4, 6, 7, 9, 11, 12, 14, 15, 22, 25, 27, 28, 29	1, 2, 7, 9, 10, 11, 12, 15, 19, 22, 23, 25, 27, 28, 29, 30			
Spect					0.8636
θ -local reduction	1, 2, 3, 4, 5, 9, 11, 12, 13, 14, 15, 17, 18, 20, 21, 22	1, 2, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22			
θ -reduction	1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 17, 18, 20, 21, 22	1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22			

Indices

They are (1) the number of selected attributes in the reduct, (2) classification accuracy of the reduct.

Parameter Specification

From confidences of all decision rules, we randomly choose a confidence which is greater than 0.5 as our experimental parameter on a specific dataset.

Accuracy

The accuracy in this paper is calculated by A/B , while A is the number of samples classified correctly in a certain decision class, and B is the number of samples in this decision class.

First Table 4 shows the detailed comparison of condition attributes in θ -local reduction and θ -reduction. Clearly θ -local reductions are different to θ -reductions. In most cases, θ -local reduction is subset of a certain θ -reduction. That is to say, the number of condition attributes in the θ -local reduction is often smaller than the one in the θ -reduction.

Next, Table 5 shows that though the average accuracy on θ -local reduction (i.e., 0.79466) is a little lower than average accuracy on raw data (i.e., 0.8147333), and much higher than average accuracy on θ -reduction (i.e., 0.75170666). This fact reveals that compared with θ -reduction and raw data, θ -local reduction can keep classification accuracy within a small perturbation. It is also easy to see from Table 5 that the number of attributes in the θ -local reduction (10.33333) is obviously less than the one in θ -reduction (12.066666) and far less than the one of raw data (20.66666). In particular for the dataset "BreastTissue," the number of attributes in the θ -local reduction is far less than the one in the θ -reduction and raw data set on every decision class.

Table 5: Comparison between θ -local reduction and θ -reduction.

	Accuracy		Raw data	Number of selected attributes		Raw data	Parameter
	θ -local-reduction	θ -reduction		θ -local-reduction	θ -reduction		
Breast tissue			0.6415			9	0.6667
Class 1	0.6132	0.6151		2	6		
Class 2	0.6208	0.6038		3	6		
Class 3	0.5717	0.6528		2	6		
Class 4	0.6472	0.6038		4	6		
Wine			0.9640			13	0.75
Class 1	0.9685	0.9719		8	9		
Class 2	0.9730	0.9652		9	10		
Class 3	0.9337	0.9663		8	9		
Credit approval			0.8071			15	0.8125
Class 1	0.8224	0.7994		14	15		
Class 2	0.8009	0.7995		13	14		
Ionosphere			0.8889			34	0.8750
Class 1	0.9090	0.9089		14	15		
Class 2	0.8917	0.9032		16	17		
Spect			0.6590			22	0.8636
Class 1	0.6932	0.6511		15	17		
Class 2	0.6170	0.6510		17	18		
Wdbc			0.9279			31	0.8333
Class 1	0.9297	0.9122		15	17		
Class 2	0.9279	0.9225		15	16		
Average	0.79466	0.75170666	0.8147333	10.33333	12.066666	20.66666	

These results initially imply that idea of θ -local reduction is effective to deal with some massive data. However, we select different parameter θ for different data set, and how to select a suitable parameter for certain data set is a complex problem. We omit detailed discussion on this topic in this paper.

5. Conclusion

Attribute reduction is a key topic in rough set theory. And the existing methods of attribute reduction ignore possible rules and cannot capture key condition attribute for special decision classes. In this paper, we develop the concept of θ -local reduction, by which possible rules with larger confidence are considered and key conditional attributes related to some special decision classes can be selected. Approach of discernibility matrix is employed to find θ -local reductions. Experiments are performed to demonstrate the effectiveness of the idea of θ -local reduction in this paper.

Acknowledgment

This paper is supported by a Grant of NSFC (71171080).

