Information Sciences 179 (2009) 4140–4150

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Relative reducts in consistent and inconsistent decision tables of the Pawlak rough set model

D.Q. Miao [a], Y. Zhao [b,*], Y.Y. Yao [b], H.X. Li [b,c], F.F. Xu [a,b]

[a] Department of Computer Science and Technology, Tongji University, Shanghai 201804, PR China
[b] Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
[c] School of Management and Engineering, Nanjing University, Nanjing, Jiangsu 210093, PR China

A R T I C L E I N F O

Article history:
Received 17 October 2008
Received in revised form 23 May 2009
Accepted 13 August 2009

Keywords:
Attribute reduction
Pawlak rough set model
Pawlak regions
Certainty of decision making
General decision
Relative relationship
Classification quality
Consistent and inconsistent decision tables

A B S T R A C T

A relative reduct can be considered as a minimum set of attributes that preserves a certain classification property. This paper investigates three different classification properties, and suggests three distinct definitions accordingly. In the Pawlak rough set model, while the three definitions yield the same set of relative reducts in consistent decision tables, they may result in different sets in inconsistent tables.

Relative reduct construction can be carried out based on a discernibility matrix. The study explicitly stresses a fact, that the definition of a discernibility matrix should be tied to a certain property. Regarding the three classification properties, we can define three distinct definitions accordingly.

Based on the common structure of the specific definitions of relative reducts and discernibility matrices, general definitions of relative reducts and discernibility matrices are suggested.

1. Introduction

Rough set theory is a useful tool for data analysis, dependency analysis and rule mining. It is typically assumed that we have a finite set of objects described by a finite set of attributes. The values of objects on attributes can be conveniently represented by an information table. Decision tables are one type of information tables with a decision attribute that gives the decision classes for all objects. A decision table is consistent if all object pairs that have the same condition values also have the same decision value; otherwise, it is inconsistent.

Attribute reduction is an important problem of rough set theory. The notion of a reduct plays an essential role in analyzing an information table. A reduct is a minimum subset of attributes that provides the same descriptive or classification ability as the entire set of attributes [12]. In other words, attributes in a reduct are jointly sufficient and individually necessary. Many methods have been proposed and examined for finding the set of all reducts or a single reduct [3,5,12,13,15,16,21,22,24,30,31]. One important class of reduct construction methods is based on the notion of a discernibility matrix [15]. One can construct a Boolean discernibility function from a discernibility relation. Skowron and Rauszer [15] show that the set of reducts are in fact the set of prime implicants of the reduced disjunctive form of the discernibility function. Many researchers studied reduct construction by using the discernibility information in a discernibility matrix [5,19,22,25,28,29,31].

* Corresponding author. Tel.: +1 306 585 5695; fax: +1 306 585 4745.
E-mail addresses: miaoduoqian@163.com (D.Q. Miao), yanzhao@cs.uregina.ca (Y. Zhao), yyao@cs.uregina.ca (Y.Y. Yao), huaxiongli@gmail.com (H.X. Li), xufeifei1983@hotmail.com (F.F. Xu).

0020-0255/$ - see front matter
doi:10.1016/j.ins.2009.08.020

This article emphasizes and unifies various types of relative reducts in the Pawlak rough set model. This is a fundamental issue that requires a careful investigation for the following three reasons. First, in the literature, there are many definitions for a reduct, and each focuses on preserving one specific property. For example, for classification tasks, we can consider the following three important properties: the certainty of decision making for all objects, the decision making for all objects, and object pair relationships regarding the decision attribute. Regarding the three properties, they are equivalent in consistent decision tables in the Pawlak rough set model, but are not necessarily equivalent in inconsistent tables. Misusing a reduct definition that is for consistent tables only for an inconsistent table will cause a problem.

Second, these existing reduct definitions are characterized by the same structure. However, there is still a lack of abstraction of a higher level. With the increasing requirements of data analysis, one needs to consider more properties of an information table. This naturally leads to more new definitions in different forms. Due to these reasons, a general definition of an attribute reduct is necessary and useful.

Third, there are confusions and misuses in the published papers on matrix-based approaches for reduct construction. Based on the three different properties and their corresponding reduct definitions, one needs to define distinct discernibility matrices accordingly. We examine the properties embedded in some of the existing definitions of discernibility matrices, such as the definitions provided by Hu and Cercone [5], Ye and Chen [29], Yang and Sun [25]. It should be emphasized that decoupling a certain property from its matching matrix will lead to a failure of finding the required reducts, or a result of finding some unwanted reducts. The paper aims to explain and clarify the reasons and results of the mismatch of properties and matrices. Based on the analysis, a general definition of discernibility matrices is suggested.

Many definitions of relative reducts and discernibility matrices presented in this paper are not new. Comparisons between some properties and reduct definitions have already been studied by many researchers [7,9,17,20,30]. The novelty of this study is to put all these definitions into a unified framework, and emphasize the coupling issues between properties and reduct definitions, also between properties and matrix definitions. Through the study of various properties, the relationships among reduct definitions and matrix definitions are more easily recognized. The significance of the generalized definitions is more easily identified. The study provides a complete view of relative reducts in consistent and inconsistent decision tables. This view can enhance the theoretical and logical understanding of the concept of relative reducts. Based on this understanding, efficient heuristics and algorithms can be designed and implemented.

The rest of the paper is organized as follows. Section 2 reviews the basic notations of the Pawlak rough set model. Section 3 discusses the definitions of relative reducts in both consistent and inconsistent decision tables. Section 4 provides a constructive view of relative reducts based on discernibility matrices. Section 5 is the conclusion. To improve the readability of the paper, the proofs are provided in Appendix.

2. Basic notions

Information tables are defined by Pawlak for raw data representation [12]. For classification tasks, we consider a special information table with a set of decision attributes. Such an information table also is called a *decision table*.

Definition 1. A decision table is the following tuple:

$$S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\}),$$

where U is a finite nonempty set of objects, and At is a finite nonempty set of attributes including a set $\mathbf{C}$ of condition attributes that describe the objects, and a set D of decision attributes that indicate the classes of objects. V_a is a nonempty set of values of $a \in At$, and $I_a : U \rightarrow V_a$ is an information function that maps an object in U to exactly one value in V_a.

For simplicity, we assume $D = \{d\}$ in this paper, where d is a decision attribute which labels the decision for each object. A table with multiple decision attributes can be easily transformed into a table with a single decision attribute by considering the Cartesian product of the original decision attributes.

Consistency and inconsistency are two important concepts for classification tasks.

Definition 2. A decision table is *consistent* if all object pairs that have the same condition values on $\mathbf{C}$ also have the same decision value on d, i.e.,

$$\forall (x,y) \in U \times U[\forall a \in \mathbf{C}[I_a(x) = I_a(y)] \Rightarrow I_d(x) = I_d(y)];$$

otherwise, it is *inconsistent* if there exists at least one object pair that have the same condition values on $\mathbf{C}$ but different decision values on d, i.e.,

$$\exists (x,y) \in U \times U[\forall a \in \mathbf{C}[I_a(x) = I_a(y)] \wedge I_d(x) \neq I_d(y)].$$

For a subset $A \subseteq At$ of attributes, Pawlak defines that two objects in U are A-indiscernible if and only if they have the same values on all attributes in A [12]. As a dual relation to A-indiscernibility, two objects are considered A-discernible if and only if they have different values on at least one attribute of A [33].

