Information Sciences 228 (2013) 175–191

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Related family: A new method for attribute reduction of covering information systems

Tian Yang [a], Qingguo Li [a,*], Bilei Zhou [b]

[a] College of Mathematics and Econometrics, Hunan University, Changsha 410082, China
[b] Business School, Central South University, Changsha 410083, China

ARTICLE INFO

Article history:
Received 7 December 2009
Received in revised form 13 October 2011
Accepted 10 November 2012
Available online 12 December 2012

Keywords:
Rough set
Related family
Attribute reduct
Covering
Granular computing
Feature selection

ABSTRACT

In terms of attribute reduction of covering based rough sets, the discernibility matrix is used as a conventional method to compute all attribute reducts. However, it is inapplicable to attribute reduction in certain circumstances. In this article, a new method, referred to as the related family, is introduced to compute all attribute reducts and relative attribute reducts for covering rough sets. Its core idea is to remove superfluous attributes while keeping the approximation space of covering information system unchanged. The related family method is more powerful than the discernibility matrix method, since the former can handle complicated cases that could not be handled by the latter. In addition, a simplified version of the related family and its corresponding heuristic algorithm are also presented.

1. Introduction

Pioneered by Pawlak [24,25] in 1982, rough set theory has become a powerful mathematical tool for dealing with the vagueness and uncertainty inherent in various practical problems [9,23,26–28,48]. Rough set theory is also a substantial constituent of granular computing [45]. Among its various applications, data mining caught most attentions. Attribute reduction of rough sets can be used to conduct data processing without any prior knowledge, which is the most significant advantage of rough set theory over other approaches. Arguably, attribute reduction is the most important research problem in rough set theory.

Covering generalized rough sets [2–6,10,8,11,13,14,17,18,30,33,34,39,44,46,47,49–51] and binary relation generalized rough sets [7,15,19,31,35,41–43] are two main extensions of the Pawlak's rough set. The existing studies on covering rough sets mainly focus on set approximations. There are currently seven types of covering rough approximations that can be classified into three categories [39]. Bonikowski et al. [3] first extended the Pawlak's rough set theory from a partition to a covering. The second type of covering rough set model was presented by Pomykala [29], followed by Tsang et al. [33] with the development of the third type. Zhu defined the fourth and fifth types of covering-based approximations [51,52], while Xu and Wang [38] provided the definition for the sixth type. The seventh type of approximation operations was defined in [37]. Unfortunately, despite all the defined types of covering rough approximations, research on their corresponding attribute reduction algorithms is still stagnant and unsystematic to date.

Two types of reduction should be done on covering rough sets. One is granular reduction, reducing redundant elements from a covering; and the other is attribute reduction including relative attribute reduction as a special case that reduces redundant coverings from a family. Zhu et al. [49,50] proposed a granular reduction method for four types of set approximations. Yang and

* Corresponding author. Tel.: +86 073188822755.
E-mail addresses: math_yangtian@126.com (T. Yang), liqingguo1i@yahoo.com.cn (Q. Li), billyzhou@msn.com (B. Zhou).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2012.11.005

Li [39] studied granular reductions and presented their corresponding algorithms. In [5,6,33], the third upper approximation operation was regarded as the most reasonable set approximations among those in [3,22], i.e. from the first to fourth types of approximations sorted in [39]. Besides, the fifth type of approximation is regarded to be more accurate than the third one, though its advantage is weaken due to the loss of information. Accordingly, Tsang et al. [33] developed an algorithm by employing discernibility matrices to compute all attribute reducts for the fifth type of covering rough sets. In their approach, the fifth type of covering rough sets is mistreated as a special form of the third type. However, we found that the reduction theories for these two types are not identical. Moreover, the fifth type is not applicable in certain circumstances, as shown Section 2 in this paper. Therefore, the attribute reduction algorithm in [33] cannot work for the third type, and thus practical applications of the third type of approximations are limited. Apparently, it is necessary and essential to develop a corresponding attribute reduction theory for the third type of rough set model. In this paper, we focus on the attribute reduction of the third type of rough set model.

As shown below, the discernibility matrix is not suitable for the third type of covering set approximations. For the purpose of reducing the more complex covering information systems, we propose the related family to replace discernibility matrices in computing all attribute reducts. we show that the related family is more effective than the discernibility matrix in reducing covering generalized rough sets. Moreover, the simplified version of the related family, i.e, the minimal set family, is also proposed to compute all attribute reducts, which significantly speed up the computation process. It has been proved that finding the set of all reducts or finding an optimal reduct (i.e, a reduct with the minimum number of attributes) is NP [36]. Thus, we also develop a heuristic algorithm accordingly.

The remainder of this paper is organized as follows. Section 2 introduces relevant background knowledge and existing gaps in the literature on the third type of covering rough set model. Section 3 describes the introduced related family method to bridge the gaps existing in applications of the discernibility matrix. Sections 4 and 5 present the design of the relative attribute reduction algorithms for consistent and inconsistent covering decision systems, respectively. Section 6 simplifies the processes for the related family. Section 7 shows experimental analysis results. Finally, Discussion and conclusion remarks are presented in Section 8. It is noteworthy that the concept of related family (described in Section 3) was originally proposed in one of our previous conference papers [40].

2. Background knowledge and existing gaps

Let U be a finite and nonempty set, and R be an equivalence relation on U. R generates a partition $U/R = \{[x_R|x\in X\}$ on U, where x_R is an equivalence class of x generated by the equivalence relation R. We call it the elementary sets of R in rough set theory. For any set X, we describe it by the elementary sets of R and the two sets

$$R_*(X) = \cup\{[x]_R | [x]_R \subseteq X\}, R^*(X) = \cup\{[x]_R | [x]_R \cap X \neq \emptyset\}$$

are called the lower and upper approximations of X, respectively. If $R_*(X) = R^*(X)$, X is an R-exact set. Otherwise, it is an R-rough set.

Let $\mathbb{R}$ be a family of equivalence relations, then $(U, \mathbb{R})$ is called a knowledge base (or an information system). Let $A \in \mathbb{R}$, denoted as IND $(\mathbb{R}) = \cap\{R : R \in \mathbb{R}\}$. A is dispensable in $\mathbb{R}$ if and only if IND $(\mathbb{R})$ =IND $(\mathbb{R} - A)$. Otherwise A is indispensable in $\mathbb{R}$. The family $\mathbb{R}$ is independent if every $A \in \mathbb{R}$ is indispensable in $\mathbb{R}$. Otherwise $\mathbb{R}$ is dependent. $\mathbb{Q} \subseteq \mathbb{P}$ is a reduct of $\mathbb{P}$ if $\mathbb{Q}$ is independent and IND $(\mathbb{Q})$ =IND $(\mathbb{P})$. The set of all indispensable relations in $\mathbb{P}$ is called the core of $\mathbb{P}$, which is denoted as CORE $(\mathbb{P})$. Evidently, CORE $(\mathbb{P}) = \cap$RED$(\mathbb{P})$, where RED$(\mathbb{P})$ is the family of all reducts of $\mathbb{P}$.

If $\mathbb{R} = \mathbb{A} \cup D, \forall A \in \mathbb{A}$ is called a conditional attribute, D is called the decision attribute, and then $(U, \mathbb{A} \cup D)$ is called a decision system, denoted by $(U, \mathbb{A}, D)$. $POS_{\mathbb{A}}(D) = \cup_{X \in U/D} \mathbb{A}_*(X)$ is called the positive region of D relative to $\mathbb{A}$. A is dispensable in $\mathbb{A}$ relative to D if and only if $POS_{\mathbb{A}}(D) = POS_{\mathbb{A}-A}(D)$. Otherwise A is indispensable in $\mathbb{A}$ relative to D. The family $\mathbb{A}$ is independent if every $A \in \mathbb{A}$ is indispensable in $\mathbb{A}$ relative to D. Otherwise $\mathbb{A}$ is dependent. $\mathbb{Q} \subseteq \mathbb{P}$ is a relative reduct of $\mathbb{P}$ if $\mathbb{Q}$ is independent and $POS_{\mathbb{A}}(D) = POS_{\mathbb{A}-A}(D)$. The set of all indispensable relations in $\mathbb{P}$ is called the relative core of $\mathbb{P}$, which is denoted as CORE $(\mathbb{P})$. Evidently, CORE $(\mathbb{P}) = \cap$RED$(\mathbb{P})$, where RED$(\mathbb{P})$ is the family of all relative reducts of $\mathbb{P}$.

The discernibility matrix method is proposed to compute all reducts of information systems and the relative reducts of decision systems in [32].

$\mathcal{C}$ is called a covering of U, where U is a nonempty domain of discourse, if and only if $\mathcal{C}$ is a family of nonempty subsets of U and $\cup\mathcal{C} = U$.

It is clear that a partition of U is certainly a covering of U, so the concept of a covering is an extension of the concept of a partition.

Definition 2.1 [3]. Minimal description. Let $\mathcal{C}$ be a covering of U, $Md_{\mathcal{C}}(x) = \{K \in \mathcal{C} | x \in K \wedge (\forall S \in \mathcal{C} \wedge x \in S \wedge S \subseteq K \Rightarrow K = S)\}$ is called the minimal description of x. When there is no confusion, we omit $\mathcal{C}$ from the subscript.

Definition 2.2 ([3,51]. Neighborhood). Let $\mathcal{C}$ be a covering of U, $N_{\mathcal{C}}(x) = \cap\{C \in \mathcal{C} | x \in C\}$ is called the neighborhood of x. Generally, we omit the subscript $\mathcal{C}$ when there is no confusion.

Since it shows that $N(x) = \cap\{C \in \mathcal{C} | x \in C\} = \cap Md(x)$ in [3], the neighborhood of x can be seen as the minimum description of x and it is the most precise description (referred to [3] for more details).

Definition 2.3 [33]. Covering lower and upper approximation operations. Let $\mathcal{C}$ be a covering of U. The operations $CL_{\mathcal{C}}, CH_{\mathcal{C}} : P(U) \to P(U)$ are defined as follows: $\forall\ X \in P(U)$

$CL_{\mathcal{C}}(X) = \cup\{K \in \mathcal{C} | K \subseteq X\} = \cup\{K | \exists x$, s.t. $(K \in Md(x)) \wedge (K \subseteq X)\}$,

$CH_{\mathcal{C}}(X) = \cup\{K \in Md(x) | x \in X\}$we call $CL_{\mathcal{C}}$ the covering lower approximation operations and $CH_{\mathcal{C}}$ the covering upper approximation operations, with respect to the covering $\mathcal{C}$. We leave out $\mathcal{C}$ from the subscript when there is no confusion.

We review the reduction theory introduced in [49], referred to as the union reduction theory in this paper. Accordingly, a reducible element is called a union reducible element. Other concepts are labeled similarly, such as the union reducible covering and union reduct. The detailed descriptions and proofs can be found in [49,51].

Definition 2.4 [49]. Union reducible covering. Let $\mathcal{C}$ be a covering of U and $C \in \mathcal{C}$. If C is a union of some sets in $\mathcal{C} - \{C\}$, we say that C is a union reducible element in $\mathcal{C}$. Otherwise, C is a union irreducible element in $\mathcal{C}$. If every element in $\mathcal{C}$ is union irreducible, we say that $\mathcal{C}$ is union irreducible; otherwise $\mathcal{C}$ is a union reducible.

Definition 2.5 [49]. Union reduct. For a covering $\mathcal{C}$ of U, the new union irreducible covering through the above reduction is called the union reduct of $\mathcal{C}$ and denoted by $ured(\mathcal{C})$.

A reducible element is merely union reducible, as shown in [49]. However, union reducible elements and union reducts are special cases of those in [39].

Definition 2.6 [50]. Covering lower and upper approximation operations. Let $\mathcal{C}$ be a covering of U. The operations $NL_{\mathcal{C}}, NH_{\mathcal{C}} : P(U) \to P(U)$ are defined as follows: $\forall\ X \in P(U)$

$$NL_{\mathcal{C}}(X) = \{x | N(x) \subseteq X\} = \cup\{N(x) | N(x) \subseteq X\},$$

$$NH_{\mathcal{C}}(X) = \cup\{N(x) | x \in X\}$$

we call $NL_{\mathcal{C}}$ the covering lower approximation operations and $NH_{\mathcal{C}}$ the covering upper approximation operations, with respect to the covering $\mathcal{C}$. We leave out $\mathcal{C}$ from the subscript when there is no confusion.

The approximations in Definition 2.3 are usually referred to as the third type of covering approximations, while the approximations in Definition 2.6 are called the fifth type of covering approximations. Suppose Δ is a family of coverings on U, then $\cup\Delta$ is also a covering on U. In [5,6,33], the third type of upper approximation proves to be a more reasonable set approximation than those in [3,22]. Tsang et al. [33] designed a reduction algorithm for the third type of covering rough sets. However, this algorithm [33] does not correspond to the third one as shown in the following example.

