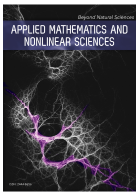
sciendo Applied Mathematics and Nonlinear Sciences 8(1) (2023) 2103–2116



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

Innovations to Attribute Reduction of Covering Decision System Based on Conditional Information Entropy

Xiuyun Xia[1], Hao Tian[2], Ye Wang[3]

[1]School of General Education, Hunan University of Information Technology, Changsha 410005, China

[2]Electronic Information College, Hunan University of Information Technology, Changsha 410005, China

[3]School of Computer Science, Huaiyin Normal University, Huaian 223000, China

Submission Info

Communicated by Juan Luis García Guirao
Received October 28th 2020
Accepted November 26th 2020
Available online April 15th 2022

Abstract

Traditional rough set theory is mainly used to reduce attributes and extract rules in databases in which attributes are characterised by partitions, which the covering rough set theory, a generalisation of traditional rough set theory, covers. In this article, we posit a method to reduce the attributes of covering decision systems, which are databases incarnated in the form of covers. First, we define different covering decision systems and their attributes' reductions. Further, we describe the necessity and sufficiency for reductions. Thereafter, we construct a discernible matrix to design algorithms that compute all the reductions of covering decision systems. Finally, the above methods are illustrated using a practical example and the obtained results are contrasted with other results.

Keywords: discernible matrix; information entropy; decision system; attribute.

1 Introduction

The Rough Set theory was proposed by Polish mathematician Zdzisław Pawlak in 1982 [1, 3, 17]. It can effectively handle uncertain, inaccurate and incompSupposinge information . Recently, rough set theory has successfully been applied in many fields, including machine learning, pattern recognition, decision analysis, process control and data mining [4, 5, 10–14]. Therefore, this theory has received great attention from the international information science community, computer science and mathematics. A large number of studies in the literature have witnessed the development of rough sets in theory and application [26–29]. Therefore, some scholars have extended the partition to encompass aspects explained in the above reasons, and the scope of rough set theory research has been greatly expanded. Although much attention is paid to the set approximation of cover, sparse research has been carried out in relation to the attribute reduction of covering rough sets. The

sciendo
ISSN 2444-8656 doi:10.2478/AMNS.2021.1.00110

attribute reduction under the algebraic point of view is generalised on the basis of the attribute reduction of conditional information entropy.

At present, certain scholars have studied the rough set theory based on their conclusions drawn from the information theory, and have proposed the information theory about rough set theory. Wang et al. [2] posited the reduction of the decision table and the common properties and different characteristics of information. Yang [6] proposed an approximate reduction method on basis of conditional information entropy in a decision table. On this basis, an approximate reduction method was proposed for vertical multidistribution decision tables [8]. Wang Yan et al also used a reduction algorithm on information entropy and identifiable Matrix, and presented a new combination algorithm [9]. After Hudan et al. added a probability measure to the rough set theory [7], some concepts and properties of information theory and rough set theory were compared, and a new method of rule extraction was obtained. Hu et al. [15] proposed a rough entropy method based on generalised rough set coverage reduction. Chen et al. [16] proposed an optimal section for reducing the superfluous cover. Yang [18] performed research on rough set methods, from the attribute reduction problem on inconsistent decision systems to the attribute reduction problem on consistent decision systems. Guo [19] studied knowledge reduction based on rough set theory for the inconsistent decision systems, such as generalised decision table, relative resolution and knowledge reduction. Shi et al. [20] proposed attribute reduction based on the Boolean matrix. Li and Yin [21] proposed a reduction algorithm of covering system on information theory. Ma [22] constructed a decision tree based on the covering rough set theory. Chen et al. [23] got a multi-label attribute reduction algorithm on neighbourhood rough set. Zhang et al. [24] developed the belief and plausibility functions from the evidence theory and these are employed to characterise attribute reductions in the covering decision information system. Zhang et al. [25] posited confidence-preserved attribute reduction and algorithms of rule acquisition in covering decision systems. Jiang et al. [30] presented an accelerator for multi-granularity attribute reduction knowledge-based systems from another angle. However, we resolve and analyse the problem in consistent and inconsistent covering decision systems based on conditional information entropy in this article.

In this article, we propose a method to reduce the attributes of covering decision systems, which are databases characterised by covers. First, we define two scenarios of covering decision systems and their attributes' reductions. Second, we state the necessity and sufficiency for reductions. Third, we construct a discernible matrix to design algorithms that compute all the reductions of different covering decision systems. Finally, the above methods are illustrated using a practical example and the obtained results are in contrast to other results.

2 Preliminaries

We go over the basic concepts related to covering rough sets which can be found in the literature [1, 4, 11, 16, 17, 19, 21, 22, 24, 25].

Definition 1. The ordered pair (U,C), where U is any nonempty set called a universe, and C its finite covering (i.e. C is a finite family of nonempty subsets of U and $\cup C = U$), is what we define as the covering approximation space, or in short, the approximation space; the covering C is called then the family of approximating sets.

Definition 2. Supposing that (U,C) be a covering approximation space, and x belong to any element of U. The following set:

$$M(x) = \{K \in C : x \in K \wedge \forall S \in C (x \in S \wedge S \subseteq K \Rightarrow K = S)\}$$

is called the *minimal description* of the object x.

Definition 3. Supposing that $A = \{A_1, A_2, \cdots, A_n\}$ be a set of covers of U, if $\forall x \in U$, $A_x = \cap\{A_j; A_j \in A, x \in A_j\}$ holds; then $Cov(A) = \{A_x; x \in U\}$ is also a covering of U; so, we call it the leaded covering of A.