Definition 3. The indiscernibility and discernibility relations with respect to $A \subseteq At$ are defined as:

$$IND(A) = \{(x,y) \in U \times U | \forall a \in A[I_a(x) = I_a(y)]\};$$
$$DIS(A) = \{(x,y) \in U \times U | \exists a \in A[I_a(x) \neq I_a(y)]\}. \quad (1)$$

An indiscernibility relation is reflexive, symmetric and transitive, and thus is an equivalence relation. A discernibility relation is irreflexive, symmetric, but not transitive. The indiscernibility relation $IND(A)$ induces a partition of U, denoted by $U/IND(A)$ or π_A. The equivalence class of $U/IND(A)$ containing x is given by $[x]_{IND(A)} = [x]_A = \{y \in U | (x,y) \in IND(A)\}$, or $[x]$ if $IND(A)$ is understood.

For classification tasks, the relative indiscernibility relation and the dual relative discernibility relation are defined as follows.

Definition 4. The relative indiscernibility and discernibility relations defined by $A \subseteq \mathbf{C}$ with respect to D are defined as:

$$IND(A|D) = \{(x,y) \in U \times U | \forall a \in A[I_a(x) = I_a(y)] \vee I_d(x) = I_d(y)\};$$
$$DIS(A|D) = \{(x,y) \in U \times U | \exists a \in A[I_a(x) \neq I_a(y)] \wedge I_d(x) \neq I_d(y)\}. \tag{2}$$

A relative indiscernibility relation is reflexive, symmetric, but not transitive. A relative discernibility relation is irreflexive, symmetric, but not transitive. According to the definition of consistency and inconsistency, it is intuitive that given a decision table S, if there exists an object pair $(x,y) \in U \times U$ such that $(x,y) \in IND(\mathbf{C})$ and $(x,y) \notin IND(D)$, then S is inconsistent; it is consistent, otherwise.

After introducing the indiscernibility relation, objects share the same condition values can be treated as a whole rather than individuals. For consistent decision tables, all objects in the same equivalence class $[x]_\mathbf{C}$ satisfying one and only one class; while for inconsistent decision tables, objects in an equivalence class $[x]_\mathbf{C}$ may take different classes. It is necessary to use a set to indicate all decision classes labelled by some objects in an equivalence class. Skowron proposes a generalized decision $\delta : U \rightarrow 2^{V_d}$ as the set of decision classes of an object [14].

Definition 5. Given a condition attribute set $A \subseteq \mathbf{C}$, the generalized decision of an object $x \in U$ is denoted as:

$$\delta(x|A) = \{I_d(y) | y \in [x]_A\}. \tag{3}$$

The set of generalized decisions of all objects in U is denoted as the general decision $\Delta(A)$, such that

$$\Delta(A) = (\delta(x_1|A), \delta(x_2|A), \ldots, \delta(x_n|A)), \tag{4}$$

where $n = |U|$.

Regarding the definition of consistency and inconsistency, given a decision table S, if there exists one object $x \in U$ such that $|\delta(x|\mathbf{C})| > 1$ then S is inconsistent; if for all objects $x \in U | \delta(x|\mathbf{C})| = 1$ then S is consistent.

Similar to the generalized decision, Slezak [17] proposes the notion of *majority decision* that uses a binary vector for each equivalence class to indicate the decision classes to which its member objects may belong. In general, a *membership distribution function* over decision classes may be used to indicate the degree to which an equivalence class belongs [16]. Kryszkiewicz also distinguishes the generalized decision and the membership distribution function, termed as μ-decision [7]. Zhang et al. [9,30] propose the *maximum distribution criterion* based on the membership distribution function. In the following discussion, we only concentrate on the generalized decision defined in Eq. (3).

Consider a partition $\pi_D = \{D_1, D_2, \ldots, D_m\}$ of the universe U defined by the decision attribute set D, and another partition π_A defined by a condition attribute set A. The equivalence classes induced by the partition π_A are the basic blocks to construct the Pawlak rough set approximations. For a decision class $D_i \in \pi_D$, the lower and upper approximations of D_i with respect to a partition π_A are defined by Pawlak [12]:

$$\underline{apr}_{\pi_A}(D_i) = \{x \in U | [x]_A \subseteq D_i\};$$
$$\overline{apr}_{\pi_A}(D_i) = \{x \in U | [x]_A \cap D_i \neq \emptyset\}. \tag{5}$$

Based on the rough set approximations of D_i defined by π_A, one can divide the universe U into three disjoint regions: the positive region $\mathrm{POS}_{\pi_A}(D_i)$, the boundary region $\mathrm{BND}_{\pi_A}(D_i)$ and the negative region $\mathrm{NEG}_{\pi_A}(D_i)$.

Definition 6. For a decision class $D_i \in \pi_D$, the three regions of D_i with respect to a partition π_A are defined by Pawlak [12]:

$$\mathrm{POS}_{\pi_A}(D_i) = \underline{apr}_{\pi_A}(D_i),$$
$$\mathrm{BND}_{\pi_A}(D_i) = \overline{apr}_{\pi_A}(D_i) - \underline{apr}_{\pi_A}(D_i),$$
$$\mathrm{NEG}_{\pi_A}(D_i) = U - \mathrm{POS}_{\pi_A}(D_i) \cup \mathrm{BND}_{\pi_A}(D_i) = U - \overline{apr}_{\pi_A}(D_i) = \left(\overline{apr}_{\pi_A}(D_i)\right)^c. \tag{6}$$

For the partition $\pi_D = \{D_1, D_2, \ldots, D_m\}$, we can compute its lower and upper approximations in terms of m two-class problems. Then, $\mathrm{POS}_{\pi_A}(\pi_D)$ indicates the union of all the equivalence classes defined by π_A that each for sure can induce a certain decision. $\mathrm{BND}_{\pi_A}(\pi_D)$ indicates the union of all the equivalence classes defined by π_A that each can induce a partial decision. $\mathrm{NEG}_{\pi_A}(\pi_D)$ indicates the union of all the equivalence classes defined by π_A that each cannot induce any decision.

Definition 7. The three regions of the partition π_D with respect to a partition π_A are defined by Yao [26]:

$$\begin{aligned}
&\mathrm{POS}_{\pi_A}(\pi_D) = \bigcup_{1\leqslant i\leqslant m} \mathrm{POS}_{\pi_A}(D_i),\\
&\mathrm{BND}_{\pi_A}(\pi_D) = \bigcup_{1\leqslant i\leqslant m} \mathrm{BND}_{\pi_A}(D_i),\\
&\mathrm{NEG}_{\pi_A}(\pi_D) = U - \mathrm{POS}_{\pi_A}(\pi_D) \cup \mathrm{BND}_{\pi_A}(\pi_D).
\end{aligned} \tag{7}$$

If the decision table is consistent, then

$$\mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D) = U, \quad \text{and } \mathrm{BND}_{\pi_{\mathbf{C}}}(\pi_D) = \mathrm{NEG}_{\pi_{\mathbf{C}}}(\pi_D) = \emptyset.$$

If the decision table is inconsistent, then

$$\mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D) \cup \mathrm{BND}_{\pi_{\mathbf{C}}}(\pi_D) = U, \quad \mathrm{BND}_{\pi_{\mathbf{C}}}(\pi_D) \neq \emptyset, \quad \text{and } \mathrm{NEG}_{\pi_{\mathbf{C}}}(\pi_D) = \emptyset.$$

We can conclude that given a decision table S, if $\mathrm{BND}_{\pi_{\mathbf{C}}}(\pi_D) \neq \emptyset$ then S is inconsistent; it is consistent, otherwise.