Example 2.1. Let $U = \{a,b,c,d\}$ be a universe, $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4$ be coverings on U, and $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4\}$, where

$$\mathcal{C}_1 = \{\{a\}, \{b, c, d\}\},$$

$$\mathcal{C}_2 = \{\{a, b\}, \{c, d\}\},$$

$$\mathcal{C}_3 = \{\{a, b, c\}, \{d\}\},$$

$$\mathcal{C}_4 = \{\{b\}, \{b, d\}, \{a, c\}\}$$

Since $Cov(\Delta - \{\mathcal{C}_4\}) = Cov(\Delta)$ and $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ are indispensable in $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$ (refer to [33] for details), it is easy to see that $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$ is a reduct of $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4\}$ according to [33].

Since $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$ is a reduct of $\Delta, \Delta - \{\mathcal{C}_4\}$ contains the same information as Δ. However, for $b \in U$, $Md_{\cup\Delta}(b) = \{\{b\}\}$, while $Md_{\cup(\Delta - \{\mathcal{C}_4\})}(b) = \{\{a, b\}, \{b, c, d\}\}$; for $c \in U$, $Md_{\cup\Delta}(c) = \{\{a,c\}, \{c,d\}\}$, while $Md_{\cup(\Delta - \{\mathcal{C}_4\})}(c) = \{\{a, b, c\}, \{c, d\}\}$.

These differences generate the discrete approximations.

Let $A = \{b\}$. $CH_{\cup\Delta}(A) = \{b\}$, while $CH_{\cup(\Delta - \{\mathcal{C}_4\})}(A) = \{\{a, b, c, d\}\}$.

Let $B = \{c\}$. $CH_{\cup\Delta}(B) = \{a, c, d\}$, while $CH_{\cup(\Delta - \{\mathcal{C}_4\})}(B) = \{a, b, c, d\}$.

That is, for subsets $A \subseteq U$ and $B \subseteq U, CH_{\cup\Delta}(A) \neq CH_{\cup(\Delta - \{\mathcal{C}_4\})}(A)$ and $CH_{\cup\Delta}(B) \neq CH_{\cup(\Delta - \{\mathcal{C}_4\})}(B)$. We can conclude that, the attribute reduction method is not fit for the approximation in Definition 2.3.

We can easily find that, the algorithm in [33] is designed for the fifth type. Therefore, the attribute reduction theory in [33] is still unavailable for the third one. As a result, practical applications of the third type of approximations are limited. Therefore, it is necessary and essential to build a corresponding attribute reduction theory for the reasonable approximations in Definition 2.3. Although the fifth type is regarded as being more accurate than the pair in Definition 2.3, its advantage is weakened by the loss of information. Moreover, the fifth type is not applicable in certain circumstances, as shown in the following example.

Example 2.2. Table 1 represents an incomplete information system [16].

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ is a set of six houses,

Table 1
A house information system.

Houses	Color	Surrounding	Structure	Price
X_1	Good	Good	Bad	High
X_2	$*$	Good	Bad	High
X_3	Bad	Good	Good	$*$
X_4	Bad	$*$	Good	Low
X_5	Bad	Bad	$*$	Low
X_6	$*$	Good	$*$	$*$

$A = \{color, surrounding, structure, price\}$ is a set of attributes, where the values of "*color*", "*surrounding*", and "*structure*" are $\{good, bad\}$, and the values of "*price*" are $\{high, low\}$.

It may happen that some attribute values for an object are missing. To indicate such a situation, a distinguished value, the so-called null value, is usually assigned to these attributes. We denote all null values by $*$.

Let $SIM(A) = \{(x,y) \in U \times U - \forall\ a \in A,\ a(x) = a(y), \text{ or } a(x) = *, \text{ or } a(y) = *\}$. $SIM(A) = \cap_{a \in A} SIM(\{a\})$, is a tolerance relation.

Let $SIM(A)(x) = \{y|(x,y) \in SIM(A)\}$, then $\mathcal{C} = \{SIM(A)(x_i)|x_i \in U\} = \{\{x_1,x_2,x_6\},\{x_3,x_4,x_6\},\{x_3,x_4,x_5,x_6\},\{x_4,x_5\},\{x_1,x_2,x_3,x_4,x_6\}\}$ is a covering on U. We obtain neighborhoods easily.

$N(x_1) = \{x_1,x_2,x_6\}$, $N(x_2) = \{x_1,x_2,x_6\}$, $N(x_3) = \{x_3,x_4,x_6\}$, $N(x_4) = \{x_4\}$, $N(x_5) = \{x_4,x_5\}$, $N(x_6) = \{x_6\}$.

In this example, due to a loss of information, x_6 belongs simultaneously to several elements of the covering $\mathcal{C}$ as a potential equivalent element of others. When taking the intersection of these covering elements containing x_6, x_6 is retained as the component of some neighborhoods. However, other elements with specific information are ignored. Moreover, the more information an object loses, the more likely it is to be retained in the neighborhood structure, as shown in the above example. Neighborhoods are the approximation elements of the approximation operations in Definition 2.6, as the core knowledge carrying information regarding various covering elements. Apparently, we are unable to obtain complete information in knowledge discovery by employing the core knowledge of those elements that have lost information. Therefore, the approximation operations in Definition 2.6 are not applicable to incomplete information systems. Based on this motivation, the subsequent sections examine the attribute reduction of the approximations in Definition 2.3.

3. Attribute reduction of covering information systems

In this section, the attribute reduction theory of the third type of covering rough sets (defined in Definition 2.3) is extended. We propose the notions for attribute reduction with the third type rough sets. Then, the limitation of the discernibility matrix is pointed out, that is, it is not applicable to the third type of rough set model. Accordingly, the related family method is proposed to bridge the gap.

3.1. The definition of attribute reduction with covering rough sets

Suppose $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ is a family of coverings of U, (U, Δ) is called a covering information system. The target of an attribute reduction of a covering information system is to find the minimal subset $\mathbb{P}$ of Δ such that the approximations of any $X \subseteq U$ are invariant.

Let $\mathcal{M}_{\cup\Delta} = \cup\{Md_{\cup\Delta}(x)|x \in U\}$. Since $CH_{\cup\Delta}(X) = \cup\{K \in Md_{\cup\Delta}(x)|x \in X\}$ and $CL_{\cup\Delta}(X) = \cup\{C|C \subseteq X, C \in \mathcal{M}_{\cup\Delta}\}$, $\mathcal{M}_{\cup\Delta}$ is the collection of all approximation elements of $CH_{\cup\Delta}$ and $CL_{\cup\Delta}$. Thus, $CH_{\cup\Delta}$ and $CL_{\cup\Delta}$ are determined by $\mathcal{M}_{\cup\Delta}$, given a subset X of U. That is, $CH_{\cup\Delta}$ and $CL_{\cup\Delta}$ are invariant if and only if $\mathcal{M}_{\cup\Delta}$ is invariant.

In short, the target of attribute reduction is to find the minimal subsets $\mathbb{P}$ of Δ such that $\mathcal{M}_{\cup\Delta}$ is variant.

Definition 3.1 [40]. Let $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ be a family of coverings on U, $\mathcal{C}_i \in \Delta$. If $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$, we say $\mathcal{C}_i$ is dispensable in Δ. Otherwise, $\mathcal{C}_i$ is indispensable in Δ. For every $\mathbb{P} \subseteq \Delta$, if $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\mathbb{P}}$ and every covering in $\mathbb{P}$ is indispensable, we say $\mathbb{P}$ is a reduct of Δ. The collection of all indispensable coverings in Δ is denoted by $CORE(\Delta)$. The collection of all reducts of Δ is denoted by $RED(\Delta)$.

Proposition 3.1 [49]. *Let $\mathcal{C}$ be a covering of U, and K be a union reducible element of $\mathcal{C}$. Then, $\mathcal{C} - \{K\}$ and $\mathcal{C}$ have the same $Md(x)$ for all $x \in U$.*

Proposition 3.2 [53]. *Let $\mathcal{C}$ be a covering of U. $K \in \mathcal{C}$ is union reducible if and only if for any $x \in U$, $K \notin Md(x)$.*

Proposition 3.3 [40]. $CORE(\Delta) = \cap RED(\Delta)$.

Proof. For any $\mathcal{C} \in CORE(\Delta)$, and any $\Delta' \in RED(\Delta)$, we have $\mathcal{M}_{\cup\Delta} \neq \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$ and $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\Delta'}$. Suppose $\mathcal{C} \notin \Delta'$, then $\Delta' \subseteq (\Delta - \{\mathcal{C}\})$. Since $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\Delta'}, \mathcal{M}_{\cup\Delta} \subseteq \cup\Delta'$. For $\cup\Delta' \subseteq \cup(\Delta - \{\mathcal{C}\}), \mathcal{M}_{\cup\Delta} \subseteq \cup(\Delta - \{\mathcal{C}\})$. According to Proposition 3.2, it is evident that, for $K \in \cup\Delta$, K is irreducible if and only if $K \in \mathcal{M}_{\cup\Delta}$. Therefore, every element in $\mathcal{C}$ is reducible in $\cup\Delta$. According to Proposition 3.1, we get that $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$, which is a contradiction. Thus, $\mathcal{C} \in \Delta'$, that is $\mathcal{C} \in \cap RED(\Delta)$. Therefore, $CORE(\Delta) \subseteq \cap RED(\Delta)$.

If $\mathcal{C} \in \cap RED(\Delta)$, suppose $\mathcal{C} \notin CORE(\Delta)$, then $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$. Obviously, there is $\mathbb{P} \subseteq \Delta - \{\mathcal{C}\}$ such that $\mathbb{P} \in RED(\Delta)$, and then $\mathcal{C} \notin \mathbb{P} \in RED(\Delta)$. It is a contradiction. Therefore, $\mathcal{C} \in CORE(\Delta)$ and $\cap RED(\Delta) \subseteq CORE(\Delta)$. □

Proposition 3.4 [40]. *Let $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ be a family of coverings on U. If $\mathbb{P}$ is a minimal subset of Δ such that $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\mathbb{P}}$, then $\mathbb{P} \in RED(\Delta)$.*

Proof. For any $\mathcal{C} \in \mathbb{P}$, $\mathbb{P} - \{C\} \notin RED(\Delta)$. That is, $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\mathbb{P}} \neq \mathcal{M}_{\cup(\mathbb{P}-\{\mathcal{C}\})}$. Thus, any $\mathcal{C} \in \mathbb{P}$ is indispensable in $\mathbb{P}$, and then $\mathbb{P} \in RED(\Delta)$. □

3.2. The limitation of the discernibility matrix

Next we examine the attribute reduction algorithm of the third type of covering rough sets. In Pawlak's rough set theory, for every pair of $x, y \in U$, if y belongs to the equivalence class that contains x, we say x and y are indiscernible. Otherwise, they are discernible. Let $\mathbb{R} = \{R_1, R_2, \ldots, R_n\}$ be a family of equivalence relation on $U, R_i \in \mathbb{R}$. R_i is indispensable in $\mathbb{R}$ if and only if there is a pair of $x, y \in U$ such that the relation between x and y is altered after deleting R_i from $\mathbb{R}$. The attribute reduction of Pawlak's rough sets is to find minimal subsets of $\mathbb{R}$ that keep the relations invariant for any $x, y \in U$. Based on this statement, a method for computing all reducts of Pawlak's rough sets using discernibility matrix was proposed in [32]. However, the statement does not hold in covering rough set theory, since it is very different and more complex for covering rough sets. Let $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ be a family of coverings of U, $Md_{\cup\Delta}(x) = \{K \in \cup\Delta | x \in K \wedge (\forall\ S \in \cup\Delta \wedge\ x \in S \wedge S \subseteq K \Rightarrow K = S)\}$. For the third type of covering rough sets, minimal descriptions are the approximation elements of the operations, and they carry the information of the elements in U. Thus, for any $x, y \in U$, we examine the relation between x and y from the inclusion relation between $Md_{\cup\Delta}(x)$ and $Md_{\cup\Delta}(y)$.

The following example shows that all relations may be invariant even if we delete an indispensable element from Δ.

Example 3.1 [40]. Let $U = \{a, b, c, d\}, \Delta = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$,

$$\mathcal{C}_1 = \{\{a\}, \{a, b\}, \{b, c, d\}\},$$

$$\mathcal{C}_2 = \{\{a\}, \{c\}, \{a, d\}, \{b, c, d\}\},$$

$$\mathcal{C}_3 = \{\{a\}, \{b, c\}, \{b, d\}, \{c, d\}\},$$

then $\mathcal{M}_{\cup\Delta} = \{\{a\}, \{a, b\}, \{a, d\}, \{b, c\}, \{c\}, \{b, d\}, \{c, d\}\}$.