Definition 4. Supposing that $\pi = \{A_i : i = 1, \ldots m\}$ be a family of covers of U, if $\forall x \in U$, $\pi_x = \cap\{A_{ix}; A_{ix} \in Cov(A_i), x \in A_{ix}\}$ holds; then $Cov(\pi) = \{\pi_x; x \in U\}$ is also a covering of U; so we call it the leaded covering of π.

3 Attribute reduction of consistent covering decision systems based on conditional information entropy

In this section, we focus on investigating the basic concepts and key results of consistent covering decision systems [1,4,16,19,21].

3.1 Basic definition of consistent covering decision systems based on conditional information entropy

Definition 5. Supposing that $\pi = \{A_i : i = 1, \dots m\}$ be a set of covers of U, d is a decision attribute and U/d is a decision partition on U. If $\forall$ and $\exists E_j \in U/d$ such that $\pi_x \subseteq d_j$, then decision system $S = (U, \pi, d)$ is called a consistent covering decision system, and recorded as $H(d|\pi) = 0$, i.e. otherwise, $S = (U, \pi, d)$ is called an inconsistent covering decision system.

$$\left(\text{Appendix :} H(d|\pi) = -\frac{1}{|U|}\sum_{x \in U} \log \frac{|[x]_d \cap \pi_x|}{\pi_x}\right)$$

The positive region of d relative to π is defined as $POS_\pi(d) = \underset{X \in U/D}{\cup} \underline{\pi}(X)$.

For $\forall X \in U/d$, if we believe every $\bar{\pi}X \Rightarrow X$to be the possible rule and every $\underline{\pi} - X \Rightarrow X$ to be a certain rule, then all the decision rules extracted from a consistent covering decision system are consistent.

If every cover in π is a partition, then $\text{cov}(\pi)$ is also a partition, and $H(d|\pi) = 0$ is just the case of a consistent decision system in traditional rough set theory. Supposing that $f_d(x)$ is a decision function $f_d : U \to V_d$ of the universe U into value set V_d; then, for $\forall x_i, x_j \in U$,if$\pi_{xi} \subseteq \pi_{xj}$, $f_d(x_i) = f_d([x_i]_d) = f_d(\pi_{xi}) = f_d(\pi_{xj}) = f_d(x_j) = f_d([x_j]_d)$.

1. If $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$, then $\pi_{xi} \cap \pi_{xj} = \emptyset$, i.e. $\pi_{xi} \not\subset \pi_{xj}$ and $\pi_{xj} \not\subset \pi_{xi}$; on the other hand, if $\pi_{xi} \not\subset \pi_{xj}$and $\pi_{xj} \not\subset \pi_{xi}$, then either $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$ or $f_d(\pi_{xi}) = f_d(\pi_{xj})$ are possible.

2. On the other hand, if $\pi_{xi} \cap \pi_{xj} \neq \emptyset$, we get $f_d(\pi_{xi}) = f_d(\pi_{xj})$; to the contrary, if $f_d(\pi_{xi}) = f_d(\pi_{xj})$, then we have $\pi_{xi} \not\subset \pi_{xj}$ and $\pi_{xj} \not\subset \pi_{xi}$, or $\pi_{xi} \subseteq \pi_{xj}or\pi_{xi} \supseteq \pi_{xj}$ are possible.

Further, we define the relative reduction of a consistent covering decision system.

Definition 6. Supposing that $S = (U, \pi, d)$ be a consistent covering decision system and there exists $A_i \in \pi$ if $H(d|\pi - \{A_i\}) = 0$, then A_i is called a superfluous relative to d in π; otherwise, A_i is known as a indispensable relative to d in π. For every $B \subseteq \pi$ satisfying $H(d|B) = 0$, if every element in $B$is indispensable, i.e. for $\forall A_i \in B$,$H(d|B - \{A_i\}) = 0$ being not true, then $B$is called a reduction ofπ relative to$d$, in short known as relative reduction. The collection of all the indispensable elements in π is called the core of π relative to d, denoted as $Core_d(\pi)$.

The relative reduction of a consistent covering decision system is the minimal set of conditional covers (attributes) which ensure that every decision rule is still consistent. For a single cover A_i, we give out some equivalence conditions to judge whether it is indispensable.

3.2 The key points of consistent covering decision system based on conditional entropy

Theorem 1. *Suppose $H(d|\pi) = 0$, if and only if for every $x, y \in U$, if $x \in \pi_y$, then $x \in [y]_d$.*

Proof. Since for every $y \in U$,$H(d|\pi) = -\frac{1}{|U|}\sum_{y \in U} \log \frac{|[y]_d \cap \pi_y|}{|\pi_y|} = 0$, such that $\pi_y \subseteq [y]_D$. If for every $x, y \in U$,$x \in \pi_y$, then $x \in [y]_d$. On the other hand, for every $x, y \in U$, if $x \in \pi_y$, we have $x \in [y]_d$, hence $\pi_y \subseteq [y]_d$. Therefore, for every$y \in U$, $H(d|\pi) = 0$.□

Theorem 2. *Suppose $H(d|\pi) = 0$ if and only if $POS_\pi(d) = U$.*

Proof. $\Rightarrow$ For every $y \in POS_\pi(d)$, then $\pi_y \subseteq [x]_d$, therefore $y \in [x]_d$, and so we have $[y]_d = [x]_d$, i.e. $\pi_y \cap [y]_d \neq \emptyset$. For every $x \in \pi_y$, then $x \in [y]_d$ and by the Theorem 1, we get $H(d|\pi) = 0$.