Pawlak defines a measure to evaluate the *quality of classification*, or the *degree of dependency of D*, on an attribute set $A \subseteq \mathbf{C}$ [13]:

$$\gamma(\pi_D|\pi_A) = \frac{|\mathrm{POS}_{\pi_A}(\pi_D)|}{|U|}. \tag{8}$$

The γ measure is used by many researchers based on its simple quantitative evaluation of the positive region [1,4,12,15,20].

3. Relative reducts

There are various definitions of relative reducts in rough set theory. Some are only applicable to consistent tables, and others are applicable to both consistent and inconsistent tables. Some researchers apply the definitions for consistent tables only to inconsistent tables without noticing it. This has led to much confusion. In this section we consider three classification properties, and investigate the corresponding definitions of relative reducts. It can be proved that the three definitions of relative reducts are equivalent in consistent tables, but are not equivalent in inconsistent tables.

3.1. Interpretations of relative reducts

Intuitively speaking, a relative reduct of a decision table is a subset of attributes that has the same or similar classification property as that of the entire set of condition attributes. A first step in defining a relative reduct is to examine such properties. We consider the following three properties:

- An attribute set $A \subseteq \mathbf{C}$ is said to preserve the positive region if and only if it produces the same positive region as $\mathbf{C}$ does, i.e., $\mathrm{POS}_{\pi_A}(\pi_D) = \mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D)$. In the Pawlak rough set model, if A preserves the positive region defined by $\mathbf{C}$, it must also preserve the boundary region defined by $\mathbf{C}$. An attribute set $A \subseteq \mathbf{C}$ that preserves both the positive region and the boundary region also is said to preserve the classification quality. In the Pawlak model, we can use the quantitative relation $\gamma(\pi_D|\pi_A) = \gamma(\pi_D|\pi_{\mathbf{C}})$ to reflect the qualitative equivalence.
- An attribute set $A \subseteq \mathbf{C}$ is said to preserve the general decision if and only if it produces the same generalized decisions for all objects as the ones produced by $\mathbf{C}$, i.e., $\Delta(A) = \Delta(\mathbf{C})$.
- An attribute set $A \subseteq \mathbf{C}$ is said to preserve the relative indiscernibility relation if and only if it produces the same relation as $\mathbf{C}$ does, i.e., $IND(A|D) = IND(\mathbf{C}|D)$. If A preserves the relative indiscernibility relation defined by $\mathbf{C}$, it must also preserve the relative discernibility relation defined by $\mathbf{C}$.

A relative reduct can be defined based on these three properties. It should be a minimum set of attributes that preserve a certain property. Pawlak defines a relative reduct by requiring that the positive region of π_D is unchanged [12]. Therefore, a Pawlak relative reduct also is a region preservation reduct, or say, a classification quality reservation reduct. Similarly, we can define a decision preservation reduct, and a relationship preservation reduct.

Definition 8. Given a decision table $S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\})$, we define the following three types of relative reducts:

Region preservation reduct:

An attribute set $R \subseteq \mathbf{C}$ is a region preservation reduct of $\mathbf{C}$ with respect to D if it satisfies the following two conditions:

(s) $\mathrm{POS}_{\pi_R}(\pi_D) = \mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D)$,
(n) for any attribute subset $R' \subset R, \mathrm{POS}_{\pi_{R'}}(\pi_D) \neq \mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D)$;

Decision preservation reduct:

An attribute set $R \subseteq \mathbf{C}$ is a decision preservation reduct of $\mathbf{C}$ with respect to D if it satisfies the following two conditions:

(s) $\Delta(R) = \Delta(\mathbf{C})$,
(n) for any attribute subset $R' \subset R, \Delta(R') \neq \Delta(\mathbf{C})$;

Relationship preservation reduct:

An attribute set $R \subseteq \mathbf{C}$ is a relationship preservation reduct of $\mathbf{C}$ with respect to D if it satisfies the following two conditions:

(s) $IND(R|D) = IND(\mathbf{C}|D)$,
(n) for any attribute subset $R' \subset R, IND(R'|D) \neq IND(\mathbf{C}|D)$.

The first condition (s) indicates the joint sufficiency of the attribute set R, i.e., R is sufficient to preserve a property. The second condition (n) indicates that each attribute in R is individually necessary with respect to the property, i.e., any subset of R is not sufficient to preserve the property.

This definition is based on the following important monotocity property. Consider any two subsets of attributes $A, B \subseteq \mathbf{C}$ with $A \subseteq B$. For any $x \in U$, we have $[x]_B \subseteq [x]_A$. We immediately obtain the monotocity of the relative indiscernibility relations, general decisions and positive regions with respect to set inclusion of attributes.

Proposition 1. *For any two subsets of attributes $A, B \subseteq \mathbf{C}$ with $A \subseteq B$,*

$$A \subseteq B \Rightarrow \mathrm{POS}_{\pi_A}(\pi_D) \subseteq \mathrm{POS}_{\pi_B}(\pi_D);$$
$$A \subseteq B \Rightarrow \delta(x|B) \subseteq \delta(x|A) \text{ for all } x \in U;$$
$$A \subseteq B \Rightarrow IND(B|D) \subseteq IND(A|D).$$

According to Proposition 1, the attribute set $\mathbf{C}$ produces the largest positive region $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$, the smallest generalized decision $\delta(x|\mathbf{C})$ for all objects, and the smallest relative indiscernibility relation $IND(\mathbf{C}|D)$. A relative reduct R is able to preserve the same result regarding a certain property, and none subset of R can produce a better result than R does. Therefore, the equality relation is suitable for condition (s). According to Proposition 1, given an attribute set $R' \subset R$, if R' cannot preserve the same result as $\mathbf{C}$ does regarding a certain property, then none of the proper subsets of R' can. Thus, as stated in condition (n), one only needs to check the proper subsets $R - \{a\}$ for all $a \in R$ in order to verify a relative reduct R. Therefore, Definition 8 is well-defined.

A region preservation reduct, also is a standard Pawlak relative reduct [12], is based on the positive regions defined by different attribute sets. For consistent decision tables, there is no need to consider the boundary region or the negative region because they both are the empty sets. It is sufficient to consider only the positive rule set in consistent tables. For inconsistent decision tables, we have $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) \cap \mathrm{BND}_{\pi_\mathbf{C}}(\pi_D) = \emptyset, \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) \cup \mathrm{BND}_{\pi_\mathbf{C}}(\pi_D) = U$, and the boundary region is not empty. The condition $\mathrm{POS}_{\pi_R}(\pi_D) = \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$ is equivalent to $\mathrm{BND}_{\pi_R}(\pi_D) = \mathrm{BND}_{\pi_\mathbf{C}}(\pi_D)$. It is sufficient to consider only the positive regions in the inconsistent tables.