If $\mathcal{C}_3$ is deleted, $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}_3\})} = \{\{a\}, \{a, b\}, \{a, d\}, \{c\}, \{b, c, d\}\}$, $\mathcal{C}_3$ is indispensable in Δ. However, the binary relation among elements in U is invariant as shown below. $Md(x)$ denotes the minimal description of $x \in U$ with respect to Δ, while $Md'(x)$ denotes the minimal description of $x \in U$ with respect to $\Delta - \{C_3\}$.

$$\Delta : Md(a) = \{\{a\}\}, Md(b) = \{\{a, b\}, \{b, c\}, \{b, d\}\},$$

$$Md(c) = \{\{c\}\}, Md(d) = \{\{a, d\}, \{b, d\}, \{c, d\}\}.$$

$$\Delta - \{\mathcal{C}_3\} : Md'(a) = \{\{a\}\}, Md'(b) = \{\{a, b\}, \{b, c, d\}\},$$

$$Md'(c) = \{\{c\}\}, Md'(d) = \{\{a, d\}, \{b, c, d\}\}.$$

a and b in $\Delta : Md(b) \cap Md(a) = \emptyset$;
a and b in $\Delta - \{\mathcal{C}_3\} : Md'(b) \cap Md'(a) = \emptyset$;
a and c in $\Delta : Md(a) \cap Md(c) = \emptyset$;
a and c in $\Delta - \{\mathcal{C}_3\} : Md'(a) \cap Md'(c) = \emptyset$;
a and d in $\Delta : Md(a) \cap Md(d) = \emptyset$;
a and d in $\Delta - \{\mathcal{C}_3\} : Md'(a) \cap Md'(d) = \emptyset$;
b and c in $\Delta : Md(c) \cap Md(b) = \emptyset$;
b and c in $\Delta - \{\mathcal{C}_3\} : Md'(c) \cap Md'(b) = \emptyset$;
b and d in $\Delta : Md(d) \not\subseteq Md(b), Md(b) \not\subseteq Md(d)$ and $Md(b) \cap Md(d) \neq \emptyset$;
b and d in $\Delta - \{\mathcal{C}_3\} : Md'(d) \not\subseteq Md'(b), Md'(b) \not\subseteq Md'(d)$ and $Md'(b) \cap Md'(d) \neq \emptyset$;
c and d in $\Delta : Md(c) \cap Md(d) = \emptyset$;
c and d in $\Delta - \{\mathcal{C}_3\} : Md'(c) \cap Md'(d) = \emptyset$;

In short, any relation among U is invariant after deleting an indispensable covering from Δ.

As a result, discernibility matrices are not applicable to compute attribute reducts based on $\mathcal{M}$-approximation spaces. We therefore develop a new method to compute attribute reducts.

3.3. The related family

Definition 3.2 [40] Related family. Let $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ be a family of coverings of $U = \{x_1, x_2, \ldots, x_n\}$, and $\mathcal{M}_{\cup\Delta} = \{K_1, K_2, \ldots, K_m\}$. For any $K_i \in \mathcal{M}_{\cup\Delta}$, we define $r_i = \{\mathcal{C} | K_i \in \mathcal{C} \in \Delta\}$ is the related set of K_i, and $R(U,\Delta) = \{r_i | i = 1, 2, \ldots, m\}$ is the related family of (U,Δ).

Minimal descriptions can be obtained in linear time $O(n)$, whereas the time complexity of the related family is $O(m)$, where n is the number of samples and m is the cardinality of $\mathcal{M}_{\cup\Delta}$.

Definition 3.3 [40]. Let $\Delta = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ be a family of coverings of U, $R(U,\Delta) = \{r_i | i = 1, 2, \ldots, m\}$. A related function $f(U,\Delta)$ is a Boolean function of n boolean variables $\overline{\mathcal{C}_1}, \overline{\mathcal{C}_2}, \ldots, \overline{\mathcal{C}_n}$ corresponding to the coverings $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n$, respectively, which is defined as: $f(U,\Delta)(\overline{\mathcal{C}_1}, \overline{\mathcal{C}_2}, \ldots, \overline{\mathcal{C}_n}) = \wedge\{\vee(r_i | r_i \in R(U,\Delta)\}$.

Proposition 3.5 [40]. *$CORE(\Delta) = \{\mathcal{C} \in \Delta | r_i = \{\mathcal{C}\}$ for some $r_i \in R(U,\Delta)\}$*

Proof. Suppose $\mathcal{C} \in CORE(\Delta), \mathcal{C}$ is indispensable in Δ, then $\mathcal{M}_{\cup\Delta} \neq \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$, which implies there is $K_i \in \mathcal{M}_{\cup\Delta}$ such that $K_i \notin \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$.

Since $\mathcal{M}_{\cup\Delta} \subseteq \cup\Delta$ and $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})} \subseteq \cup(\Delta - \{\mathcal{C}\})$, $K_i \in \cup\Delta$ and $K_i \notin \cup(\Delta - \{\mathcal{C}\})$. (Or else $K_i \in \cup(\Delta - \{\mathcal{C}\})$, since $K_i \in \mathcal{M}_{\cup\Delta}$, there is no $\mathcal{B} \subseteq \cup\Delta - \{K_i\}$ such that $K_i = \cup\mathcal{B}$. That is, K_i is union irreducible in $\cup\Delta$, so K_i is also union irreducible in $\cup(\Delta - \{\mathcal{C}\})$. Thus, $K_i \in \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$, which is a contradiction.)

So $K_i \in \mathcal{C}$ and $K_i \notin \mathcal{C}_j$ for any $\mathcal{C}_j \in \Delta - \{\mathcal{C}\}$. Thus $r_i = \{\mathcal{C}\}$.

Suppose $r_{i0} = \{\mathcal{C}\}$ and $r_{i0} \in R(U,\Delta)$, then there is $K_{i0} \in \mathcal{M}_{\cup\Delta}$ such that r_{i0} is the related set of K_{i0}. In other words, $K_{i0} \in \mathcal{C}$ and $K_{i0} \notin \mathcal{C}' \in \Delta - \{\mathcal{C}\}$. If $\mathcal{C}$ is deleted from Δ, $K_{i0} \notin \mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})}$, then $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})} \neq \mathcal{M}_{\cup\Delta}$. Thus $\mathcal{C}$ is indispensable in Δ. That is, $\mathcal{C} \in CORE(\Delta)$. □

Proposition 3.6 [40]. *Suppose $\Delta' \subseteq \Delta$, then $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\Delta'}$, if and only if $\Delta' \cap r_i \neq \emptyset$ for every $r_i \in R(U,\Delta)$.*

Proof. Suppose $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\Delta'}$, then for any $K_i \in \mathcal{M}_{\cup\Delta}$, we have $K_i \in \mathcal{M}_{\cup\Delta'}$. It implies there is $\mathcal{C} \in \Delta'$ such that $K_i \in \mathcal{C}$. On the other hand, $K_i \in \mathcal{C}$ implies $\mathcal{C} \in r_i$. So $\mathcal{C} \in r_i \cap \Delta'$, that is $\Delta' \cap r_i \neq \emptyset$.

Suppose $\Delta' \cap r_i \neq \emptyset$ for every $r_i \in R(U,\Delta)$, then, for any $K_i \in \mathcal{M}_{\cup\Delta}$, $K_i \in \cup\Delta'$. That is, $\mathcal{M}_{\cup\Delta} \subseteq \cup\Delta' \subseteq \cup\Delta$. Thus $\mathcal{M}_{\cup\Delta} = \mathcal{M}_{\cup\Delta'}$. □

Corollary 3.1 [40]. *Let $\Delta' \subseteq \Delta$, then Δ' is a reduct of Δ, if and only if Δ' is the minimal subset of Δ such that $\Delta' \cap r_i \neq \emptyset$ for $r_i \in R(U,\Delta)$.*

Theorem 3.1 [40]. *Let Δ be a family of coverings on U, and $f(U,\Delta)$ be the related function. If $g(U,\Delta) = (\wedge\Delta_1) \vee \ldots \vee (\wedge\Delta_l)$ is the reduced disjunctive form transferred from $f(U,\Delta)$ via the laws of multiplication and absorption. That is, for any $\Delta_k \subseteq \Delta, k = 1, 2, \ldots, l$, there is no repeated element in Δ_k. Then $RED(\Delta) = \{\Delta_1, \Delta_2, \ldots, \Delta_l\}$.*

Proof. For every $k = 1, 2, \ldots, l, \wedge\Delta_k \leqslant \vee r_i$ for any $r_i \in R(U,\Delta)$, so $\Delta_k \cap r_i \neq \emptyset$. Let $\Delta'_k = \Delta_k - \{\mathcal{C}\}$ for any $\mathcal{C} \in \Delta_k$, then $g(U,\Delta) \precneqq \vee_{t=1}^{k-1}(\wedge\Delta_t) \vee (\wedge\Delta'_k) \vee (\vee_{t=k+1}^{l}(\wedge\Delta_t))$. If for every $r_i \in R(U,\Delta)$, we have $\Delta'_k \cap r_i \neq \emptyset$, then $\wedge\Delta'_k \leqslant \vee r_i$ for every $r_i \in R(U,\Delta)$. That is, $g(U,\Delta) \geqslant \vee_{t=1}^{k-1}(\wedge\Delta_t) \vee (\wedge\Delta'_k) \vee (\vee_{t=k+1}^{l}(\wedge\Delta_t)$, which is a contradiction. It implies there is $r_{i0} \in R(U,\Delta)$ such that $\Delta'_k \cap r_{i0} = \emptyset$. Thus, Δ_k is a reduct of Δ.

For any $X \in RED(\Delta)$, we have $X \cap r_i \neq \emptyset$ for every $r_i \in R(U,\Delta)$, so $f(U,\Delta) \wedge (\wedge X) = \wedge(\vee r_i) \wedge (\wedge X) = \wedge X$, which implies $\wedge X \leqslant f(U,\Delta) = g(U,\Delta)$. Suppose, for every $k = 1, 2, \ldots, l$, we have $\Delta_k - X \neq \emptyset$. Then, for every k, there is $\mathcal{C}_k \in \Delta_k - X$. By rewriting $g(U,\Delta) = (\vee_{k=1}^{l}\mathcal{C}_k) \wedge \Phi, \wedge X \leqslant \vee_{k=1}^{l}\mathcal{C}_k$. Thus, there is C_{k_0} such that $\wedge X \leqslant \mathcal{C}_{k_0}$, that is $\mathcal{C}_{k_0} \in X$, which is a contradiction. So $\Delta_{k_0} \subseteq X$ for some k_0, since both X and Δ_{k_0} are reducts, it is evident that $X = \Delta_{k_0}$. Consequently, $RED(\Delta) = \{\Delta_1, \Delta_2, \ldots, \Delta_l\}$. □

From the above theorem, we can compute all attribute reducts of Δ with related functions.

Example 3.2. Suppose $\{a, b, c, d, e, f\}$ are six interviewees for a job position. They are assessed from four characteristic attributes, E = {Education, Height, Weight, Ability}, whose values are:

Education: {High, ordinary, low};

Height: {Tall, Average, Short};

Weight: {Heavy, Average, Light};

Ability: {Very strong, Strong, Normal}.

A panel of specialists evaluate the attributes of these interviewees. We have four coverings to describe these six interviewees.

Education: $\mathcal{C}_1 = \{High = \{a,b,c\},\ Ordinary = \{c,d,e\},\ Low = \{a,e,f\}\}$,

Height: $\mathcal{C}_2 = \{Short = \{a,b\},\ Average = \{c,d,e\},\ Tall = \{e,f\}\}$,

Weight: $\mathcal{C}_3 = \{Light = \{a\},\ Average = \{b,c,d,e\},\ Heavy = \{f\}\}$,

Ability: $\mathcal{C}_4 = \{Normal = \{a,b,c\},\ Strong = \{a,e,f\},\ Verystrong = \{c,d,e,f\}\}$.

Let $\Delta = \{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3,\mathcal{C}_4\}$, then $\mathcal{M}_{\cup\Delta} = \{\{a\},\{a,b\},\{a,b,c\},\{c,d,e\},\{b,c,d,e\},\{e,f\},\{f\}\}$.

$r(\{a\}) = \{\mathcal{C}_3\}; r(\{a,b\}) = \{\mathcal{C}_2\};\quad r(\{a,b,c\}) = \{\mathcal{C}_1,\mathcal{C}_4\};\quad r(\{c,d,e\}) = \{\mathcal{C}_1,\mathcal{C}_2\};\quad r(\{b,c,d,e\}) = \{\mathcal{C}_3\}; r(\{e,f\}) = \{\mathcal{C}_2\}; r(\{f\}) = \{\mathcal{C}_3\}$.

$$R(U,\Delta) = \{\{\mathcal{C}_3\},\{\mathcal{C}_2\},\{\mathcal{C}_1,\mathcal{C}_4\},\{\mathcal{C}_1,\mathcal{C}_2\}\}$$

$$f(U,\Delta)(\overline{\mathcal{C}_1},\overline{\mathcal{C}_2},\overline{\mathcal{C}_3},\overline{\mathcal{C}_4}) = \wedge\{\vee(r_i|K_i \in \mathcal{M}_{\cup\Delta}\} = \mathcal{C}_3 \wedge \mathcal{C}_2 \wedge (\mathcal{C}_1 \vee \mathcal{C}_4) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2) = (\mathcal{C}_1 \wedge \mathcal{C}_2 \wedge \mathcal{C}_3) \vee (\mathcal{C}_2 \wedge \mathcal{C}_3 \wedge \mathcal{C}_4)$$

So $RED(\Delta) = \{\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3\},\{\mathcal{C}_2,\mathcal{C}_3,\mathcal{C}_4\}\}$, $CORE(\Delta) = \{\mathcal{C}_2,\mathcal{C}_3\}$.