$\Leftarrow$ Since $POS_\pi(d) = \bigcup_{X \in U/d} \underline{\pi}(X)$, we have $\underline{\pi}(X) \subseteq X$, hence $\cup\underline{\pi}(X) \subseteq \cup X = U$, and so we get $POS_\pi(d) \subseteq U$. On the other hand, since $POS_\pi(d) = \bigcup_{X \in U/d} \underline{\pi}(X) = \cup\underline{\pi}([x]_d)$, for every $y \in POS_\pi(d)$, then $\pi_y \subseteq [x]_d$. By Theorem 1, $H(d|\pi) = 0$ if and only if for every $x, y \in U, x \in \pi_y$, then $x \in [y]_d$; hence $U \subseteq [y]_d$; then by $\pi_y \subseteq [x]_d$ and $[y]_d \subseteq [x]_d$, we have $U \subseteq [x]_d$; hence $x \in POS_\pi(d)$ such that $U \subseteq POS_\pi(d)$. So, the result is true.□

By the above two theorems of discussions, we obtain the following two corollaries.

Corollary 1. *Suppose $H(d|\pi) = 0$, $B \subseteq \pi$ is a positive-region consistent set on S; then (U, B, d) is a consistent covering decision system.*

Corollary 2. *Suppose $H(d|\pi) = 0$, $B \subseteq \pi$ is a positive-region reduction on S; then B is a minimal set such that decision system (U, B, d) is a consistent covering decision system.*

Theorem 3. *Suppose $H(d|\pi) = 0$, $A_i \in \pi$ and $Cov(\pi - \{A_i\}) = \{B_x : x \in U\}$; then A_i is indispensable, i.e. $H(d|\pi - \{A_i\}) = 0$ is not true if and only if there exists $x \in U$ such that $B_x \subseteq [x]_D$ is not true.*

Proof. If there exists $x \in U$ such that $B_x \subseteq [x]_D$ is not true, by $x \in B_x$, then for every $y \in U$ such that $B_x \subseteq [y]_D$ is not true. So $H(d|\pi - \{A_i\}) = 0$ is not true.

If $H(d|\pi - \{A_i\}) = 0$ is not true, then there exists , for $\forall y \in U$ such that $B_x \subseteq [y]_D$ is not true. Especially, $B_x \subseteq [x]_D$ is also not true.

It should be indicated that $B_x \subseteq [x]_D$ not being true means that $(U, \pi - \{A_i\}, d)$ is an inconsistent decision system, i.e. $H(d|\pi - \{A_i\}) \neq 0$, A_i is thus indispensable implies it is a key cover to ensure (U, π, d) is a consistent decision system, i.e. $H(d|\pi) = 0$.□

Theorem 4. *Suppose $H(d|\pi) = 0, A_i \in \pi$, and $Cov(\pi - \{A_i\}) = \{B_x : x \in U\}$, then $H(\{A_i\}|\pi - \{A_i\}) > 0$, i.e. A_i is absolutely indispensable.*

In other words, $H(\{A_i\}|\pi - \{A_i\}) > 0$ is not true if and only if A_i is called superfluous in B.

Proof. We assume that $Cov(\pi - \{A_i\}) = \{B_x : x \in U\}$ and $Cov(\{A_i\}) = \{A_{ix} : x \in U\}$, If $A_i$is called superfluous in B, for $x \in U, B_x$ have equivalence values on A_i, i.e. for every $B_x (x \in U)$ there is $A_i (x \in U)$ such that $B_x \subseteq A_{ix}$, so $H(\{A_i\}|\pi - \{A_i\})$=$-\frac{1}{|U|} \sum_{x \in U} \log \frac{|A_{ix} \cap P_x|}{|P_x|}$ =0.

If $H(\{A_i\}|\pi - \{A_i\}) > 0$ is not true, then $H(\{A_i\}|\pi - \{A_i\})$=$-\frac{1}{|U|} \sum_{x \in U} \log \frac{|A_{ix} \cap B_x|}{|B_x|}$, and for every x, i, $\frac{1}{|U|} \sum_{x \in U} \log \frac{|A_{ix} \cap B_x|}{|B_x|} \leq$ 0 and $\frac{1}{|U|} > 0$, so $\sum_{x \in U} \log \frac{|A_{ix} \cap B_x|}{|B_x|} \leq 0$ is true; if there exists a certain j such that $0 < \frac{|A_{jx} \cap B_x|}{|B_x|} < 1$; otherwise it implies $H(\{A_i\}|\pi - \{A_i\})$=$-\frac{1}{|U|} \sum_{x \in U} \log \frac{|A_{jx} \cap B_x|}{|B_x|}$>0, which is a contradiction. So for every x, i there is $\frac{|A_{ix} \cap B_x|}{|B_x|} = 1$, i.e. $P_x \subseteq A_{ix}$, hence $H(\{A_i\}|\pi - \{A_i\}) > 0$ such that $H(\{A_i\}|\pi - \{A_i\}) > 0$ is not true.□

Theorem 4 implies that the superfluous knowledge in question could not supply new and useful information to the concerned information system. However, the necessary knowledge could give helpful information for information systems.