Various types of relative reducts also have been studied by many researchers. For decision preservation, besides the generalized decisions, one also can address other decision related functions, such as the membership distribution function and the maximum distribution function. Alternative decision preservation reducts have been studied [7,9,16]. For relationship preservation, besides the relative indiscernibility relation, one can also preserve the relative discernibility relation, or both. Relative reducts for object pair relationship preservation also have been studied [18,33]. The quantitative alternative of region preservation reducts, also called classification quality preservation reducts, is studied by many authors [1,3,4,12,20]. The conditional entropy reflects the classification accuracy from the information viewpoint, and can be used for reduct computation. In general, any monotonic measure f can be used to evaluate positive region preservation if it satisfies the condition

$$(f(\pi_D|\pi_A) = f(\pi_D|\pi_\mathbf{C})) \iff \left(\mathrm{POS}_{\pi_A}(\pi_D) = \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)\right).$$

For example, Shannon's entropy and many of its variations have been explored to measure the uncertainty in rough set theory [1,2,6,8,23], and thus can be understood as different forms of the f measure.

There may exist more than one reduct in an information table. We denote $RED_{\mathrm{region}}(\mathbf{C}|D)$ as the set of region preservation reducts, $RED_{\mathrm{decision}}(\mathbf{C}|D)$ as the set of decision preservation reducts, and $RED_{\mathrm{relationship}}(\mathbf{C}|D)$ as the set of relationship preservation reducts.

The following theorem states the equivalence relation among the three properties in consistent decision tables.

Theorem 1. *Given a consistent decision table $S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\})$, for an attribute set $A \subseteq \mathbf{C}$, the following conditions are equivalent:*

(i) $\mathrm{POS}_{\pi_A}(\pi_D) = \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$,
(ii) $\Delta(A) = \Delta(\mathbf{C})$,
(iii) $IND(A|D) = IND(\mathbf{C}|D)$.

The equivalence of the three properties stated in Theorem 1 indicates that the three definitions of relative reducts are in fact equivalent in consistent decision tables. Different forms of the definitions yield the same set of relative reducts. In other words, the definition of the Pawlak relative reducts preserves the positive region, relative relationship, and general decision that are defined by the attribute set **C** at the same time.

The following theorem indicates that the three properties may not be equivalent in inconsistent decision tables. Furthermore, there is an ordering among them.

Theorem 2. *Given an inconsistent decision table* $S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\})$*, for an attribute set* $A \subseteq \mathbf{C}$*, consider the following conditions:*

(i) $\mathrm{POS}_{\pi_A}(\pi_D) = \mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D)$,
(ii) $\Delta(A) = \Delta(\mathbf{C})$,
(iii) $IND(A|D) = IND(\mathbf{C}|D)$.

Condition (ii) implies condition (i), and condition (iii) implies conditions (i) and (ii).

Theorem 2 indicates that in inconsistent decision tables, a Pawlak relative reduct preserves the positive region defined by **C**, but does not necessarily preserve the general decision or the object pair relationship. Furthermore, the relationship preservation is the strongest criterion, the region preservation is the weakest, and the decision preservation is in the middle. Preserving a stronger property guarantees that a weaker property also is preserved, while the reversed proposition does not hold. The sets $RED_{\mathrm{decision}}(\mathbf{C}|D)$ and $RED_{\mathrm{relationship}}(\mathbf{C}|D)$ might be different, and they might not be equivalent to $RED_{\mathrm{region}}(\mathbf{C}|D)$.

3.2. A general definition of a relative reduct

In the rough set literature, there exist many different definitions, just as the above three we have discussed. Along with the increasing requirements of data analysis, we may need to define more properties of an information table; this naturally leads to more definitions of reducts. All these definitions have the same structure. Therefore, a general definition of an attribute reduct is necessary and useful [32]. Here, we provide a generalized definition for relative reducts.

Definition 9. Given a decision table $S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\})$ and a certain property $\mathbb{P}$ of S regarding the decision attribute set D. The attribute set $R \subseteq \mathbf{C}$ is a relative reduct of **C** if it satisfies the following three conditions:

(p) evaluate $\mathbb{P}$ by a function $e : 2^{At} \to L$, which maps an attribute set of 2^{At} to an element of a poset L;
(s) $e(R) = e(\mathbf{C})$;
(n) for any $R' \subset R, e(R') \neq e(\mathbf{C})$.

According to the above analysis, the property $\mathbb{P}$ can be interpreted as region preservation, relationship preservation or decision preservation. By applying the function e, we are able to pick the attribute set that preserves the property $\mathbb{P}$. The mapping yields different sets of reducts, such as $RED_{\mathrm{region}}(\mathbf{C}|D), RED_{\mathrm{decision}}(\mathbf{C}|D)$ and $RED_{\mathrm{relationship}}(\mathbf{C}|D)$.

Skowron and Rauszer define [15] that an attribute $a \in \mathbf{C}$ is relatively indispensable in **C** if

$$\mathrm{POS}_{\pi_{\mathbf{C}-\{a\}}}(\pi_D) \neq \mathrm{POS}_{\pi_{\mathbf{C}}}(\pi_D),$$

otherwise a is said to be relatively dispensable in **C**. The set of relatively indispensable attributes is called a *relative core*, or more specifically, a region preservation core, denoted as $CORE_{\mathrm{region}}(\mathbf{C}|D)$. Similarly, we can define a relative core for decision preservation and relationship preservation as follows:

$$CORE_{\mathrm{decision}}(\mathbf{C}|D) = \{a \in \mathbf{C} | \Delta(\mathbf{C} - \{a\}) \neq \Delta(\mathbf{C})\};$$

$$CORE_{\mathrm{relationship}}(\mathbf{C}|D) = \{a \in \mathbf{C} | IND(\mathbf{C} - \{a\}|D) \neq IND(\mathbf{C}|D)\}.$$

For various relative reducts and cores, the following proposition indicates that a relative core is in fact the intersection of the corresponding set of relative reducts.

Proposition 2. $CORE_{\mathbb{P}}(\mathbf{C}|D) = \bigcap RED_{\mathbb{P}}(\mathbf{C}|D)$.

It is crucial to note that the above general definition of relative reducts is based on the assumption that the monotocity of the function e is held with respect to set inclusion of attributes. In this case, R can preserve the biggest value as the entire condition attribute set **C** does regarding the function e, and no proper subset of R can reach the same value. For simplicity, only the proper subsets $R - \{a\}$ for all $a \in R$ need to be checked. This important assumption is true in the Pawlak rough set model, but is not true in probabilistic rough set models [27]. It means that the above general definition is applicable for the Pawlak model only.

4. Relative reduct construction based on discernibility matrices

A theoretical result has been developed by Skowron and Rauszer based on the notion of a discernibility matrix [15]. In this section, we first review the standard definition of a discernibility matrix, and the inference of the corresponding discernibility function. Definitions of different discernibility matrices are then introduced and compared. Finally, a generalized definition of discernibility matrices is proposed.

4.1. Classical discernibility matrix and discernibility function

Skowron and Rauszer suggest a matrix representation for storing the sets of attributes that discern pairs of objects, called a *discernibility matrix* [15]. Both the rows and columns of the matrix correspond to the objects. An element of the matrix is the set of all attributes that distinguishes the corresponding object pairs.