If these six interviewees are the training samples, we have two different kinds of evaluation references for other input samples:

$\{Education, Height, Weight\}, \{Height, Weight, Ability\}, and \{Height, Weight\}$ are the key attributes for the evaluation.

4. The attribute reduction of consistent covering decision systems

Suppose $(U, \Delta \cup \{D\})$ is an information system, where $\Delta = \{\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_n\}$ (the conditional attribute set) is a family of coverings on U, and D (the decision attribute) is an equivalence relation on U, then $(U, \Delta \cup \{D\})$ is called a covering decision system, which is a special information system that contains a decision attribute, also denoted by (U,Δ,D). In an information system based on Pawlak's rough sets, the collections of objects with the same conditional attribute values do not overlap, so every conditional attribute can induce a partition. However, the attributes of decision systems induce coverings, so there are more complex cases than a partition. Covering decision systems can be divided into consistent and inconsistent ones. We begin with the definition of attribute reduction of consistent covering systems.

Definition 4.1. Let $\Delta = \{\mathcal{C}_i : i = i, 2, \ldots, m\}$ be a family of coverings on U, D is a decision attribute, U/D is a decision partition on U. If for $\forall\, x \in U$, $\exists D_j \in U/D$ and $\exists K \in Md(y)$ such that $x \in K \subseteq D_j$, where $y \in U$, then the decision system (U,Δ,D) is called a consistent covering decision system, which is denoted as $\mathcal{M}_{\cup\Delta} \ll U/D$. Otherwise, (U,Δ,D) is called an inconsistent covering decision system, denoted as $\mathcal{M}_{\cup\Delta} \not\ll U/D$. The positive region of D relative to Δ is defined as $POS_{\cup\Delta}(D) = \cup_{X\in\ U/D} CL_{\cup\Delta}(X)$.

For every $X \in U/D$, we consider every $CH_{\cup\Delta}(X) \Rightarrow X$ as a possible rule and every $CL_{\cup\Delta}(X) \Rightarrow X$ as a certain rule. Even if (U,Δ,D) is a consistent covering decision system, the rules extracted from (U,Δ,D) may be possible, which is different from Pawlak's rough sets. However, for a decision system, we usually only focus on the certain rules. For a consistent decision system (U,Δ,D), every object in U can be divided into a decision class by some certain rules as shown in the following proposition. Although there is a more strict consistent covering decision system and a corresponding more complex relative attribute reduction, we ignore it since this is not the main focus of this study.

Proposition 4.1. *$\mathcal{M}_{\cup\Delta} \ll U/D$ if and only if $POS_{\cup\Delta}(D) = U$.*

Proof. Suppose $\mathcal{M}_{\cup\Delta} \ll U/D$, obviously for $\forall\ x \in U$, $\exists D_j \in U/D$ and $\exists\ K \in Md(y)$ such that $x \in K \subseteq D_j$, where $y \in U$, that is $[x]_D = \cup\{K{:}K \in Md(y) \wedge y \in [x]_D \wedge K \subseteq [x]_D\}$. So $CL_{\cup\Delta}([x]_D) = [x]_D$ and $POS_{\cup\Delta}(D) = U$.

Suppose $POS_{\cup\Delta}(D) = U$, then for any $x \in U$, there is $K \in \mathcal{M}_{\cup\Delta}$ and $\exists D_j \in U/D$ such that $x \in K \subseteq D_j$, that is $\mathcal{M}_{\cup\Delta} \ll U/D$. $\square$

Next we define the relative attribute reduction of a consistent covering decision system.

Definition 4.2. Let $(U,\Delta,D = \{d\})$ be a consistent covering decision system. For $\mathcal{C}_i \in \Delta$, if $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})} \ll U/D$, then $\mathcal{C}_i$ is called superfluous relative to D in Δ. Otherwise $\mathcal{C}_i$ is called indispensable relative to D in Δ. For every $\mathbb{P} \subseteq \Delta$ satisfying $\mathcal{M}_{\cup\mathbb{P}} \ll U/D$, if every element of $\mathbb{P}$ is indispensable relative to D in $\mathbb{P}$, i.e. for every $\mathcal{C}_i \in \mathbb{P}, \mathcal{M}_{\mathbb{P}-\{\mathcal{C}_i\}} \not\ll U/D$, then $\mathbb{P}$ is called a reduct of Δ relative to D, or relative reduct in short. The collection of all indispensable elements relative to D in Δ is called the core of Δ relative to D, denoted as $CORE(\Delta)$.

The relative reduct of a consistent covering decision system is the minimal set of conditional coverings (attributes) that ensures that every object can be decided by some certain rules. For a single covering $\mathcal{C}_i$, we present an equivalence condition to judge whether it is indispensable.

Theorem 4.1. *Suppose $\mathcal{M}_{\cup\Delta} \ll U/D$, and $\mathcal{C}_i \in \Delta$, then $\mathcal{C}_i$ is indispensable relative to D in Δ, i.e. $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})} \not\ll U/D$ if and only if $\exists x \in U$ such that $x \notin \cup\{K : K \in \cup(\Delta - \{\mathcal{C}_i\})$ and $\exists X \in U/D$ s.t. $K \subseteq X\}$.*

Proof. $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})} \not\ll U/D$

$\Leftrightarrow POS_{\cup(\Delta-\{\mathcal{C}_i\})}(D) \neq U \Leftrightarrow$ There exists $x \in U$ such that $x \notin POS_{\cup(\Delta-\{\mathcal{C}_i\})}(D) \Leftrightarrow x \notin \cup\{K : K \in \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$ *and* $\exists X \in U/D$ *s.t.* $K \subseteq X\}$

Let $\mathbb{A} = \{K : K \in \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$ *and* $\exists X \in U/D$ *s.t.* $K \subseteq X\}$,

$\mathbb{B} = \{K : K \in \cup(\Delta - \{\mathcal{C}_i\})$ *and* $\exists X \in U/D$*s.t.* $K \subseteq X\}$, we prove that $\cup\mathbb{A} = \cup\mathbb{B}$. On the one hand, it is obvious that $\cup\mathbb{A} \subseteq \cup\mathbb{B}$. On the other hand, for any $C \in \mathbb{B} - \mathbb{A}$, we have $C \notin \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$ and $\exists X \in U/D$ such that $C \subseteq X$. From Proposition 3.2, C is union reducible in $\cup(\Delta - \{\mathcal{C}_i\})$. That is, there is $\mathcal{B} \subseteq \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$ such that $C = \cup\mathcal{B}$. Since $C \subseteq X$, for any $C\prime \in \mathcal{B}, C\prime \subseteq X$. So we have $C \subseteq \cup\mathbb{A}$. Thus, $\cup\mathbb{A} = \cup\mathbb{B}$. That is,

$x \notin \cup\{K : K \in \mathcal{M}_{\cup(\Delta-\{\mathcal{C}_i\})}$ *and* $\exists X \in U/D$*s.t.* $K \subseteq X\} \Leftrightarrow x \notin \cup\{K : K \in \cup(\Delta - \{\mathcal{C}_i\})$ *and* $\exists X \in U/D$*s.t.* $K \subseteq X\}$. □

It should be noted that $x \notin \cup\{K : K \in \cup(\Delta - \{\mathcal{C}_i\})$ *and* $\exists X \in U/D$ s.t. $K \subseteq X\}$ means that $(U, \Delta - \{\mathcal{C}_i\}, D)$ is an inconsistent decision system. Thus, the fact that $\mathcal{C}_i$ is indispensable implies that it is a key covering to ensure that (U,Δ,D) is a consistent decision system. Assuming $\mathbb{P} \subseteq \Delta$, that ensures $(U,\mathbb{P},D)$ is a consistent decision system, we need to ensure that, for $\forall x \in U$, x should be covered by the elements in $\{K : K \in \cup\mathbb{P}$ *and* $\exists X \in U/D$ s.t.$K \subseteq X\}$. Then the target of the attribute reduction based on covering decision systems is to find a minimal subset $\mathbb{P}$ of Δ, such that $\{K : K \in \cup\mathbb{P}$ *and* $\exists X \in U/D$ s.t. $K \subseteq X\}$ is a covering of U. Next we define the related family based on covering decision systems.

Definition 4.3. Let $(U,\Delta,D = \{d\})$ be a consistent covering decision system, where $U = \{x_1,x_2,\ldots,x_n\}$. For any $x_i \in U, r(x_i) = \{\mathcal{C} \in \Delta | \exists C_k \in \cup\Delta$ *and* $\exists X \in U/D$ s.t. $x_i \in C_k \in \mathcal{C}$ *and* $C_k \subseteq X\}$. Then $R(U,\Delta,D) = \{r(x_i)|x_i \in U\}$ is called the related family of (U,Δ,D).

In this method, we first obtain the set $\mathcal{A} = \{C_k \in \cup\Delta | \exists X \in U/D$ s.t. $C_k \subseteq X\}$. Then, $r(x_i) = \{\mathcal{C} \in \Delta | \exists C_k \in \mathcal{A}$ s.t. $x_i \in C_k \in \mathcal{C}\}$. The time complexity of the related family of a consistent covering decision system is $O(n)$, where n is the number of samples.

We propose judging whether a covering decision is consistent, by the components of the related family.

Proposition 4.2. *(U,Δ,D) is a consistent covering decision system if and only if there is no empty set in $R(U,\Delta,D)$.*

Proof. (U,Δ,D) is a consistent covering decision system. $\Leftrightarrow POS_{\cup\Delta}(D) = U$. $\Leftrightarrow$ For any $x \in U, r(x_i) = \{\mathcal{C} \in \Delta | \exists C_k \in \cup\Delta$ *and* $\exists X \in U/D$ s.t. $x_i \in C_k \in \mathcal{C}$ *and* $C_k \subseteq X\} \neq \emptyset$. $\Leftrightarrow$ There is no empty set in $R(U,\Delta,D)$. □

Theorem 4.2. *Let $(U,\Delta,D = \{d\})$ be a consistent covering decision system, $\mathbb{P} \subseteq \Delta$, then we have:*

(1). *$\mathcal{M}_{\cup\mathbb{P}} \ll U/D$ if and only if $P \cap r(x_i) \neq \emptyset$ for any $x_i \in U$.*
(2) *$CORE(\Delta) = \{\mathcal{C} \in \Delta | r(x_i) = \mathcal{C}$ for some $x_i \in U\}$.*

Proof

(1) Suppose $\mathcal{M}_{\cup\mathbb{P}} \ll U/D$, then for any $x_i \in U, x_i \in \cup\{K : K \in \cup\mathbb{P}$ *and* $\exists X \in U/D$*s.t.* $K \subseteq X\}$. So there is a $C \in \mathcal{M}_{\cup\mathbb{P}} \subseteq \cup\mathbb{P}$ such that $x_i \in C \subseteq [x_i]_D$. Then there is a $\mathcal{C} \in \mathbb{P}$ such that $C \in \mathcal{C}$. It is evident that $\mathcal{C} \in r(x_i)$, thus $\mathcal{C} \in r(x_i) \cap \mathbb{P}$. Suppose $\mathbb{P} \cap r(x_i) \neq \emptyset$, then there is $\mathcal{C} \in \mathbb{P}$ such that $\mathcal{C} \in r(x_i)$, that is, there is $C \in \mathcal{C}$ such that $x_i \in C \subseteq [x_i]_D$. That is $x_i \in \cup\{K : K \in \cup\mathbb{P}$ *and* $\exists X \in U/D$ *s.t.* $K \subseteq X\}$. Thus, $\mathcal{M}_{\cup\mathbb{P}} \ll U/D$.
(2) Suppose $\mathcal{C} \in CORE(\Delta)$, then $\mathcal{C}$ is indispensable in Δ, then $\mathcal{M}_{\cup(\Delta-\{\mathcal{C}\})} \not\ll U/D$. That is, there is $x_i \in U$ such that $x_i \notin \cup\{K : K \in \cup(\Delta - \{\mathcal{C}\})$ *and* $\exists X \in U/D$*s.t.* $K \subseteq X\}$. Thus, $r(x_i) = \{\mathcal{C}\}$. If $r(x_i) = \{\mathcal{C}\}$ for some $x_i \in U$, it is evident that $\mathcal{C} \in CORE(\Delta)$. □

Corollary 4.1. *Let $\mathbb{P} \subseteq \Delta$, then $\mathbb{P}$ is a relative reduct of Δ if and only if it is the minimal subset satisfying $\mathbb{P} \cap r(x_i) \neq \emptyset$, for any $x_i \in U$.*

Definition 4.4. Let (U,Δ,D) be a consistent covering decision system, where $\Delta\{\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m\}$, $U = \{x_1,x_2,\ldots,x_n\}$. A related function $f(U,\Delta,D)$ for (U,Δ,D) is a boolean function of m boolean variables $\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m$ corresponding to coverings of $\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m$, respectively, which is defined as $f(U,\Delta,D)(\overline{\mathcal{C}}_1,\overline{\mathcal{C}}_2,\ldots,\overline{\mathcal{C}}_m) = \wedge\{\vee(r(x_i))|x_i \in U\}$, where $\vee(r(x_i))$ is the disjunction of all elements in $r(x_i)$.