Corollary 3. *Suppose $H(d|\pi) = 0$, $A_i \in \pi$, then A_i is superfluous relative tod in π if and only if $H(d|\pi - \{A_i\}) = H(d|\pi) = 0$.*

Theorem 5. *Assume that $H(d|\pi) = 0$, $B \subseteq \pi$ is called a reduction of π relative todif and only if*

1. *$H(d|B) = H(d|\pi)$, and*

2. *for every $A_i \in B$, then $H(\{A_i\}|B - \{A_i\}) > 0$.*

Proof. $B \subseteq \pi$ is called a reduction of π relative to d if and only if for every $A_i \in B$ such that $H(d|B - \{A_i\}) = 0$ is not true and B is independent. As we know, it implies that A_i is indispensable attribute in B such that $H(d|B) = 0$; therefore we get $H(d|B) = H(d|\pi)$.

By Theorem 4, A_i is independent in B if and only if for every $A_i \in B$, then $H(\{A_i\}|B - \{A_i\}) > 0$.□

Based on the discussion of the above theorems, we could consider indicial form of information entropy as being equivalent to expressions of algebra for attribute reduction.

Theorem 6. Assume that $H(d|\pi)=0$,$A_i \in \pi$,A_i is then indispensable, i.e. $H(d|\pi-\{A_i\})=0$ is not true if and only if there is at least a pair of $x_i, x_j \in U$ satisfying $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$, and the factor which has the foremost relation toπtransfers behind A_i and is removed from π.

Proof. $\Rightarrow$ We note that if $Cov(\pi-\{A_i\})=\{B_x : x \in U\}$, if $H(d|\pi-\{A_i\})=0$ is not true, then there are $x_0, y_0 \in U$ such that $y_0 \in B_{x0}$ and $y_0 \notin [x_0]_d$, which implies $B_{y0} \subseteq B_{x0}$ and $\pi_{x0} \not\subset \pi_{y0}$,$\pi_{y0} \not\subset \pi_{x0}$, respectively, and x_0, y_0 satisfying$f_d(\pi_{x0}) \neq f_d(\pi_{y0})$, so the factor that has the foremost relation of x_0, y_0 with regard to π transfers behind A_i and is removed from π.

$\Leftarrow$ Suppose $x_0, y_0 \in U$ satisfying $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$, which implies $[x_0]_d \cap [y_0]_d = \emptyset$, and $\pi_{x0} \not\subset \pi_{y0}$,$\pi_{y0} \not\subset \pi_{x0}$, if their foremost relation x_0, y_0 with regard to π transfers behind A_i removing from π, which implies $B_{x0} \not\subset B_{y0}$or $B_{y0} \not\subset B_{x0}$ is not effective. Thus if $B_{y0} \not\subset B_{x0}$, then $B_{y0} \subseteq [y_0]_d$ is not effective, otherwise it points out $x_0 \in [y_0]_d$ which is a contradiction; if $B_{x0} \supseteq B_{y0}$, then $B_{x0} \subseteq [x_0]_d$ is not virtual; if $B_{x0} = B_{y0}$ then $B_{y0} \subseteq [y_0]_d$ and $B_{x0} \subseteq [x_0]_d$ are not effective. Hence $H(d|\pi-\{A_i\})=0$ is not true.□

Theorem 6 implies that an indispensable cover can be characterised by the foremost relation between two elements in the universe. Thus, we have the following theorem to characterise a consistent decision system.

Theorem 7. *Suppose $H(d|\pi)=0$,$B \subseteq \pi$, then$H(d|B)=0$ if and only if for every $x_i, x_j \in U$ satisfying $f_d(\pi_{x0}) \neq f_d(\pi_{y0})$, the relation x_i, x_j with regard to π is equivalent to their relation with regard to B, i.e. $\pi_{xi} \not\subset \pi_{xj}, \pi_{xj} \not\subset \pi_{xi} \Leftrightarrow B_{xi} \not\subset B_{xj}, B_{xj} \not\subset B_{xi}$.*

Proof. Since the proof is similar, here there is no need to repeat it.□

3.3 Attribute reduction of consistent covering decision system

The intention of relative reduction of covering attribute π is to discover the minimal subset of π to preserve every decision rule invariant. Through the theorems stated in the previous section, we understand that it is coordinative to preserve the foremost relation of every two elements to which variant decision values are invariant. Based on this understanding, we could construct an algorithm to compute all the relative reductions. Definition 7 can be found in the literature [16, 21, 23].

Definition 7. Supposing (U, π, d) be a consistent decision system. Assume $U=\{x_1, x_2, \dots x_n\}$; by $M(U, \pi, d)$, we mark a $n \times n$ matrix (m_{ij}), called the discernibility matrix of (U, π, d), and defined as follows: $A \in \pi : (A_{xi} \neq A_{xj}) \wedge (A_{xi} \not\subset A_{xj}) \wedge (A_{xj} \not\subset A_{xi})$, $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$=$m_{ij}$ π, $f_d(\pi_{xi}) = f_d(\pi_{xj})$ for .$x_i, x_j \in U$

Theorem 8. Supposing (U, π, d) be a consistent covering decision system, then we obtain

(1) For every $B \subseteq \pi$, $B \cap m_{ij} \neq \emptyset$ is a valid result for every $i, j \leq n$ if and only if $H(d|\pi)=0$.

(2) $Core_d(\pi)=\{\forall A \in \pi \wedge (\exists m_{ij} \in M_{n\times n}(S) \wedge m_{ij}=\{A\}\}$ for some i, j.

Proof.