Definition 10. Given an information table S, its discernibility matrix $M = (M(x,y))$ is a $|U| \times |U|$ matrix, in which the element $M(x,y)$ for an object pair (x,y) is defined by:

$$M(x,y) = \{a \in At | I_a(x) \neq I_a(y)\}. \tag{9}$$

The physical meaning of the matrix element $M(x,y)$ is that objects x and y can be distinguished by any attribute in $M(x,y)$. The pair (x,y) can be discerned if $M(x,y) \neq \emptyset$. A discernibility matrix M is symmetric, i.e., $M(x,y) = M(y,x)$, and $M(x,x) = \emptyset$. Therefore, it is sufficient to consider only the lower triangle or the upper triangle of the matrix.

From a discernibility matrix, one can define the notion of a discernibility function [15].

Definition 11. The discernibility function of a discernibility matrix is defined by:

$$f(M) = \bigwedge \left\{ \bigvee (M(x,y)) | \forall x,y \in U[M(x,y) \neq \emptyset] \right\}. \tag{10}$$

The expression $\bigvee(M(x,y))$ is the disjunction of all attributes in $M(x,y)$, indicating that the object pair (x,y) can be distinguished by any attribute in $M(x,y)$. The expression $\bigwedge\{\bigvee(M(x,y))\}$ is the conjunction of all $\bigvee(M(x,y))$, indicating that the family of discernible object pairs can be distinguished by a set of attributes satisfying $\bigwedge\{\bigvee(M(x,y))\}$.

The discernibility function can be used to state an important result regarding the set of reducts of an information table, as shown by the following theorem from Skowron and Rauszer [15].

Theorem 3. *The reduct set problem is equivalent to the problem of transforming the discernibility function to a reduced disjunctive form. Each conjunctor of the reduced disjunctive form is called a prime implicant. Given the discernibility matrix M of an information table S, an attribute set $R = \{a_1, \ldots, a_p\}$ is a reduct if and only if the conjunction of all attributes in R, denoted as $a_1 \wedge \ldots \wedge a_p$, is a prime implicant of $f(M)$.*

In order to derive the reduced disjunctive form, the discernibility function $f(M)$ is transformed by using the absorption and distribution laws. Accordingly, finding the set of reducts can be modelled based on the manipulation of a Boolean function. The set reducts of an information table is equivalent to the set of prime implicants of the discernibility function.

4.2. Related work

For a decision table, the discernibility matrix also can be defined. A typical definition of a discernibility matrix for decision tables is proposed by Hu and Cercone, in which an element is defined as [5]:

$$M(x,y) = \{a \in \mathbf{C} | [I_a(x) \neq I_a(y)] \wedge [I_d(x) \neq I_d(y)]\}. \tag{11}$$

That is, we only consider the discernibility of objects in different decision classes.

Table 1
A simple decision table and its discernibility matrix defined by Eq. (11).

U	a_1	a_2	d
o_1	1	0	1
o_2	1	0	2
o_3	0	0	1
o_4	0	0	0
o_5	1	1	1

	o_1	o_2	o_3	o_4	o_5
o_1	$\emptyset$				
o_2	$\emptyset$	$\emptyset$			
o_3	$\emptyset$	$\{a_1\}$	$\emptyset$		
o_4	$\{a_1\}$	$\{a_1\}$	$\emptyset$	$\emptyset$	
o_5	$\emptyset$	$\{a_2\}$	$\emptyset$	$\{a_1, a_2\}$	$\emptyset$

An interesting counter example has been offered by Ye et al., and has led to the doubt on the descriptive power of the discernibility matrix [29]. Let us first have a look at the decision table and its discernibility matrix shown in Table 1. The table is slightly modified from the original example. According to the discernibility matrix defined in Eq. (11) and the discernibility function, we can easily verify that $\{a_1, a_2\}$ is a relative reduct. According to Definition 7, we obtain that $\mathrm{POS}_{\pi_{\{a_1,a_2\}}}(\pi_D) = \{o_5\}, \mathrm{POS}_{\pi_{\{a_1\}}}(\pi_D) = \emptyset$ and $\mathrm{POS}_{\pi_{\{a_2\}}}(\pi_D) = \{o_5\}$, and thus $\{a_2\}$ is a relative reduct. This is a contradiction. If $\{a_2\}$ is a reduct then $\{a_1, a_2\}$ cannot be a reduct. How to explain this result?

We can do some computation regarding the relative indiscernibility relations. According to the definition, we obtain:

$$\begin{aligned}IND(\{a_1,a_2\}|D) = \{&(o_1,o_1),(o_1,o_2),(o_1,o_3),(o_1,o_5),(o_2,o_1),(o_2,o_2),(o_3,o_1),\\&(o_3,o_3),(o_3,o_4),(o_3,o_5),(o_4,o_3),(o_4,o_4),(o_5,o_1),(o_5,o_3),(o_5,o_5)\};\\IND(\{a_1\}|D) = \Big\{&(o_1,o_1),(o_1,o_2),(o_1,o_3),(o_1,o_5),(o_2,o_1),(o_2,o_2),\underline{(o_2,o_5)},\\&(o_3,o_1),(o_3,o_3),(o_3,o_4),(o_3,o_5),(o_4,o_3),(o_4,o_4),(o_5,o_1),\underline{(o_5,o_2)},(o_5,o_3),(o_5,o_5)\Big\};\\IND(\{a_2\}|D) = \Big\{&(o_1,o_1),(o_1,o_2),(o_1,o_3),\underline{(o_1,o_4)},(o_1,o_5),(o_2,o_1),(o_2,o_2),\underline{(o_2,o_3)},\underline{(o_2,o_4)},(o_3,o_1),\\&\underline{(o_3,o_2)},(o_3,o_3),(o_3,o_4),(o_3,o_5),\underline{(o_4,o_1)},\underline{(o_4,o_2)},(o_4,o_3),(o_4,o_4),(o_5,o_1),(o_5,o_3),(o_5,o_5)\Big\}.\end{aligned}$$

This indicates that neither $\{a_1\}$ nor $\{a_2\}$ is sufficient to keep the relative indiscernibility relation. Hence, $\{a_1, a_2\}$ is a relationship preservation reduct.

We also can do some computation regarding the general decisions. According to the definition, we obtain:

$$\begin{aligned}&\delta(o_1|\{a_1,a_2\}) = \{1,2\}, \quad \delta(o_1|\{a_1\}) = \{1,2\}, \quad \delta(o_1|\{a_2\}) = \underline{\{0,1,2\}},\\&\delta(o_2|\{a_1,a_2\}) = \{1,2\}, \quad \delta(o_2|\{a_1\}) = \{1,2\}, \quad \delta(o_2|\{a_2\}) = \underline{\{0,1,2\}},\\&\delta(o_3|\{a_1,a_2\}) = \{0,1\}, \quad \delta(o_3|\{a_1\}) = \{0,1\}, \quad \delta(o_3|\{a_2\}) = \underline{\{0,1,2\}},\\&\delta(o_4|\{a_1,a_2\}) = \{0,1\}, \quad \delta(o_4|\{a_1\}) = \{0,1\}, \quad \delta(o_4|\{a_2\}) = \underline{\{0,1,2\}},\\&\delta(o_5|\{a_1,a_2\}) = \{1\}; \quad \delta(o_5|\{a_1\}) = \underline{\{1,2\}}; \quad \delta(o_5|\{a_2\}) = \{1\}.\end{aligned}$$

This indicates that neither $\{a_1\}$ nor $\{a_2\}$ is sufficient to keep the same generalized decisions for all objects. Hence, $\{a_1, a_2\}$ is a decision preservation reduct.