Theorem 4.3. *Let $(U,\Delta,D = \{d\})$ be a consistent decision system, $R(U,\Delta,D) = \{r(x_i)|x_i \in U\}$ be the related family of (U,Δ,D), $f(U,\Delta,D)$ is the related function of (U,Δ,D). If $g(U,\Delta,D) = \vee_{k=1}(\wedge\Delta_k)(\Delta_k \subseteq \Delta)$ is the reduced disjunctive form obtained from $f(U,\Delta,D)$ by using the multiplication and absorption laws. That is, every element in Δ_i appears only once, then $RED(\Delta,D) = \{\Delta_1,\ldots,\Delta_m\}$.*

Proof. The proof is similar to that of Theorem 3.1. □

Example 4.1. $(U, \Delta, D = \{d\})$ is a covering decision system, where $U = \{x_1, x_2, \ldots, x_9\}\Delta = \{\mathcal{C}_i : i = 1, 2, 3, 4, 5, 6\}$,

$$\mathcal{C}_1 = \{\{x_1, x_2\}, \{x_2, x_3, x_4\}, \{x_3\}, \{x_4\}, \{x_5, x_6\}, \{x_6, x_7, x_8\}, \{x_7, x_8, x_9\}\},$$

$$\mathcal{C}_2 = \{\{x_1, x_3, x_4\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_6\}, \{x_7, x_8, x_9\}\},$$

$$\mathcal{C}_3 = \{\{x_1\}, \{x_1, x_2, x_3\}, \{x_2, x_3\}, \{x_3, x_4, x_5, x_6\}, \{x_7, x_8\}, \{x_9\}\},$$

$$\mathcal{C}_4 = \{\{x_1, x_2, x_4\}, \{x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_6\}\{x_7, x_8, x_9\}\},$$

$$\mathcal{C}_5 = \{\{x_1\}, \{x_1, x_2, x_3\}, \{x_4\}, \{x_5\}, \{x_5, x_6\}, \{x_5, x_6, x_8\}, \{x_8\}, \{x_7, x_8, x_9\}\},$$

$$\mathcal{C}_6 = \{\{x_1, x_5, x_9\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_6\}, \{x_7, x_8\}\},$$

$$U/D = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8, x_9\}\}.$$

$\mathcal{M}_{\cup\Delta} = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7, x_8\}, \{x_8\}, \{x_9\}\}, POS_{\cup\Delta}(D) = U$, then $(U, \Delta, D = \{d\})$ is a consistent covering decision system.

$$r(x_1) = \{\mathcal{C}_1, \mathcal{C}_3, \mathcal{C}_5\},$$

$$r(x_2) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_3) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_4) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_5) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_5\},$$

$$r(x_6) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_7) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_8) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\},$$

$$r(x_9) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5\}.$$

The related family of $(U, \Delta, D = \{d\})$:

$$R(U, \Delta, D = \{d\}) = \{\{\mathcal{C}_1, \mathcal{C}_3, \mathcal{C}_5\}, \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\}, \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_5, \mathcal{C}_6\}, \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_5\}, \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5\}\}.$$

The related function of $(U, \Delta, D = \{d\})$:

$$\begin{aligned} f(U, \Delta, D = \{d\}) &= (\mathcal{C}_1 \vee \mathcal{C}_3 \vee \mathcal{C}_5) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_3 \vee \mathcal{C}_4 \vee \mathcal{C}_5 \vee \mathcal{C}_6) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_4 \vee \mathcal{C}_5 \vee \mathcal{C}_6) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_4 \vee \mathcal{C}_5) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_3 \\ &\quad \vee \mathcal{C}_4 \vee \mathcal{C}_5) \\ &= (\mathcal{C}_1 \vee \mathcal{C}_3 \vee \mathcal{C}_5) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_4 \vee \mathcal{C}_5) = \mathcal{C}_1 \vee \mathcal{C}_5 \vee (\mathcal{C}_2 \wedge \mathcal{C}_3) \vee (\mathcal{C}_3 \wedge \mathcal{C}_4). \end{aligned}$$

Thus, $RED(\Delta) = \{\{\mathcal{C}_1\}, \{\mathcal{C}_5\}, \{\mathcal{C}_2, \mathcal{C}_3\}, \{\mathcal{C}_3, \mathcal{C}_4\}\}, CORE(\Delta) = \emptyset$.

5. Attribute reduction of inconsistent covering decision systems

In practice, there are many inconsistent covering decision systems. In this section we propose relative attribute reduction for inconsistent covering decision systems. By the definition of inconsistent covering decision systems, we know that some objects cannot be sorted by a certain rule. Similar to the traditional rough sets, we define the attribute reduct as the minimal set of conditional attributes to keep the positive region of the decision attribute invariant. Thus, we have the following definition of attribute reduction. In what follows we always suppose U is a finite universe and $\Delta = \{\mathcal{C}_i : i = 1, 2, \ldots, m\}$ is a family of coverings of U. Additionally, we always suppose $POS_{\cup\Delta}(D) \neq \emptyset$ in this section.

Definition 5.1. Suppose U is a finite universe and $\Delta = \{\mathcal{C}_i : i = 1, 2, \ldots, m\}$ is a family of coverings of $U, \mathcal{C}_i \in \Delta, D$ is a decision attribute relative to Δ on U and $d:U \rightarrow V_d$ is the decision function V_d, defined as $d(x) = [x]_D, (U, \Delta, D)$ is an inconsistent covering decision system, i.e. $POS_{\cup\Delta}(D) \neq U$. If $POS_{\cup\Delta}(D) = POS_{\cup(\Delta - \{\mathcal{C}_i\})}(D)$, then $\mathcal{C}_i$ is superfluous relative to D in Δ. Otherwise $\mathcal{C}_i$ is indispensable relative to D in Δ. For every $\mathbb{P}\subseteq\Delta$, if every element in $\mathbb{P}$ is indispensable relative to D, and $POS_{\cup\mathbb{P}}(D)=POS_{\cup\Delta}(D)$, then $\mathbb{P}$ is a reduct of Δ relative to D, or relative reduct in short. The collection of all the indispensable elements relative to D in Δ is the core of Δ relative to D, denoted by $CORE(\Delta)$.

To keep $POS_{\cup\Delta}(D)$ invariant, we must ensure that, after reducing some coverings, $\{C\in\cup\Delta'|\ \exists X_i\in U/D\ s.t.\ C\subseteq X_i\}$ is still a covering of $POS_{\cup\Delta}(D)$.

Definition 5.2. Let $(U,\Delta,D=\{d\})$ be an inconsistent decision system, $\mathcal{A}=\{C_k\in\cup\Delta|\exists X_i\in U/D s.t. C_k\subseteq X_i\}$.

(1) $r(x_i)=\{\mathcal{C}\in\Delta|\exists C_k\in\mathcal{A}$ s.t. $x_i\in C_k\in\mathcal{C}\}$, for $x_i\in POS_{\cup\Delta}(D)$;

(2) $r(x_i)=\varnothing$, for $x_i\notin POS_{\cup\Delta}(D)$;

then $R(U,\Delta,D)=\{r(x_i)|x_i\in U\}$.

The time complexity of $POS_{\cup\Delta}(D)$ of a covering decision system is $O(N)$, where N is the number of conditional attributes. Moreover, the time complexity of the related family of an inconsistent covering decision system is $O(n')$, where n' is the cardinality of $POS_{\cup\Delta}(D)$.

Theorem 5.1. *Suppose $(U,\Delta,D=\{d\})$ is an inconsistent decision system, $\mathbb{P}\subseteq\Delta$, then*

(1) $POS_{\cup\Delta}(D)=POS_{\cup\mathbb{P}}(D)$ if and only if $\mathbb{P}\cap r(x_i)$ for any nonempty set $r(x_i$;

(2) $CORE(\Delta)=\{\mathcal{C}|\exists x_i\in POS_{\cup\Delta}(D) s.t. r(x_i)=\{\mathcal{C}\}\}$.

Proof. The process is similar to that of Theorem 4.2. □

Corollary 5.1. *Suppose $\mathbb{P}\subseteq\Delta$, then $\mathbb{P}$ is a relative reduct of Δ, if and only if it is a minimal set satisfying $\mathbb{P}\cap r(x_i)\neq\varnothing$ for any nonempty set $r(x_i)\in R(U,\Delta,D)$.*

Definition 5.3. Let (U,Δ,D) be an inconsistent covering decision system, where $\Delta=\{\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m\}$, and $U=\{x_1,x_2,\ldots,x_n\}$. A related function $f(U,\Delta,D)$ for (U,Δ,D) is a boolean function of m boolean variables $\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m$ corresponding to coverings of $\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_m$, respectively, which is defined as

$$f(U,\Delta,D)(\overline{\mathcal{C}}_1,\overline{\mathcal{C}}_2,\ldots,\overline{\mathcal{C}}_m)=\wedge\{\vee(r(x_i))|r(x_i)\neq\varnothing,r(x_i)\in R(U,\Delta,D)\}$$

where $\vee(r(x_i))$ is the disjunction of all elements in $r(x_i)$.

Theorem 5.2. *Let $(U,\Delta,D=\{d\})$ be an inconsistent decision system, $R(U,\Delta,D)=\{r(x_i)|x_i\in U\}$ is the related family of (U,Δ,D), $g(U,\Delta,D)$ is the related function. If $f(U,\Delta,D)=\vee_{k=1}(\wedge\Delta_k)(\Delta_k\subseteq\Delta)$ is obtained from $f(U,\Delta,D)$ by applying the multiplication and absorption laws as many times as possible such that every element in Δ_i appears only once, then the set $\{\Delta_k{:}k\leqslant l\}$ is the collection of all reducts of system (U,Δ,D), i.e. if $RED(\Delta,D)$ is the collection of all reducts of system (U,Δ,D), then $RED(\Delta,D)=\{\Delta_1,\ldots,\Delta_m\}$.*

Proof. The proof is similar to that of Theorem 3.1. □

As described above, the attribute reduction of a covering decision system using the related family is more convenient than using a discernibility matrix. Moreover, the related family is applicable to those circumstances in which a discernibility matrix is inapplicable.

Example 5.1. Suppose $(U,\Delta,D=\{d\})$ is a covering decision system, where $U=\{x_1,x_2,\ldots,x_9\}$, $\Delta=\{\mathcal{C}_i:i=1,2,3,4\}$,

$$\mathcal{C}_1=\{\{x_1,x_2,x_3,x_4\},\{x_3,x_6,x_7\},\{x_4,x_5\},\{x_4,x_5,x_6\},\{x_6\},\{x_7,x_8,x_9\}\},$$

$$\mathcal{C}_2=\{\{x_1\},\{x_2,x_3,x_4\},\{x_4,x_5\},\{x_4,x_5,x_6\},\{x_6,x_7,x_8,x_9\}\},$$

$$\mathcal{C}_3=\{\{x_1\},\{x_1,x_3,x_4\},\{x_2,x_3,x_4\},\{x_3,x_4,x_5,x_6,x_7\},\{x_7,x_8,x_9\}\},$$

$$\mathcal{C}_4=\{\{x_1,x_4,x_5\},\{x_2,x_3,x_4,x_5\},\{x_4,x_5,x_6\},\{x_7,x_8,x_9\}\}.$$

$$U/D=\{\{x_1,x_2,x_3\},\{x_4,x_5,x_6\},\{x_7,x_8,x_9\}\}.$$

Then $\mathcal{M}_{\cup\Delta}=\{\{x_1\},\{x_1,x_3,x_4\},\{x_2,x_3,x_4\},\{x_3,x_6,x_7\},\{,x_4,x_5\},\{x_6\},\{x_7,x_8,x_9\}\}$, $POS_{\cup\Delta}(D)=\{x_1,x_4,x_5,x_6,x_7,x_8,x_9\}$. It is evident that $(U,\Delta,D=\{d\})$ is an inconsistent covering decision system.

$$r(x_1)=\{\mathcal{C}_2,\mathcal{C}_3\},r(x_2)=\varnothing,r(x_3)=\varnothing,r(x_4)=\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_4\},r(x_5)=\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_4\},r(x_6)=\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_4\},r(x_7)=\{\mathcal{C}_1,\mathcal{C}_3,\mathcal{C}_4\},r(x_8)=\{\mathcal{C}_1,\mathcal{C}_3,\mathcal{C}_4\},r(x_9)=\{\mathcal{C}_1,\mathcal{C}_3,\mathcal{C}_4\}.$$

The related family of $(U, \varDelta, D = \{d\})$:

$$R(U, \varDelta, D = \{d\}) = \{\{\mathcal{C}_2, \mathcal{C}_3\}, \emptyset, \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_4\}, \{\mathcal{C}_1, \mathcal{C}_3, \mathcal{C}_4\}\}.$$

The related function of $(U, \varDelta, D = \{d\})$:

$$f(U, \varDelta, D = \{d\}) = (\mathcal{C}_2 \vee \mathcal{C}_3) \wedge (\mathcal{C}_1 \vee \mathcal{C}_2 \vee \mathcal{C}_4) \wedge (\mathcal{C}_1 \vee \mathcal{C}_3 \vee \mathcal{C}_4) = (\mathcal{C}_1 \wedge \mathcal{C}_2) \vee (\mathcal{C}_2 \wedge \mathcal{C}_3) \vee (\mathcal{C}_2 \wedge \mathcal{C}_4) \vee (\mathcal{C}_1 \wedge \mathcal{C}_3) \vee (\mathcal{C}_3 \wedge \mathcal{C}_4).$$

Thus, $RED(\varDelta) = \{\{\mathcal{C}_1, \mathcal{C}_2\}, \{\mathcal{C}_2, \mathcal{C}_3\}, \{\mathcal{C}_2, \mathcal{C}_4\}, \{\mathcal{C}_1, \mathcal{C}_3\}, \{\mathcal{C}_3, \mathcal{C}_4\}\}$,

$CORE(\varDelta) = \emptyset$.