1. Assume $H(d|\pi)=0$. If $f_d(\pi_{xi}) = f_d(\pi_{xj})$, then $m_{ij}=\pi$, hence $B \cap m_{ij} \neq \emptyset$ is true for every $B \subseteq \pi$. If $f_d(\pi_{xi}) \neq f_d(\pi_{xj})$, and since $\pi_{xi} \subseteq B_{xi} \subseteq [x_i]_d$, $\pi_{xj} \subseteq B_{xj} \subseteq [x_j]_d$, then, we have $f_d(B_{xi}) \neq f_d(B_{xj})$. By the Theorem 7, $\pi_{xi} \not\subset \pi_{xj}, \pi_{xj} \not\subset \pi_{xi} \Leftrightarrow B_{xi} \not\subset B_{xj}, B_{xj} \not\subset B_{xi}$, there is a $A_{i0} \in B$ such that $A_{i0x} \neq A_{j0y}$, i.e. $A_{i0x} \not\subset A_{j0y}$,$A_{j0y} \not\subset A_{i0x}$ and so we have $B \cap m_{ij} \neq \emptyset$.

   On the contrary, if $B \cap m_{ij} \neq \emptyset$, for $x, y \in U$ satisfying $f_d(\pi_x) \neq f_d(\pi_y)$, it points out that there are enough covers in B to maintain the relation of x, y with regard to π equivalent to the relation of x, y with regard to B. Thus we obtain $H(d|\pi)=0$.

2. If every , then $H(d|\pi-\{A_i\})=0$ is not true, which means there exists at least a pair of $x, y \in U$ satisfying $f_d(\pi_x) \neq f_d(\pi_y)$ whose foremost relation with regard to π transfer behind A is removed from π. Thus A is the only cover in π satisfying $(A_{xi} \not\subset A_{xj}) \wedge (A_{xj} \not\subset A_{xi})$. By the Definition 7, we have $m_{ij}=\{A\}$.

Hence we have $Core_D(\pi)$$\subseteq$$\{\forall A \in \pi \wedge (\exists m_{ij} \in M_{n\times n}(S) \wedge m_{ij}=\{A\}\}$. If $m_{ij}=\{A\}$ for some i, j, it is obvious $A \in Core_D(\pi)$.□

The value of core in information system is exclusive, which is the most important part of knowledge category in the information system.

By the Theorem 8(2), a method is used which can directly obtain $Core_d(\pi)$, for any $A \in Core_d(\pi)$ if and only if there is at least a m_{ij} satisfying $m_{ij} = \{A\}$for $1 \leq j \leq i \leq n$ of the discernibility matrix $M(S)$. Suppose $E = \pi - Core_d(\pi)$, then covering attribute on E can be reduced from the known values of core set, but it cannot be reduced at the same time. If π have a minimal subset, i.e. $Core(\pi) = Min(\pi)$, then, we could reduce covering attribute synchronously. Generally speaking, π have several minimal subsets.

Theorem 9. *Supposing (U, π, d) be a consistent covering decision system. Suppose $E_1 \subseteq E, E = \pi - Core_d(\pi)$, then $H(d|\pi - \{E_i\}) = 0$ is true if and only if $m_{ij} \not\subset E_1$ for all $m_{ij}, 1 \leq j \leq i \leq n$.*

Proof. $\Leftarrow$ Since $E_1 \subseteq E$, then any element of E cannot contain in $Core_d(\pi)$, or in other words, it could be reduced. $m_{ij} \not\subset E_1$ for all m_{ij}, $1 \leq j \leq i \leq n$ means that we could guarantee $H(d|\pi - \{E_i\}) = 0$ after the attributes on E_1 are reduced.

$\Rightarrow$ If the attributes on E_1 could be reduced at the same time, and there also exists a m_{ij} such that $m_{ij} \subseteq E_1$, then we obtain $m_{ij} = \emptyset$ after the attributes on E_1 are reduced at the same time, which point out the attributes on E_1 could not be reduced at the same time.□

4 Attribute reduction of inconsistent covering decision systems based on information entropy

In many of practical problems, we always have inconsistent covering decision system. In this section, we propose attribute reductions for inconsistent covering decision systems. We understand that some rules extracted from inconsistent decision systems may not be consistent. As to covering decision system, experts can still give the decision-making in the case of inconsistent information, so it can be assumed that the decision-making property is not empty. So we have the following definition of attribute reduction. We always suppose U is a finite universe and $\pi = \{A_i : i = 1, \cdots m\}$ is a family of covers of U. Then, the induced cover of π is defined as in terms of $Cov(\pi) = \{\pi_x : x \in U\}$, $U/d = \{[x]_d : x \in U\}$ is the decision partition, d is a decision attribute and $POS_\pi(d) \neq \emptyset$.

4.1 Key definitions of inconsistent covering decision systems based on information entropy

In this subsection, we discuss the key definition of inconsistent covering decision systems which can be found in the literature [7, 16, 17, 19].

Definition 8. Suppose U is a finite universal set and $\pi = A_i : i = 1, \ldots, m$ is a family of covers of U, $A_i \in \pi$, d is a decision attribute relative to π on U and $f_d : U \to V_d$ is the decision function, denoted by $f_d(x) = [x]_d$, then, $Icds = U, \pi, d$ is an inconsistent covering decision system, i.e. $H(d|\pi) \neq 0$ or $POS_\pi(d) \neq U$.