Conclusively, both $\{a_2\}$ and $\{a_1, a_2\}$ are relative reducts. While $\{a_2\}$ is a region preservation reduct, $\{a_1, a_2\}$ is both a decision preservation reduct and a relationship preservation reduct. To judge if an attribute set is a reduct or not without considering the property it preserves is not meaningful, and may lead to a wrong judgement.

4.3. Specific discernibility matrices

Regarding the general definition of a relative reduct, given a certain property $\mathbb{P}$ of S, we can define a discernibility matrix for storing the sets of attributes that discern pairs of objects regarding the property $\mathbb{P}$. In this subsection, we focus on the three property preservation matrices. They are relative relationships, general decisions and positive regions.

Regarding the region preservation property, given two objects $x, y \in U$, if $[x]_\mathbf{C} \in \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$ and $[y]_\mathbf{C} \in \mathrm{BND}_{\pi_\mathbf{C}}(\pi_D)$ then they are considered discernible; if both of them belong to the positive region, they may still be distinguished when $[x]_\mathbf{C} \in \mathrm{POS}_{\pi_\mathbf{C}}(D_i)$ and $[y]_\mathbf{C} \in \mathrm{POS}_{\pi_\mathbf{C}}(D_j)$ where $D_i \neq D_j$; if both of them belong to the boundary region, they are considered indiscernible. Hence, the discernibility matrix for region preservation can be defined as follows.

Definition 12. Given a decision table S, its discernibility matrix $M_{\mathrm{region}} = (M_{\mathrm{region}}(x,y))$ for region preservation is a $|U| \times |U|$ matrix, in which the element $M_{\mathrm{region}}(x,y)$ for an object pair (x,y) is defined by:

$$M_{\mathrm{region}}(x,y) = \begin{cases}\{a \in \mathbf{C} | I_a(x) \neq I_a(y)\}, & [x] \text{ or } [y] \in \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D),\\ \emptyset, & \text{otherwise}.\end{cases} \tag{12}$$

Skowron and Rauszer's definition of a discernibility matrix [15] is equivalent to this definition. Another version is provided by Ye and Chen [29] by using the cardinality of generalized decisions. The condition $\min\{|\delta(x|\mathbf{C})|, |\delta(y|\mathbf{C})|\} = 1$ means that at least one of the equivalence classes $[x]_\mathbf{C}$ and $[y]_\mathbf{C}$ belongs to the positive region, and thus (x,y) needs to be compared.

A discernibility matrix for decision preservation only keeps track of the object pairs that have different generalized decisions, and can be defined as follows.

Definition 13. Given a decision table S, its discernibility matrix $M_{\mathrm{decision}} = (M_{\mathrm{decision}}(x,y))$ for decision preservation is a $|U| \times |U|$ matrix, in which the element $M_{\mathrm{decision}}(x,y)$ for an object pair (x,y) is defined by:

$$M_{\mathrm{decision}}(x,y) = \begin{cases}\{a \in \mathbf{C} | I_a(x) \neq I_a(y)\}, & \delta(x|\mathbf{C}) \neq \delta(y|\mathbf{C}),\\ \emptyset, & \text{otherwise}.\end{cases} \tag{13}$$

A discernibility matrix for relationship preservation only keeps track of the object pairs that satisfy the relative discernibility relation defined by $\mathbf{C}$, and can be defined as follows.

Definition 14. Given a decision table S, its discernibility matrix $M_{\text{relationship}} = (M_{\text{relationship}}(x,y))$ for relationship preservation is a $|U| \times |U|$ matrix, in which the element $M_{\text{relationship}}(x,y)$ for an object pair (x,y) is defined by:

$$M_{\text{relationship}}(x,y) = \begin{cases} \{a \in \mathbf{C} | I_a(x) \neq I_a(y)\}, & (x,y) \in DIS(\mathbf{C}|D), \\ \emptyset, & \text{otherwise}. \end{cases} \quad (14)$$

Hu and Cercone's definition of a discernibility matrix [5] is equivalent to this definition, by having $I_d(x) \neq I_d(y)$ to simplify the relative discernibility relation.

From these three constructive definitions, one can also investigate the superset-subset relationship among the three properties. The matrix for relationship preservation compares all object pairs that have different decision classes. The matrix for decision preservation compares less pairs. Those object pairs having different decision classes are not compared if they have the same generalized decisions. The matrix for region preservation compares least pairs. Those object pairs having different generalized decisions are not compared if both the equivalence classes belong to the boundary region.

4.4. A general definition of a discernibility matrix

To summarize the common structure of the discernibility matrices for different property preservations, we here provide a general definition of discernibility matrices.

Definition 15. Given a decision table $S = (U, At = \mathbf{C} \cup D, \{V_a | a \in At\}, \{I_a | a \in At\})$ and a certain property $\mathbb{P}$ of S regarding the decision attribute set D. The discernibility matrix $M_{\mathbb{P}} = (M_{\mathbb{P}}(x,y))$ of the property $\mathbb{P}$ is a $|U| \times |U|$ matrix, in which the element $M_{\mathbb{P}}(x,y)$ for an object pair (x,y) is defined by:

$$M_{\mathbb{P}}(x,y) = \begin{cases} \{a \in \mathbf{C} | I_a(x) \neq I_a(y)\}, & (x,y) \text{ are distinguishable w.r.t. } \mathbb{P}, \\ \emptyset, & \text{otherwise}. \end{cases} \quad (15)$$

Two fundamental concepts – reduct and core – can be constructed in the discernibility matrix. For classification problems, one can verify that the relative core, $CORE_{\mathbb{P}}(\mathbf{C}|D)$, can be characterized by a discernibility matrix $M_{\mathbb{P}}$ in the following way:

Proposition 3

$$CORE_{\mathbb{P}}(\mathbf{C}|D) = \{a \in \mathbf{C} | M_{\mathbb{P}}(x,y) = \{a\} \text{ for some } x \text{ and } y\}.$$

The discernibility function in Definition 11 offers a logical and systematic way to compute the set of all reducts. However, a difficulty exists when applying the theoretical results. It is not very clear how to design an efficient algorithm for constructing *one* reduct based on manipulating the discernibility function. Many efforts have been made to construct reducts based on the matrices. For example, Miao et al. introduce Wu's method to transform a matrix problem to a linear algebra problem [10]. Yao and Zhao propose a reduct construction method based on discernibility matrix simplification [28]. The method works in a similar way to the classical Gaussian elimination method for solving a system of linear equations. Various heuristic algorithms also can be applied to all types of discernibility matrices [8,11,21,22,31].

5. Conclusion

A relative reduct is a minimum set of attributes that keeps a particular classification property. Three different properties are discussed in this paper for classification tasks. In consistent decision tables, they yield the same set of relative reducts. In inconsistent decision table, they may result in different sets. To suit with different properties, different discernibility matrices are defined. Uncoupling the definitions of a reduct from its discernibility matrix may lead to the failure of reduct construction or wrong results. To summarize the common structures of the specific definitions, the general definitions of relative reducts and discernibility matrices are suggested.