Example 2.2 argues that the third type of covering rough set model is suit to incomplete information systems, unlike the fifth, sixth [38] or seventh types [37]. As shown in Examples 2.1 and 3.1, the discernibility matrices are not suit to the third type. As a result, algorithms based on discernibility matrices cannot be applied to reduce incomplete information systems. Moreover, algorithms based on discernibility matrices are not applicable to the first [3], second [29], third [33] or fourth types [52]. Consequently, the reduction algorithms based on discernibility matrices cannot be applied to instances in which the above four types are used, for example, incomplete information systems and the Great Wall security policy model in [20,21]. Fortunately, the related family method proposed in this paper can bridge this gap.

6. The simplification of a related family

For the purpose of finding the set of all attribute reducts, we propose a method of related family. However, this proves that finding the set of all reducts, or finding an optimal reduct (i.e, a reduct with the minimum number of attributes) is NP [36]. Thus, a simplification of the algorithms is most essential in practice. In this section, we simplify the related family.

Definition 6.1. Let $R(U, \varDelta) = \{r_i | i = 1, 2, \ldots, n\}$ be the related family of $(U, \varDelta)$. For any $r_i \in R(U, \varDelta)$, if there is $r_j \in R(U, \varDelta) - \{r_i\}$ such that $r_j \subseteq r_i$, let $r'_i = \emptyset$; otherwise, $r'_i = r_i$. Then we get a new related family $SIR(U, \varDelta) = \{r'_i | i = 1, 2, \ldots, n\}$, which is referred to as the simplified related family of $(U, \varDelta)$.

Here, $R(U, \varDelta) = \{r_i | i = 1, 2, \ldots, n\}$ denotes the related family of a covering information system or a covering decision system. If $\varDelta = \mathbb{A} \cup D$, where $\mathbb{A}$ denotes the covering attribute set and D denotes the decision attribute, $(U, \varDelta)$ can also denote a covering decision system. This suggests that the proposed simplified method is suited to all related families of covering information systems and covering decision systems.

Theorem 6.1. *Let $R(U, \varDelta) = \{r_i | i = 1, 2, \ldots, n\}$ be the related family of $(U, \varDelta)$, $SIR(U, \varDelta)$ be the simplified related family of $(U, \varDelta)$, $\varDelta' \subseteq \varDelta$. Then $\varDelta' \cap r_i \neq \emptyset$ for every nonempty $r_i \in R(U, \varDelta)$ if and only if $\varDelta' \cap r' \neq \emptyset$ for every nonempty set $r' \in SIR(U, \varDelta)$.*

Proof. If $\varDelta' \cap r_i \neq \emptyset$ for every nonempty element $r_i \in R(U, \varDelta)$, it is clear that $\varDelta' \cap r' \neq \emptyset$ for every $r' \in SIR(U, \varDelta)$.

Suppose $\varDelta' \cap r'_i \neq \emptyset$ for every nonempty set $r'_i \in SIR(U, \varDelta)$. For any nonempty element $r_j \in R(U, \varDelta)$, if there is a set $r_0 \in R(U, \varDelta) - \{r_j\}$ such that $r_0 \subseteq r_j$, and for any set $r_1 \in R(U, \varDelta) - \{r_j, r_0\}$, $r_1 \not\subseteq r_0$, then $r'_0 = r_0 \neq \emptyset$. Since $\varDelta' \cap r'_0 \neq \emptyset$, then $\varDelta' \cap r_0 \neq \emptyset$, thus $\varDelta' \cap r_j \neq \emptyset$. If $r_0 \not\subseteq r_j$ for any nonempty set $r_0 \in R(U, \varDelta) - \{r_j\}$, then $r'_j = r_j$. Since $\varDelta' \cap r'_j \neq \emptyset$, then $\varDelta' \cap r_j \neq \emptyset$. Thus, $\varDelta' \cap r_j \neq \emptyset$ for every nonempty set $r_j \in R(U, \varDelta)$. □

Theorem 6.1 ensures that the simplified related family keeps the same power as the raw one in computing all attribute reducts of a covering information system. Two corollaries are directly obtained as follows.

Corollary 6.1. *Suppose $\varDelta' \subseteq \varDelta$, then $\varDelta' \in RED(\varDelta)$ if and only if $\varDelta'$ is a minimal set satisfying $\varDelta' \cap r'_i \neq \emptyset$ for every $r'_i \neq \emptyset$ and $r'_i \in SIR(U, \varDelta)$.*

Corollary 6.2. *Let $f'(U, \varDelta)(\overline{\mathcal{C}_1}, \overline{\mathcal{C}_2}, \ldots, \overline{\mathcal{C}_n}) = \wedge\{\vee(r' | r_i \in SIR(U, \varDelta)\}$ be the related function induced by $SIR(U, \varDelta)$, and $g(U, \varDelta)$ be the reduced disjunctive form of $f(U, \varDelta)$ by applying the multiplication and absorption laws. If $g(U, \varDelta) = (\wedge\varDelta_1) \vee \cdots \vee (\wedge\varDelta l)$ where $\varDelta_k \subseteq \varDelta$, $k = 1, 2, \ldots, l$ and every element in $\varDelta k$ only appears once, then $RED(\varDelta) = \{\varDelta_1, \varDelta_2, \ldots, \varDelta_l\}$.*

Next we illustrate the irreducibility of the simplified related family. Proposition 6.1 suggests that all useless elements in $\varDelta$ are deleted from $R(U, \varDelta)$. That is, every element in $\cup\, SIR(U, \varDelta)$ may appear in a reduct.

Proposition 6.1. *$\cup SIR(U, \varDelta) = \cup\, RED(\varDelta)$.*

Proof. Let $SIR(U, \varDelta) = \{r'_i | i = 1, 2, \ldots, n\}$. Suppose $\mathcal{C} \in \cup\{r'_i | i = 1, 2, \ldots, n\}$, then there is $r'_i \in SIR(U, \varDelta)$ such that $\mathcal{C} \in r'_i$. For any $r'_j \in SIR(U, \varDelta), r'_j \neq \emptyset$, if $\mathcal{C} \in r'_j$, let $r^1_j = \{\mathcal{C}\}$. Otherwise, $r^1_j = \{\mathcal{C}_{ji}\}$, where $\mathcal{C}_{ji} \in r'_j$. Suppose $M_1(U, \varDelta) = \{r^1_i | i = 1, 2, \ldots, n\}$, it is easy to prove that $\cup M_1(U, \varDelta)$ is a reduct of $\varDelta$ by Corollary 6.1. Since $\mathcal{C} \in \cup\{r^1_i | i = 1, 2, \ldots, n\} \in RED(\varDelta)$, $\mathcal{C} \in \cup RED(\varDelta)$.

Suppose $\mathcal{C} \in \cup RED(\varDelta)$, then there is $\varDelta_k \in RED(\varDelta)$ such that $\mathcal{C} \in \varDelta_k$. From Corollary 6.1, we know that $\varDelta k$ is a minimal set satisfying $\varDelta_k \cap r'_i \neq \emptyset$ for every $r'_i \in SIR(U, \varDelta)$. So there is an $r'_i \in SIR(U, \varDelta)$ such that $\mathcal{C} \in r'_i$, or else $\mathcal{C}$ is redundant in $\mathcal{C}_k$. Thus, $\mathcal{C} \in \cup\{r'_i | r'_i \in SIR(U, \varDelta)\}$. □

In summary, $\cup\{r'_i | r'_i \in SIR(U, \varDelta)\} = \cup RED(\varDelta)$.

Theorem 6.2. *Let $SIR(U,\Delta)=\{r'_i|i=1,2,\ldots,n\}$ be the simplified related family of (U,Δ), then $SIR(U,\Delta)$ is the minimal set family to compute all attribute reducts of Δ. That is, for any set family $R^0(U,\Delta)=\{r^0_i|i=1,2,\ldots,n\}$, where $r^0_i\subseteq r'_i$, $R^0(U,\Delta)$, can compute all attribute reducts of Δ if and only if $r^0_i=r'_i$ for $i=1,2,\ldots,n$.*

Proof. If $r^0_i=r'_i$ for $i=1,2,\ldots,n$, then $R^0(U,\Delta)=SIR(U,\Delta)$. Therefore, $R^0(U,\Delta)$ can compute all attribute reducts of Δ.

Suppose there is a nonempty element $r'_i\in SIR(U,\Delta)$ and a nonempty element $r^0_i\in R^0(U,\Delta)$ such that $r^0_i\subset r'_i$. If $|r'_i|=1$ ($|.|$denotes the cardinality of a set), suppose $r'_i=\{\mathcal{C}\}$, then $r^0_i=\emptyset$ and $\mathcal{C}\in CORE(\Delta)$. From the definition of the simplified related family, we know $\mathcal{C}\notin r'_j$ for any r'_j, where $r'_j\in SIR(U,\Delta)-\{r'_i\}$, then $\mathcal{C}\notin r^0_i$ for any $r^0_i\in R^0(U,\Delta)$. So $R^0(U,\Delta)$ cannot compute any attribute reducts of Δ.

If $|r'_i|\geqslant 2$, since $r^0_i\subset r'_i$, then there is an $\mathcal{C}\in(r'_i-r^0_i)$. Let $r^1_i=\{\mathcal{C}\}$. For any nonempty element $r'_j\in SIR(U,\Delta)-\{r'_i\}$, if $\mathcal{C}\in r'_j$, let $r^1_j=\emptyset$. Otherwise, let $r^1_j=\{\mathcal{C}_0\}$ where $\mathcal{C}_0\in r'_j-r'_i$. Let $R^1(U,\Delta)=\{r^1_i|i=1,2,\ldots,n\}$ and $\Delta'=\cup\{r^1_i|i=1,2,\ldots,n\}$, it is easy to prove that $\Delta'\in RED(\Delta)$ by Corollary 6.1. However, $\Delta'\cap r^0_i=\emptyset$, that is, $R^0(U,\Delta)$ cannot compute all attribute reducts of Δ. Thus, if $R^0(U,\Delta)$ can compute all attribute reducts of Δ, then $r^0_i=r'_i$ for $i=1,2,\ldots,n$. □

From Proposition 6.1 and Theorem 6.2 we know that there are no redundant elements in $SIR(U,\Delta)$ and it is a minimal set family to compute all attribute reducts of a covering information system. Since it is based on the reduction algorithm, the simplified related family greatly shortens the time for reduction processes.

Unfortunately, it proves that finding the set of all reducts, or finding an optimal reduct (i.e. a reduct with the minimum number of attributes), is NP [36]. So it is necessary to develop heuristic algorithms. The following is a heuristic algorithm to obtain an attribute reduct from the related family.

Heuristic Algorithm

(Note: $|\cdot|$ denotes the cardinality of a set, while $||\cdot||$ denotes the number of times for a covering appeared in a related family)

Step1. Simplify the related family $R(U,\Delta)=\{r_i|i=1,2,\ldots,n\}$ to $SIR(U,\Delta)=\{r'_i|i=1,2,\ldots,n\}$.