Definition 9. Suppose $Icds = U, \pi, d$ is an inconsistent covering decision system. For $B \subseteq \pi$, we define the limitary entropy, d, and limitary conditional entropy, π, as to attribute decision by

$$LH(d|\pi) = -\frac{1}{|K|} \sum_{x \in K} \log \frac{|[x]_d \cap \pi_x|}{|\pi_x|};$$

$$LH(d) = -\frac{1}{|K|} \sum_{x \in K} \log \frac{[x]_d}{|K|};$$

the limitary mutual information of B and d is defined as follows

$$LI(B;d) = LH(d) - LH(d|B) = -\frac{1}{|K|} \sum_{x \in K} \log \frac{[x]_d}{|K|} - \frac{1}{|K|} \sum_{x \in K} \log \frac{|[x]_d \cap B_x|}{|B_x|}$$

where $K = \{x \in U : \pi_x \subseteq [x]_d\}$.

Definition 10. Suppose $Icds = U, \pi, d$ is an inconsistent covering decision system. For every $A_i \in \pi$ and $LH(d|\pi - \{A_i\}) = LH(d|\pi)$, then A_i is superfluous relative to d in π. For every and , then is independent relative to in .

Definition 11. Suppose is an inconsistent covering decision system. For every $B \subseteq \pi$, if two following conditions are met:

1. $LI(B;d) = LI(\pi;d)$,
2. If $\forall A_i \in B$, then $LH(d|B) < LH(d|B - \{A_i\})$,

Then B is a reduct of π relative to d, denoted by $B \in Red_\pi(d)$.

4.2 Basic properties of inconsistent covering decision system based on positive region

Proposition 1. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system. For every $W \subseteq B \subseteq \pi$, then $POS_W(d) \subseteq POS_B(d)$.*
Proof. Suppose $Cov(B) = \{B_x : x \in U\}$ and $Cov(W) = \{W_x : x \in U\}$. For every $x \in POS_W(d)$, then there exists $X \in U/d$, such that $x \in \underline{W}(X)$. There exists $y \in W_x$, for every x such that $W_x \subseteq X$. Suppose $B = W \cup \{b_1, \dots, b_t\}$, then the following attribute sets are made:

$$B_1 = W \cup \{b_1\};$$

$$B_2 = B_1 \cup \{b_2\};$$

$$\dots$$

$$B_t = B_{t-1} \cup \{b_t\};$$

Obviously, we can get $W \subseteq B_1 \subseteq \ldots \subseteq B_t = B$. For $\forall b \in B$, there exists $B_x \subseteq (B - \{b\})_x$ such that $B_x = B_{tx} \subseteq B_{(t-1)x} \subseteq \cdots \subseteq B_{1x} \subseteq W_x$, i.e. $B_x \subseteq W_x$. Based on the covering lower approximation, we can have $x \in \underline{B}(X)$, then $x \in POS_B(d)$. So for every $W \subseteq B$, then $POS_W(d) \subseteq POS_B(d)$.□
Theorem 10. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system. For $B \subseteq \pi, |POS_\pi(d)| = |POS_B(d)|$ is true if and only if $POS_\pi(d) = POS_B(d)$.*
Proof. Here we will not prove it.□
Theorem 11. *Suppose $Icds = (U, \pi, d)$ is an inconsistent covering decision system, for every $B \subseteq \pi$, if b is a superfluous attribute relative to d in π, then $Core_B(d) \subseteq Core_{B-\{b\}}(d)$ is true.*
Proof. If $a \in Core_B(d)$ and $B - \{a\} \subset B \subseteq \pi$, then $POS_\pi(d) \supseteq POS_{B-\{a\}}(d)$. Since a is an indispensable attribute, then $POS_\pi(d) \neq POS_{B-\{a\}}(d)$, thus we can get $x_1 \notin POS_{B-\{a\}}(d)$ such that $x_1 \notin \underline{B - \{a\}}(X)$, we can further get $(B - \{a\})_{x1} \not\subset X$. For every $(B - \{a\})_{x1} \not\subset X$ and $(B - \{a\})_{x1} \subseteq \overline{((B - \{a\}) - \{b\})}_{x1}$, then $((B - \{b\}) - \{a\})_{x1}$ $((B - \{a\}) - \{b\})_{x1} \not\subset X$, which shows that $x_1 \notin POS_{(B-\{b\})-\{a\}}(d)$. Since X have arbitrary, $x_1 \neq POS_{(B-\{b\})-\{a\}}(d)$, thus $POS_{(B-\{b\})-a}(d) \neq POS_{B-\{a\}}(d)$. By the given condition, we will get $Core_{B-\{b\}}(d) = POS_\pi(d)$ and $a \in Core_{B-\{b\}}(d)$. Thus $Core_B(d) \subseteq Core_{B-\{b\}}(d)$.□

4.3 Properties of inconsistent covering decision system based on limimtary information entropy

Lemma 1. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system. For $B, Q \subseteq \pi$ and $POS_B(d) = POS_Q(d)$, then $LI(B;d) = LI(Q;d)$.*
Proof. For every $X \in U/d$, by the definition of positive region, we get $POS_\pi(d) = \underset{X \in U/d}{\cup} \underline{B}(X)$. If every $y\ inU$, there exists $x_i \in U$ then $B_y \subseteq [x_i]_d \Leftrightarrow Q_y \subseteq [x_i]_d$. Since $\underline{B}(x) = \{x : \forall y \in U(x \in B_y \to B_y \subseteq X)\}$, $Q(X) = \{x : \forall y \in (x \in Q_y \to Q_y \subseteq X)\}$ and $POS_B(d) = POS_Q(d)$, thus we obtain $\underline{B}(X) = \underline{\pi}(X)$. Clearly, we get $LD(d|B) = LD(d|Q)$. By the definition of mutual information, we can obtain $I(B;d) = I(Q;d)$.□
Lemma 2. *Suppose $Icds = (U, \pi, d)$ is an inconsistent covering decision system. For every $B \subseteq Q \subseteq \pi$, if $LI(B;d) = LI(Q;d)$ is true, then $POS_B(d) = POS_Q(d)$.*
Proof. The same is true; here we will not prove it.□
Corollary 4. *Supposing $Icds = (U, \pi, d)$ be an inconsistent decision system. For every $B \subseteq Q \subseteq \pi$, if $LI(B;d) = LI(Q;d)$ is true if and only if $POS_B(d) = POS_Q(d)$.*