The study also tries to remark on the generalization problem. A simple definition may work very well for the standard and simple model by integrating several interpretations explicitly and implicitly. Moving forward to a more general model, these interpretations may not be equivalent, and need to be treated differently and separately. A straightforward or oversimplified generalization may take the risk for missing some interpretations, or mixing up the whole concept.

From consistent decision tables to inconsistent decision tables, we encounter the problem of generalizing the definition of a relative reduct. This paper is one of the efforts to differentiate the embedded interpretations and properties that are equivalent and might not need to be distinguished in consistent tables, but are different and need to be differentiated in inconsistent tables. To generalize the Pawlak model to probabilistic models, more complicated situations will be encountered while the monotonic feature might not hold for some properties.

Acknowledgements

The authors thank the anonymous referees for the constructive comments and suggestions. This work is partially supported by the National Natural Science Foundation of China, Grant No. 60775036, the Specialized Research Fund for the Doctoral Program of Higher Education of China, Grant No. 2006024703, and a Discovery Grant from NSERC Canada.

Appendix A. Proofs of proposed theorems and propositions

A.1. Proof of Theorem 2

"(ii) $\Rightarrow$ (i)" Based on the monotocity of the positive regions with respect to set inclusion of attributes, we have $A \subseteq \mathbf{C} \Rightarrow \mathrm{POS}_{\pi_A}(\pi_D) \subseteq \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$. We thus should prove that $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) \subseteq \mathrm{POS}_{\pi_A}(\pi_D)$. For any $x \in U$, $[x]_\mathbf{C} \in \mathrm{POS}_{\pi_\mathbf{C}}(\pi_D)$ indicates that $[x]_\mathbf{C}$ has one and only one decision value denoted by v_1, i.e., $\delta(x|\mathbf{C}) = \{v_1\}$. Because $\delta(x|A) = \delta(x|\mathbf{C})$, then $\delta(x|A) = \{v_1\}$, which indicates that $[x]_A \in \mathrm{POS}_{\pi_A}(\pi_D)$. Therefore, $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) \subseteq \mathrm{POS}_{\pi_A}(\pi_D)$.

"(ii) $\Leftarrow$ (i) is not true". Proof by a counter example shown in Table 2. Investigate the attribute sets **C** and $\{a_1, a_2\}$. The computation shows that $\Delta(\{a_1, a_2\}) \neq \Delta(\mathbf{C})$, and thus $\{a_1, a_2\}$ does not preserve the general decision defined by **C**. However, also in the previous example, $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) = \mathrm{POS}_{\pi_{\{a_1,a_2\}}}(\pi_D) = \{o_1\}$. Thus, $\{a_1, a_2\}$ preserves the positive region defined by **C**. This indicates that an attribute set preserving the region does not necessarily preserve the general decision.

"(iii) $\Rightarrow$ (ii)" Proof by contradiction. Suppose there exists an object $x_1 \in U$ such that $\delta(x_1|A) \neq \delta(x_1|\mathbf{C})$. Based on the monotocity of the general decisions with respect to set inclusion of attributes, we have $A \subseteq \mathbf{C} \Rightarrow \delta(x|\mathbf{C}) \subset \delta(x|A)$. In other words, there must exist a decision value $v_1 \in V_d$ such that $v_1 \in \delta(x_1|A)$ and $v_1 \notin \delta(x_1|\mathbf{C})$. Denote an object y_1 having $I_d(y_1) = v_1$, then $y_1 \in [x_1]_A$ and $y_1 \notin [x_1]_\mathbf{C}$. This indicates that $(x_1, y_1) \in IND(A|D)$ and $(x_1, y_1) \notin IND(\mathbf{C}|D)$. Therefore $IND(A|D) \neq IND(\mathbf{C}|D)$.

"(iii) $\Leftarrow$ (ii) is not true" Proof by a counter example shown in Table 3. Investigate the attribute sets **C** and $\{a_1, a_2\}$. The computation shows that $\Delta(\{a_1, a_2\}) = \Delta(\mathbf{C})$, and thus $\{a_1, a_2\}$ preserves the general decision defined by **C**. At the same time, $\mathrm{POS}_{\pi_\mathbf{C}}(\pi_D) = \mathrm{POS}_{\pi_{\{a_1,a_2\}}}(\pi_D) = \{o_1\}$. Thus, $\{a_1, a_2\}$ preserves the positive region defined by **C**. However, by reducing **C** to $\{a_1, a_2\}$, we cannot distinguish the object pairs $(o_2, o_5), (o_3, o_4), (o_4, o_3)$ and (o_5, o_2) which are originally discernible. This indicates that an attribute set preserving the general decision, or the positive region, does not necessarily preserve the object pair relationship defined by **C**.

"(iii) $\Rightarrow$ (i)" can be proved based on the fact that (iii) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i). "(iii) $\Leftarrow$ (i) is not true" can be proved by the same counter example shown in Table 3. □

A.2. Proof of Proposition 2

We show the proof for $CORE_{\mathrm{relationship}}(\mathbf{C}|D) = \bigcap RED_{\mathrm{relationship}}(\mathbf{C}|D)$. where $M(x, y)$ is computed according to Definition 14.

"$\Rightarrow$" Proof by contradiction. According to the definition, for any core attribute $a \in CORE_{\mathrm{relationship}}(\mathbf{C}|D)$, $IND(\mathbf{C}|D) \neq IND(\mathbf{C} - \{a\}|D)$. Based on the monotocity of the relative indiscernibility relations with respect to set inclusion of attributes, $IND(\mathbf{C}|D) \subset IND(\mathbf{C} - \{a\}|D)$. It means that there exists an object pair $(x_1, y_1) \in IND(\mathbf{C} - \{a\}|D)$ and

Table 2
A counter example to show that an attribute set which is region preservable is not necessarily decision preservable.

U	a_1	a_2	a_3	D	$\Delta(\mathbf{C})$	$\Delta(\{a_1, a_2\})$
o_1	0	1	1	1	{1}	{1}
o_2	1	0	0	1	{1,2}	{1,2,3}
o_3	1	0	0	2	{1,2}	{1,2,3}
o_4	1	0	1	3	{2,3}	{1,2,3}
o_5	1	0	1	2	{2,3}	{1,2,3}

Table 3
A counter example to show that an attribute set which is decision preservable or region preservable is not necessarily relationship preservable.

U	a_1	a_2	a_3	D	$\Delta(\mathbf{C})$	$\Delta(\{a_1, a_2\})$
o_1	0	1	1	1	{1}	{1}
o_2	1	0	0	1	{1,2}	{1,2}
o_3	1	0	0	2	{1,2}	{1,2}
o_4	1	0	1	1	{1,2}	{1,2}
o_5	1	0	1	2	{1,2}	{1,2}

$(x_1, y_1) \notin IND(\mathbf{C}|D)$. This indicates that for all $c \in \mathbf{C} - \{a\}, I_c(x_1) = I_c(y_1) \vee I_d(x_1) = I_d(y_1)$ and there exists $b \in \mathbf{C}, I_b(x_1) \neq I_b(y_1) \wedge I_d(x_1) \neq I_d(y_1)$. Thus, $b = a$ such that $I_a(x_1) \neq I_a(y_1)$. Suppose $R_1 \subseteq \mathbf{C}$ is a relationship preservation reduct and $a \notin R_1$. Then, $(x_1, y_1) \in IND(R_1|D)$ and $(x_1, y_1) \notin IND(R_1 \cup \{a\}|D)$. Since $IND(\mathbf{C}|D) \subseteq IND(R_1 \cup \{a\}|D) \subset IND(R_1|D)$, then R_1 is not a relative reduct.