Step2. If there is $\mathcal{C}\in\cup SIR(U,\Delta)$ such that $||\mathcal{C}||=max\{||\mathcal{C}_i||:\mathcal{C}_i\in\cup SIR(U,\Delta)\}$, let $A=\{r'_i\in SIR(U,\Delta):\mathcal{C}\in r'_i\}$. For one of any $r'_i\in A$, let $r^1_i=\{\mathcal{C}\}$; for any other $r'_j\in A, j\neq i$, let $r^1_j=\emptyset$. Then we get $R^1(U,\Delta)=\{r^1_i|i=1,2,\ldots,n\}$.

Step3. If $max\{||\mathcal{C}_i||:\mathcal{C}_i\in\cup R^1(U,\Delta)\}\geqslant 2$, then repeat Step 2 for $R^1(U,\Delta)$, until $max\{||\mathcal{C}_i||:\mathcal{C}_i\in\cup R^n(U,\Delta)\}=1$. We get $R^n(U,\Delta)$.

Step4. For any $r^n_i\in R^n(U,\Delta)$, if $|r^n_i|\geqslant 2$, let $r^{n+1}_i=\{\mathcal{C}'\}$ for any $\mathcal{C}'\in r^n_i$. Otherwise, $r^{n+1}_i=r^n_i\}$. We get $R^{n+1}(U,\Delta)$.

Step5. $\Delta'=\cup R^{n+1}(U,\Delta)$ is a reduct of Δ.

From the above propositions we know that, the simplified related family is the minimal related family that can generate the same reducts as the original ones. Thus, hereafter we only examine simplified related families rather than general related families. The following example is used to illustrate the idea:

Example 6.1. The related family of (U,Δ) in Example 3.2 is shown as follow: $R(U,\Delta)=\{\{\mathcal{C}_3\},\{\mathcal{C}_2\},\{\mathcal{C}_1,\mathcal{C}_4\},\{\mathcal{C}_1,\mathcal{C}_2\}\}$.

So the simplified related family of $(U,\mathcal{C})$ is:

$SIR(U,\Delta)=\{\{\mathcal{C}_3\},\{\mathcal{C}_2\},\{\mathcal{C}_1,\mathcal{C}_4\},\emptyset\}$.

(1) Compute reducts by the related function:

$f(U,\Delta)(\overline{\mathcal{C}_1},\overline{\mathcal{C}_2},\overline{\mathcal{C}_3},\overline{\mathcal{C}_4})=\wedge\{\vee(r'|r'\in SIR(U,\Delta)\text{ and }r'\neq\emptyset\}=\mathcal{C}_3\wedge\mathcal{C}_2\wedge(\mathcal{C}_1\vee\mathcal{C}_4)=(\mathcal{C}_3\wedge\mathcal{C}_2\wedge\mathcal{C}_1)\vee(\mathcal{C}_3\wedge\mathcal{C}_2\wedge\mathcal{C}_4)$.

So $RED(\Delta)=\{\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3\},\{\mathcal{C}_2,\mathcal{C}_3,\mathcal{C}_4\}\}$, $CORE(\Delta)=\{\mathcal{C}_2,\mathcal{C}_3\}$.

(2) Compute a reduct by the heuristic algorithm:

For $|r'_3|=|\{\mathcal{C}_1,\mathcal{C}_4\}|=2$, let $r^1_3=\{\mathcal{C}_1\}$, then we get $\{\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3\}\in RED(\Delta)$; let $r^1_3=\{\mathcal{C}_4\}$, then we get $\{\mathcal{C}_2,\mathcal{C}_3,\mathcal{C}_4\}\in RED(\Delta)$.

From this example, it is evident that the simplified related family can remarkably simplify the computing process, while the heuristic algorithm can easily obtain a reduct.

It must be noted that, the simplified process and the heuristic algorithm can also be used to simplify the related family of a covering decision system with similar processes.

7. Experimental analysis

In this section, an experimental analysis is conducted to test the efficiency of the related family method on attribute reduction of covering rough sets and for comparison with Pawlak's rough sets. Eight standard data sets were downloaded from the UCI Repository of Machine Learning Databases [1], as described in Table 2. The conditional attributes of Iono, Sonar, WDBC, WPBC, Wine and Ecoli are numerical, while the others are hybrid.

Since Pawlak's model can only deal with discrete attributes, we transformed the numerical attributes into discrete ones by three discretized approaches: equal-width, equal-frequency, and fuzzy c-means clustering (FCM) [15]. An algorithm based on dependency was used on the discretized data sets. The reduction results using the three methods (equal-width, equal-frequency, and FCM) in [12] were compared with our results.

Before the reduction, all the attributes were standardized in the interval $[0,1]$. The radial basis function (RBF) kernel in the support vector machine (SVM) learning algorithm was employed to validate the selected features. The number of selected

Table 2
Data description.

Data sets	Abbreviation	Samples	Features	Classes
Australian credit approval	Credit	690	15	2
Heart disease	Heart	270	13	2
Ecoli	Ecoli	336	7	7
Ionosphere	Iono	351	34	2
Sonar, mines vs. rocks	Sonar	208	60	2
Wisconsin diagnostic breast cancer	WDBC	569	31	2
Wisconsin prognostic breast cancer	WPBC	198	33	2
Wine recognition	Wine	178	13	3

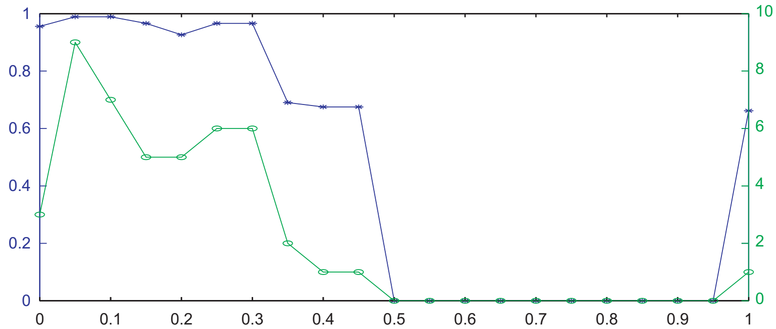

Fig. 1. Variation of number of reducts and classification accuracies with δ(data set Wine) (–∗– accuracy; –o– number).

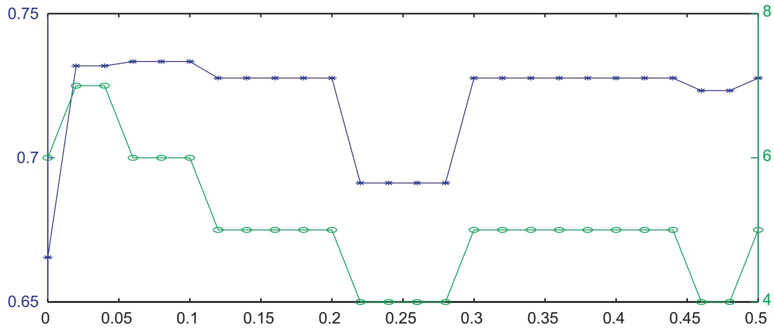

Fig. 2. Variation of number of reducts and classification accuracies with δ(data set credit) (–∗– accuracy; –o– number).

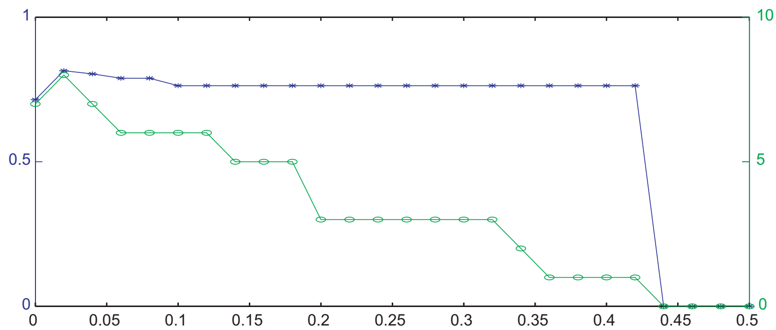

Fig. 3. Variation of number of reducts and classification accuracies with δ(data set heart) (–∗– accuracy; –o– number).

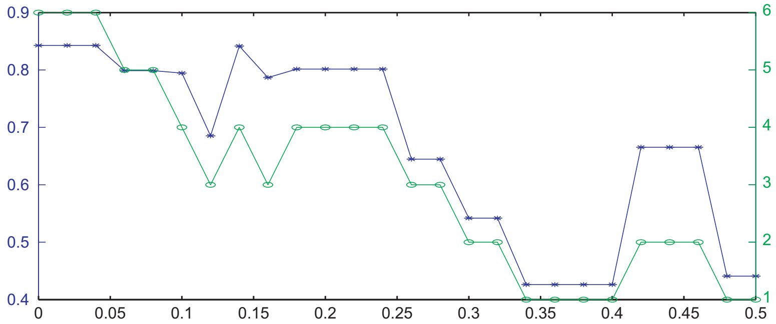

Fig. 4. Variation of number of reducts and classification accuracies with δ(data set Ecoli) (–*– accuracy; –o– number).

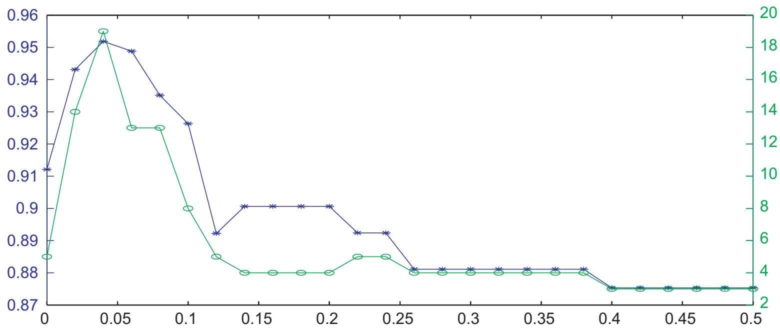

Fig. 5. Variation of number of reducts and classification accuracies with δ(data set Iono) (–*– accuracy; –o– number).

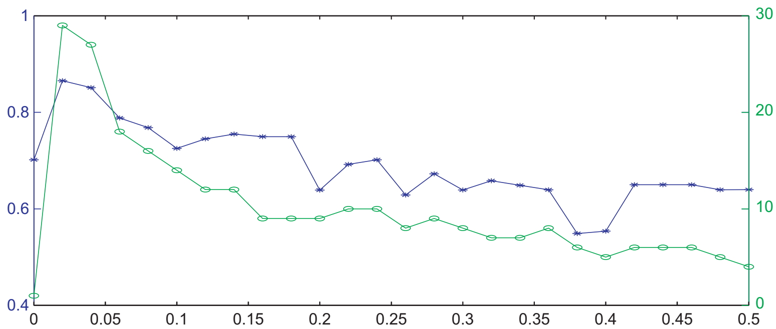

Fig. 6. Variation of number of reducts and classification accuracies with δ(data set Sonar) (–*– accuracy; –o– number).

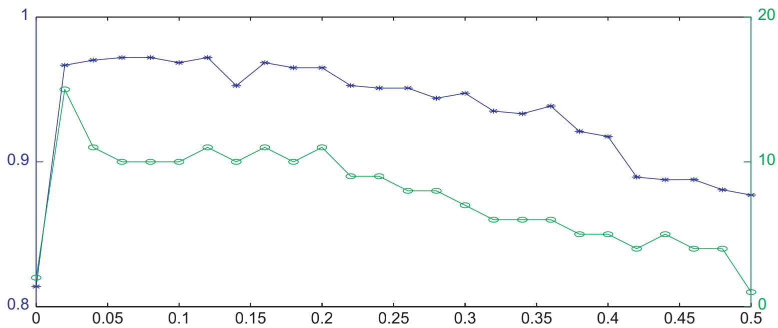

Fig. 7. Variation of number of reducts and classification accuracies with δ(data set WDBC) (–*– accuracy; –o– number).

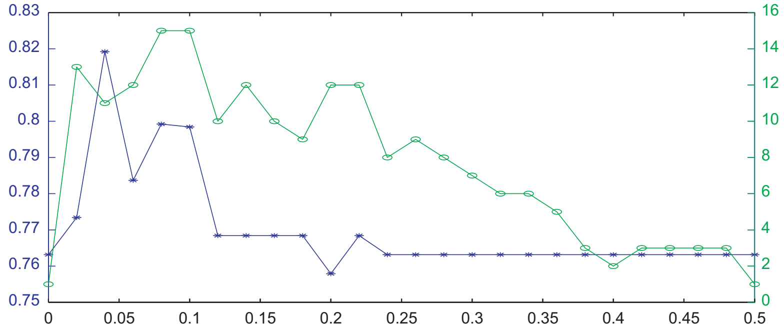

Fig. 8. Variation of number of reducts and classification accuracies with δ(data set WDPC) (–∗– accuracy; –o– number).