Theorem 12. *Suppose $Icds = (U, \pi, d)$ is an inconsistent covering decision system, for every $A_i \in \pi$, A_i is a superfluous element relative to d in π, if and only if $LH(d|\pi) = LH(d|\pi - \{A_i\})$.*
Proof. If A_i is a superfluous element relative to d in π, by the definition of positive region, we can get $POS_\pi(d) = POS_{\pi-\{Ai\}}(d)$. And by Lemma 2, we have $LI(\pi - \{A_i\}; d)$, thus $LH(d|\pi - \{A_i\}) = LH(d|\pi)$.

If $LI(\pi - \{A_i\}; d) = LH(d|\pi) \Rightarrow I(\pi; d) = I(\pi - \{A_i\}; d)$ and $\pi - \{A_i\} \in \pi$, by Theorem 10, we can obtain $POS_\pi(d) = POS_{\pi-\{Ai\}}$. Therefore, A_i is asuperfluous element relative to d in π.□
Corollary 5. *Supposing $Icds = (U, \pi, d)$ be an inconsistent decision system. For every $A_i \in \pi$, A_i is dispensable relative to d in π, if and only if $LH(d|C) < LH(d|\pi - \{A_i\})$.*
Corollary 6. *Supposing $Icds = (U, \pi, d)$ be an inconsistent decision system. For every $A_i \in \pi$, A_i is independent relative to d in π, if and only if $LH(d|C) < LH(d|\pi - \{A_i\})$.*
Theorem 13. *Suppose $Icds = (U, \pi, d)$ is an inconsistent decision system, for every $B, Q \subseteq \pi$, if $B \prec Q$, then $LH(d|B) > LH(d|Q)$.*
Proof. Suppose $Cov(B) = \{B_x : x \in U\}$ and $Cov(Q) = \{Q_x : x \in U\}$. If $B\ precQ$, there exists $B_x \subseteq Q_x$ for every $x \in U$.

Since

$$LH(d|B) - LH(d|Q) = -\frac{1}{|K|}\sum_{x\in K}\log\frac{|[x]_d \cap B_x|}{|B_x|} - \frac{1}{|K|}\sum_{x\in K}\log\frac{|[x]_d \cap Q_x|}{|Q_x|} = -\frac{1}{|K|^2}\sum_{x\in K}\log\frac{|[d]_d \cap B_x|}{|B_x|} \times \frac{|[x]_d \cap Q_x|}{|Q_x|} \tag{1}$$

If (1) met $B_x \subseteq [x]_d, Q_x \subseteq [x]_d$ and $B_x \subseteq Q_x$, then $LH(d|B) = LH(d|Q)$.
If (1) only met $B_x \subseteq Q_x$, we obviously can get $LH(d|B) > LH(d|Q)$.
Therefore, $LH(d|B) \geq LH(d|Q)$.□
Theorem 14. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system, if $B\ precQ$ for every $B, Q \subseteq \pi$, then $LI(B; d) \geq LI(Q; d)$.*
Proof. Here we will not prove it.□
Theorem 15. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system. For every $B, Q \subseteq \pi$, if $LH(d|B) > LH(d|Q)$ and $LH(Q|B) = 0$, then $B \prec 0$.*
Proof. Suppose $Cov(B) = \{B_x : x \in U\}$ and $Cov(Q) = \{Q_x : x \in U\}$. If $LH(d|B) > LH(d|Q)$ is true, then $B_x \supset Q_x$ must be wrong (If $B_x \supset Q_x$ is true, then $LH(d|B) \leq LH(d|Q)$ by Theorem 13, which is contradiction!), thus we could obtain $B_x \subseteq Q_x$ or $B_x \not\subset Q_x$. If $B_x \not\subset Q_x$ is true, there at least exists $B_{x1}(x_1 \in U)$, for every $Q_{y1}(y_1 \in U)$, then $B_{x1} \not\subset Q_{y1}$, if $B_x \cap Q_x \neq \emptyset$ is true for every $x \in U$, then $H(Q|B) = -\frac{1}{|K|}\sum_{x\in K}\frac{|Q_x \cap B_x|}{|B_x|} > 0$, which implies a contradiction, thus $B_x \subseteq Q_x$ is true.□
Similarly, we can prove the following theorem:
Theorem 16. *Suppose $Icds = (U, \pi, d)$ is an inconsistent covering decision system. If $LI(B; d) \geq LI(Q; d)$ and $LH(Q|B) = 0, B, Q \subseteq \pi$, then $B \prec Q$.*
From the description of Theorem13 through Theorem 16, we know that there exists corresponding relation for one to one between rough of knowledge and information entropy in inconsistent covering decision system.