"$\Leftarrow$" Proof by contradiction. Suppose $a \notin CORE_{\text{relationship}}(\mathbf{C}|D)$. According to the definition of a relationship preservation core, $IND(\mathbf{C}|D) = IND(\mathbf{C} - \{a\}|D)$. Furthermore, there exists $R_1 \subseteq \mathbf{C} - \{a\}$ which is a relative reduct. It indicates that $a \notin R_1$. □

A.3. Proof of Proposition 3

We show the proof of a relationship preservation core. i.e., $CORE_{\text{relationship}}(\mathbf{C}|D) = \{a \in \mathbf{C}|M(x, y) = \{a\} \text{for some } x \text{ and } y\}$, where $M(x, y)$ is computed according to Definition 14. The interpretation of a region preservation core has been proved by Skowron and Rauszer [15]. Other Interpretations can be similarly proved.

"$\Rightarrow$" For any $a \in CORE_{\text{relationship}}(\mathbf{C}|D), IND(\mathbf{C}|D) \neq IND(\mathbf{C} - \{a\}|D)$, thus $IND(\mathbf{C}|D) \subset IND(\mathbf{C} - \{a\}|D)$. It indicates that there exists $(x_1, y_1) \in IND(\mathbf{C} - \{a\}|D)$ but $(x_1, y_1) \notin IND(\mathbf{C}|D)$. Thus, $(\forall b \in \mathbf{C} - \{a\}[I_b(x_1) = I_b(y_1)]) \vee (I_d(x_1) = I_d(y_1))$ and $(\exists a \in \mathbf{C}[I_{\mathbf{C}}(x_1) \neq I_{\mathbf{C}}(y_1)]) \wedge (I_d(x_1) = I_d(y_1))$. These two results imply that $I_b(x_1) = I_b(y_1)$ for all $b \in \mathbf{C} - \{a\}$, and thus $M(x_1, y_1) \cap (\mathbf{C} - \{a\}) = \emptyset$ and $M(x_1, y_1) \cap \mathbf{C} \neq \emptyset$. Thus, $M(x_1, y_1) = \{a\}$.

"$\Leftarrow$" For any $a \in \mathbf{C}$ such that $\{a\} \in M$, there exists $(x_1, y_1) \in U \times U$ satisfying $M(x_1, y_1) = \{a\}$. This indicates that $(x_1, y_1) \notin IND(\mathbf{C}|D)$ and $(x_1, y_1) \in IND(\mathbf{C} - \{a\}|D)$. These two results imply that $IND(\mathbf{C}|D) \neq IND(\mathbf{C} - \{a\}|D)$. Thus, $a \in CORE_{\text{relationship}}(\mathbf{C}|D)$. □

References

[1] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1998) 185–195.
[2] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1998) 77–107.
[3] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition Letters 40 (2007) 3509–3521.
[4] Q.H. Hu, D.R. Yu, Z.X. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, Pattern Recognition Letters 27 (2006) 414–423.
[5] X.H. Hu, N. Cercone, Learning in relational databases: a rough set approach, Computational Intelligence 11 (1995) 323–338.
[6] J. Klir, M.J. Wierman, Uncertainty Based Information: Elements of Generalized Information Theory, Physica-Verlag, New York, 1999.
[7] M. Kryszkiewicz, Comparative study of alternative types of knowledge reduction in inconsistent systems, International Journal of Intelligent Systems 16 (2001) 105–120.
[8] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (2004) 37–46.
[9] J.S. Mi, W.Z. Wu, W.X. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (2004) 255–272.
[10] D.Q. Miao, J. Zhou, N. Zhang, N. Feng, R.Z. Wang, Research of attribute reduction based on algebraic equations, Acta Electronica Sinica, accepted for publication. (in Chinese)
[11] S.H. Nguyen, H.S. Nguyen, Some efficient algorithms for rough set methods, Proceedings of the International Conference on Information Processing and Management of Uncertainty on Knowledge Based Systems (1996) 1451–1456.
[12] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.
[13] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Boston, 1991.
[14] A. Skowron, Boolean reasoning for decision rules generation, Proceedings of the International Symposium on Methodologies for Intelligent Systems (1993) 295–305.
[15] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R. Slowiński (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer, Dordrecht, 1992.
[16] D. Slezak, Approximate reducts in decision tables, Proceedings of Information Processing and Management of Uncertainty (1996) 1159–1164.
[17] D. Slezak, Normalized decision functions and measures for inconsistent decision tables analysis, Fundamenta Informaticae 44 (2000) 291–319.
[18] R. Susmaga, Reducts and constructs in attribute reduction, Fundamenta Informaticae 61 (2) (2004) 159–181.
[19] G.Y. Wang, Calculation methods for core attributes of decision table, Chinese Journal of Computers 26 (2003) 611–615. in Chinese.
[20] G.Y. Wang, J. Zhao, J. Wu, A comparitive study of algebra viewpoint and information viewpoint in attribute reduction, Foundamenta Informaticae 68 (2005) 1–13.
[21] J. Wang, D.Q. Miao, Analysis on attribute reduction strategies of rough set, Chinese Journal of Computer Science and Technology 13 (1998) 189–192.
[22] J. Wang, J. Wang, Reduction algorithms based on discernibility matrix: the ordered attributes method, Journal of Computer Science and Technology 16 (2001) 489–504.
[23] M.J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems 28 (1999) 283–297.
[24] W.Z. Wu, M. Zhang, H.Z. Li, J.S. Mi, Knowledge reduction in random information systems via Dempster–Shafer theory of evidence, Information Sciences 174 (2005) 143–164.
[25] M. Yang, Z.H. Sun, Improvement of discernibility matrix and the computation of a core, Journal of Fudan University (Natural Science) 43 (2004) 865–868. in Chinese.
[26] Y.Y. Yao, Decision-theoretic rough set models, Proceedings of the Second International Conference on Rough Sets and Knowledge Technology LNAI 4481 (2007) 1–12.
[27] Y.Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, Information Sciences 178 (2008) 3356–3373.
[28] Y.Y. Yao, Y. Zhao, Discernibility matrix simplification for constructing attribute reducts, Information Sciences 179 (7) (2009) 867–882.
[29] D.Y. Ye, Z.J. Chen, An improved discernibility matrix for computing all reducts of an inconsistent decision table, The Proceedings of the Fifth IEEE International Conference on Cognitive Informatics (2006) 305–308.
[30] W.X. Zhang, J.S. Mi, W.Z. Wu, Knowledge reduction in inconsistent information systems, Chinese Journal of Computers 1 (2003) 12–18.
[31] K. Zhao, J. Wang, A reduction algorithm meeting users' requirements, Journal of Computer Science and Technology 17 (2002) 578–593.
[32] Y. Zhao, F. Luo, S.K.M. Wong, Y.Y. Yao, A general definition of an attribute reduct, Proceedings of the Second Rough Sets and Knowledge Technology (2007) 101–108.
[33] Y. Zhao, Y.Y. Yao, F. Luo, Data analysis based on discernibility and indiscernibility, Information Sciences 177 (2007) 4959–4976.