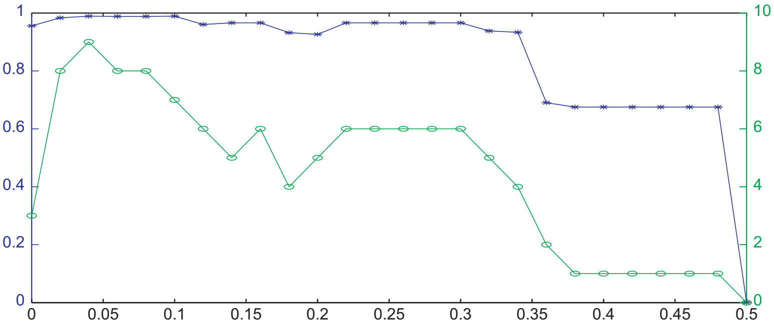

Fig. 9. Variation of number of reducts and classification accuracies with δ(data set Wine) (–∗– accuracy; –o– number).

Table 3
Comparison of the numbers of feature with different methods.

	Raw data	Equal width	Equal frequency	FCM	Related family
Credit	15	11	9	12	7
Heart	13	10	–	8	8
Ecoli	7	6	7	1	4
Iono	34	1	1	10	10
Sonar	60	7	0	6	29
WDBC	31	12	6	7	10
WPBC	33	9	0	4	11
Wine	13	5	4	4	7

features and the classification accuracy computed by 10-fold cross validation are shown in Figs. 2–9. These algorithms are implemented with OSU_SVM 3.00 obtained from http//www.ece.osu.edu/maj/osu_svm/osu_svm3.00.zip.

The size of the covering of an object, δ, must be specified in the methods based on covering rough sets. We first investigated the size of δ with a step size of 0.05. However, we found that the classification accuracy are relatively low when $\delta \in [0.5,1]$. As an example, the results of dataset Wine are shown in Fig. 1, with the patterns for other datasets being similar. The better selections are gathered in the interval [0,0.2] for all datasets. Therefore, we adjust δ from 0 to 0.5 with a step of 0.02.

As shown in Figs. 2–9, we can choose different reducts according to different levels of accuracy. The results exhibit qualitatively similar patterns. Generally, the numbers of selected features and the accuracies increase rapidly with the growth of δ at the beginning when $\delta = 0$, and then revert when δ continues to increase. Evidently, there is an optimal value for δ. In our case, the highest accuracy occurs when $\delta \in [0,0.2]$.

The comparisons in Tables 3 and 4 suggest that the related family method overwhelmingly outperforms Pawlak's rough sets. It can obtain reducts with higher accuracy, thereby maintaining or even improving the classification performance of the raw data sets. Moreover, the numbers of selected features for datasets Credit, Heart, and Iono, are less than those obtained from Pawlak's rough sets. We believe that the selected features obtained from Pawlak's rough sets contain some dispensable information in Credit, Heart, and Iono, which is reduced in the related family method. For all datasets, if the highest demanded accuracy is released, we can select a sub-high accurate reduct with much fewer elements than the optimal one. Furthermore, the greedy forward search reduction based on Pawlak's rough sets finds nothing for dataset Heart, because any attribute produces zero significance. The greedy forward search reduction performs poorly in such cases, while the

Table 4
Comparison of classification accuracies with different methods.

	Raw data	Equal width	Equal frequency	FCM	Related family
Credit	0.8144	0.8144	0.8028	0.8058	0.8534
Heart	0.7407	0.7630	0	0	0.8148
Ecoli	0.8512	0.8512	0.8512	0.4262	0.8416
Iono	0.9379	0.7499	0.7499	0.9348	0.9434
Sonar	0.8510	0.7398	0	0.7074	0.8655
WDBC	0.9808	0.9668	0.9597	0.9649	0.9720
WPBC	0.7779	0.7737	0	0.7837	0.8192
Wine	0.9889	0.9444	0.9660	0.9486	0.9889

heuristic algorithm proposed in this paper offers a good solution to this problem. Finally and more importantly, the related family method bridges the existing gap between covering reduction and incomplete information systems as discussed in Section 5.

8. Conclusion

In this paper, we presented examples to argue that discernibility matrices cannot be used to compute the reducts of the third type of covering generalized rough sets, which is regarded as the most reasonable type of them all. Therefore, a new method, referred to as the related family, was proposed to solve the problem. We constructed attribute reduction and relative attribute reduction algorithms for the third type based on the related family. The examples in this paper indicate that the related family bridges the gap where the discernibility matrix is not applicable. Moreover, we discussed the simplification of related families and the heuristic algorithms as well. Furthermore, the experimental analysis suggests that the related family method outperforms the traditional methods in dealing with both numerical and hybrid datasets. This indicates a promising path to replace the discernibility matrix in reducing covering rough sets in future studies.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant No. 11071061, No. 11201490 and 973 Program No. 2011CB311808.

References

[1] C.L. Blake, C.J. Merz, UCI Repository of machine learning databases, Tech. Rep., Univ. California, Dept. Inform. Comp. Sci., Irvine, CA, 1998. <http://www.ics.uci.edu/mlearn/MachineLearning.html>.
[2] J. Blaszczynski, R. Slowinski, M. Szelag, Sequential covering rule induction algorithm for variable consistency rough set approaches, Information Sciences 181 (2011) 987–1002.
[3] Z. Bonikowski, E. Bryniarski, U. Wybraniec-Skardowska, Extensions and intentions in the rough set theory, Information Sciences 107 (1998) 149–167.
[4] E. Bryniarski, A calculus of rough sets of the first order, Bulletin of the Polish Academy of Sciences 36 (1989) 71–77.
[5] D.G. Chen, C.Z. Wang, Q.H. Hu, A new approach to attributes reduction of consistent and inconsistent covering decision systems with covering rough sets, Information Sciences 177 (2007) 3500–3518.
[6] D. Chen, W. Zhang, D.S. Yeung, Eric C.C. Tsang, Rough approximations on a complete completely distributive latice with application to generalized rough sets, Information Science 176 (2006) 1829–1848.
[7] X.Y. Chen, Q.G. Li, Construction of rough approximations in fuzzy setting, Fuzzy Sets and Systems 158 (2007) 2641–2653.
[8] M. Diker, A.A. Ugur, Textures and covering based rough sets, Information Sciences 184 (2012) 44–63.
[9] G. Dong, J. Han, J. Lam, J. Pei, K. Wang, W. Zou, Mining constrained gradients in large databases, IEEE Transactions on Knowledge and Data Engineering 16 (2004) 922–938.
[10] Y. Du, Q.H. Hu, P.F. Zhu, P.J. Ma, Rule learning for classification based on neighborhood covering reduction, Information Sciences 181 (2011) 5457–5467.
[11] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition 40 (2007) 3509–3521.
[12] Q.H. Hu, H. Zhao, Z.X. Xie, D.R. Yu, Consistency based attribute reduction, Advvances in Knowledge Discovery and Data Mining, Publish: Springer Berlin, Heidelberg, pp. 96–107.
[13] Q.H. Hu, D.R. Yu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Information Sciences 178 (2008) 3577–3594.
[14] Q.H. Hu, D.R. Yu, Z.X. Xie, Neighborhood classifiers, Expert Systems with Applications 34 (2008) 866–876.
[15] Q.H. Hu, D.R. Yu, Z.X. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, Pattern Recognition Letters 27 (5) (2006) 414C423.
[16] M. Kryszkiewicz, Rules in incomplete information systems, Information Sciences 113 (1999) 271–292.
[17] T.J. Li, Y. Leung, W.X. Zhang, Generalized fuzzy rough approximation operators based on fuzzy coverings, International Journal of Approximation Reasoning 48 (2008) 836–856.
[18] T.J. Li, W.X. Zhang, Fuzzy rough approximations on two universes of discourse, Information Sciences 178 (2008) 892–906.
[19] T.Y. Lin, Topological and fuzzy rough sets, in: R. Slowinsk (Ed.), Intelligent Secision Support: Handbook of Applications and Advances of the Rough Set Theory, Kluwer Acedemic, Boston, 1992, pp. 287–304.
[20] T.Y. Lin, Chinese wall security model and conflict analysis, in: the 24th IEEE Computer Society International Computer Software and Applications Conference (Compsac2000) Taipei, Taiwan, October 25–27, 2000
[21] T.Y. Lin, Chinese wall security policy models: information flows and confining trojan horses, in: Sabrina de Capitani di Vimercati, Indrakshi Ray, Indrajit Ray (Eds.), Proceedings of the 17th IFIP11.3 Working Conference on Database and Applications Security, Estes Park, Colorado, USA, 2003.

[22] J.N. Mordeson, Rough set theory applied to (fuzzy) ideal theory, Fuzzy Sets and Systems 121 (2001) 315–324.
[23] S. Pal, P. Mitra, Case generation using rough sets with fuzzy representation, IEEE Transactions on Knowledge and Data Engineering 16 (2004) 292–300.
[24] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (1982) 341–356.
[25] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Acedemic Publishers, Boston, 1991.
[26] L. Polkowski, A. Skowron (Eds.), Rough Sets and Current Trends in Computing, vol. 14, Springer, Berlin, 1998.
[27] L. Polkowski, A. Skowron (Eds.), Rough Sets in Knowledge Discovery, vol. 1, Physica-Verlag, Berlin, 1998.
[28] L. Polkowski, A. Skowron (Eds.), Rough Sets in Knowledge Discovery, vol. 2, Physica-Verlag, Berlin, 1998.
[29] J.A. Pomykala, Approximation operations in approximation space, Bulletin of the Polish Academy of Sciences 35 (9–10) (1987) 653–662.
[30] Zh. H. Shi, Z.T. Gong, The further investigation of covering-based rough sets: uncertainty characterization, similarity measure and generalized models, Information Sciences 180 (2010) 3745–3763.
[31] A. Skowron, J. Stepaniuk, Tolerance approximation spaces, Fundamenta Informaticae 27 (1996) 245–253.
[32] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, intelligent decision suport, in: R. Slowinski (Ed.), Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, 1992.
[33] Eric C.C. Tsang, D. Chen, D.S. Yeung, Approximations and reducts with covering generalized rough sets, Computers and Mathematics with Applications 56 (2008) 279–289.
[34] E. Tsang, D. Chen, J. Lee, D.S.Yeung, On the upper approximations of covering generalized rough sets, in: Proc.3rd International Conf. Machine Learning and Cybernetics, 2004, pp. 4200–4203.
[35] C.Z. Wang, C.X. Wu, D.G. Chen, A systematic study on attributes reduction with rough sets based on general binary relations, Information Sciences 178 (2008) 2237–2261.
[36] S.K.M. Wong, W. Ziarko, On optimal decision rules in decision tables, Bulletin of Polish Acedemy of Sciences 33 (1985) 693–696.
[37] W.H. Xu, W.X. Zhang, Measureing roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems 158 (2007) 2443–2455.
[38] Z. Xu, Q. Wang, On the properties of covering rough sets model, Journal of Henan Normal University(Natural Science) 33 (2005) 130–132.
[39] T. Yang, Q.G. Li, Reduction about approximation spaces of covering generalized rough sets, International Journal of Aproximate Reasoning (51) (2010) 335–345.
[40] Tian Yang, Qingguo Li, Bilei Zhou, A new method of attribute reduction of covering rough sets, in: The 7th IEEE International Conference on Granular Computing, Kaohsiung, Taiwan, November 8–10, 2011 (Best Paper Award).
[41] Y.Y. Yao, On generalizing pawlak approximation operators, LNAI 1424 (1998) 298–307.
[42] Y.Y. Yao, Constructive and algebraic methods of theory of rough sets, Information Sciences 109 (1998) 21–47.
[43] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111 (1998) 239–259.
[44] Z.Q. Yun, X. Ge, X.L. Bai, Axiomatization and conditions for neighborhoods in a covering to form a partition, Information Sciences 181 (2011) 1735–1740.
[45] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (1997) 111–127.
[46] Y.L. Zhang, J.J. Li, W.Z. Wu, On axiomatic characterizations of three pairs of covering based approximation operators, Information Sciences 180 (2010) 274–287.
[47] Y.L. Zhang, M.K. Luo, On minimization of axiom sets characterizing covering-based approximation operators, Information Sciences 181 (2011) 3032–3042.
[48] N. Zhong, Y.Y. Yao, M. Ohshima, Peculiarity oriented multidatabase mining, IEEE Transactions on Knowledge and Data Engineering 15 (4) (2003) 952–960.
[49] W. Zhu, F.Y. Wang, Reduction and axiomization of covering generalized rough sets, Information Sciences 152 (2003) 217–230.
[50] W. Zhu, Topological approaches to covering rough sets, Information Sciences 177 (2007) 1499–1508.
[51] W. Zhu, Relationship between generalized rough sets based on binary relation and coverings, Information Sciences 179 (2009) 210–225.
[52] W. Zhu, F.Y. Wang, A new type of covering rough sets, in: IEEE IS 2006, London, 4–6 September 2006, pp. 444–449.
[53] W. Zhu, Relationship among basic concepts in covering-based rough sets, Information Sciences 179 (2009) 2478–2486.