References

- [1] Z. Pawlak, "Rough sets," *International Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [2] J. Bazan, "A comparison of dynamic and non-dynamic rough set methods for extracting laws from decision tables," in *Rough Sets in Knowledge Discovery*, L. Polkowski and A. Skowron, Eds., pp. 321–365, Physica, Heidelberg, Germany, 1998.
- [3] M. Beynon, "Reducts within the variable precision rough sets model: a further investigation," *European Journal of Operational Research*, vol. 134, pp. 592–605, 2001.
- [4] J. Grzymala-Busse and X. Zuo, "Classification strategies using certain and possible rules," in *Proceedings of the 1st International Conference on Rough Sets and Current Trends in Computing (RSCTC '98)*, vol. 1424 of LNAI, pp. 37–44, Springer, 1998.
- [5] M. Kryszkiewicz, "Comparative study of alternative type of knowledge reduction in inconsistent systems," *International Journal of Intelligent System*, vol. 16, pp. 105–120, 2001.
- [6] M. Kryszkiewicz, "Rough set approach to incomplete information systems," *Information Sciences*, vol. 112, no. 1–4, pp. 39–49, 1998.
- [7] Y. Leung, J.-M. Ma, W.-X. Zhang, and T.-J. Li, "Dependence-space-based attribute reductions in inconsistent decision information systems," *International Journal of Approximate Reasoning*, vol. 49, no. 3, pp. 623–630, 2008.
- [8] Y. Leung, W.-Z. Wu, and W.-X. Zhang, "Knowledge acquisition in incomplete information systems: a rough set approach," *European Journal of Operational Research*, vol. 168, no. 1, pp. 164–180, 2005.
- [9] J.-S. Mi, W.-Z. Wu, and W.-X. Zhang, "Approaches to knowledge reduction based on variable precision rough set model," *Information Sciences*, vol. 159, no. 3–4, pp. 255–272, 2004.
- [10] H. S. Nguyen and D. Slezak, "Approximation reducts and association rules correspondence and complexity results," in *Proceedings of the 7th International Workshop on New Directions in Rough Sets, Data Mining, and Granular-Soft Computing (RSFDGrC '99)*, N. Zhong, A. Skowron, and S. Oshuga, Eds., vol. 1711 of LNAI, pp. 137–145, Yamaguchi, Japan, 1999.
- [11] Z. Pawlak and R. Sets, *Theoretical Aspects of Reasoning About Data*, Kluwer Academic Publishers, Boston, Mass, USA, 1991.
- [12] A. Skowron and C. Rauszer, "The discernibility matrices and functions in information systems," in *Intelligent Decision Support-Handbook of Applications and Advances of the Rough Sets Theory*, R. Slowinski, Ed., pp. 331–362, Kluwer Academic Publishers, 1992.
- [13] D. Slezak, "Approximate reducts in decision tables," in *Proceedings of 6th International Conference Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU '96)*, vol. 3, pp. 1159–1164, Granada, Spain, 1996.
- [14] D. Slezak, "Searching for dynamic reducts in inconsistent decision tables," in *Proceedings of 7th International Conference Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU '98)*, vol. 2, pp. 1362–1369, Paris, France, 1998.
- [15] G. Y. Wang, "Algebra view and information view of rough sets theory," in *Data Mining and Knowledge Discovery: Theory, Tools, and Technology III*, B. V. Dasarathy, Ed., vol. 4384 of *Proceedings of SPIE*, pp. 200–207, 2001.
- [16] G. Y. Wang, H. Yu, and D. C. Yang, "Decision table reduction based on conditional information entropy," *Chinese Journal of Computers*, vol. 25, no. 7, pp. 759–766, 2002.
- [17] G. Y. Wang, "Rough reduction in algebra view and information view," *International Journal of Intelligent System*, vol. 18, pp. 679–688, 2003.
- [18] W.-Z. Wu, M. Zhang, H.-Z. Li, and J.-S. Mi, "Knowledge reduction in random information systems via Dempster-Shafer theory of evidence," *Information Sciences*, vol. 174, no. 3–4, pp. 143–164, 2005.
- [19] W.-X. Zhang, J.-S. Mi, and W.-Z. Wu, "Approaches to knowledge reductions in inconsistent systems," *International Journal of Intelligent Systems*, vol. 18, pp. 989–1000, 2003.
- [20] W.-X. Zhang and G.-F. Qiu, *Uncertain Decision Making Based on Rough Sets*, vol. 6 of *Uncertainty Theory and Optimization Series*, Tsinghua University Press, Beijing, China, 2005.
- [21] W.-X. Zhang, W.-Z. Wu, J.-Y. Liang, and D.-Y. Li, *Theory and Method of Rough Sets*, Science Press, Beijing, China, 2001.
- [22] W. Ziarko, "Variable precision rough set model," *Journal of Computer and System Sciences*, vol. 46, no. 1, pp. 39–59, 1993.
- [23] W. Ziarko, "Analysis of uncertain information in the framework of variable precision rough sets," *Foundations of Computing and Decision Sciences*, vol. 18, pp. 381–396, 1993.

- [24] J. Zhou, J. Y. Wang, and A. Luo, "Analysis of characteristics reducts in variable precision rough sets," *Application Research of Computers*, vol. 24, no. 7, pp. 10–15, 2007.
- [25] UCI Machine Learning Repository, <http://archive.ics.uci.edu/ml/>.
- [26] D. Yu, Q. Hu, and W. Bao, "Combining rough set methodology and fuzzy clustering for knowledge discovery from quantitative data," *Proceeding of Chinese Society of Electrical Engineering*, vol. 24, no. 6, pp. 205–210, 2004.