4.4 Based on split policy inconsistent covering decision system for attribute reduction

Well-known scholars in Poland initially proposed discernibility matrix [11], or discernibility function can be used to calculate all attribute reduction in the decision table. Although the resolution matrix and its approach have been widely used, but due to the definition of resolution matrix, the data regarding the degree of inconsistency and its effects are not fully taken into account, so there are limitations. Literature [4, 20, 21] improved methods and discussed the case of inconsistent decision table, so that the former method can obtain the correct (all) attribute reduction results. Hence, such research is of great significance and, ultimately, a new application used in inconsistent decision tables to distinguish Matrices. Further, a method is proposed to distinguish matrix in the literature [31] based on the past, that is, split-based strategies and to distinguish Matrices decision table attribute reduction. Literature [19, 30] presents rough set theory, algorithms and applications, but also specifically pointed out that the resulting matrix method to distinguish demand for inconsistent decision tables is relatively simple errors in the nuclear, and also presented the results of a detailed analysis. Please refer to literature [19, 30].

Supposing $U1 = \sum_{X\ inU/d} \underline{\pi}(X)$ and $U2 = U - U1$

Definition 12. Suppose $Icds = (U, \pi, d)$ is an inconsistent covering decision system, all subjects on universe, U, are divided into consistent covering sub-table $\pi|U_1 \to d|U_1$ and inconsistent covering sub-table $\pi|U_2 \to d|U_2$. We denote a $x \times n$ matrix (m_{ij}), called the discernibility matrix of $Icds$, such that if $x_u, x_j \in U$ satisfies

$$\{A \in \pi : (A_{xi} \not\subset A_{xj}) \wedge (A_{xj} \not\subset A_{xi})\} \vee \{A_p \wedge A_q : (A_{pxi} \subset A_{pxj}) \wedge (A_{qxj} \subset A_{qxi})\}$$

Theorem 17. Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decisive system, denoted by $IM(\pi) = \{m_{ij}\}$ where m_{ij} is a single attribute, $IM(\pi) = Core_d(\pi)$ is true if and only if $\exists m_{ij} \in$ a single attribute such that $m_{ij} \in Core_d(\pi)$.

Proof. We denote $Cov(\pi - \{A_i\}) = \{B_x : x \in U\}$, if $A_i \in IM(\pi)$ is true, then there exists $m_{ij} = \{A_i\}$ for every $A_j \in \pi - \{A_i\}$ such that $A_{yj} = A_{xj}$ is true.

Suppose $x_j \in \pi([x_s]_d)(s \in [1, l])$, we will divide them to two parts for further discussion: first we prove $IM(\pi) \subseteq Core_d(\pi)$. If $y_j \in U1$ and $f_d(x_j) \neq f_d(y_j)$ are true, then there exists $t \in [1, l]$ and $t \neq s$ such that $y_j \in \underline{\pi}([x_y]_d)$ and $B_{xj} \cap \underline{\pi}([x_t]_d) \neq \emptyset$.

By Definition 12, A_i is indispensable. If $y_j \in U2$ is true, then $B_{xj} \cap U2 \neq \emptyset$ is true such that a_i is indispensable. So $IM(\pi) \subseteq Core_d(\pi)$ is true.

On the other hand, if $A_i \in Core_d(\pi)$ is true, then there exists $x_s \in \underline{\pi}([x_i]_d)(i \in [1, k])$ such that one of the following conditions is true:

(1) there exists $j \in [1, k]$ and $j \neq i$ such that $B_{xs} \cap [x_j]_d \neq \emptyset$;

(2) $B_{xs} \cap (U - \sum_{x \in U}([x]_d)) \neq \emptyset$

If Condition (1) is true, then there exists $x_t \in [x_j]_d$ and $A_j \in \pi - \{A_i\}$ such that $A_{xsj} = A_{xij}$ and $A_{xsi} \neq A_{xti}$, and because of $x_s \in ([x_i]_d), x_t \in ([x_t]_d)$, then we can obtain $f_d(x_s) \neq f_d(x_t)$ for $x_s \in U1, x_t \in U2$ such that $m_{st} \in \{A_i\}$.

If Condition (2) is true, then there exists $x_r \in U2$, for every $A_j \in \pi - \{A_i\}$ such that $A_{xsj} = A_{xrj}$ and $A_{xsi} \neq A_{xri}$ for $x_s \in U1, x_y \in U2$, therefore we could get $m_{sr} = \{A_i\}$. So $Core_d(\pi) \subseteq IM(\pi)$ is true.

Theorem 18. *Supposing $Icds = (U, \pi, d)$ be an inconsistent covering decision system, denoted by subset $E = \pi - Core(\pi), E_1 \subseteq E$, then $POS_\pi(d) = POS_{\pi - \{E_1\}(d_1)}$ is true if and only if for every $m_{ij}(1 \leq j \leq i \leq n)$ such that $m_{ij} \not\subset E_1$ is true.*

Proof. If there exists $E_1 \subseteq E$ such that any element of E_1 does not belong to $Core_d(\pi)$, so E_1 can be reduced. If for every m_{ij} and $1 \leq j \leq i \leq n$, then we could get $m_{ij} \not\subset E_1$, which still ensure $POS_\pi(d) = POS_{\pi - \{E_1\}}(d_1)$ after reduction of properties on E_1 at the same time.

If there exists m_{ij} after reduction by properties on E_1 contemporarily such that m_{ij} and $m_{ij} \neq \emptyset$ is true, which implies properties on E_1 cannot be reduced at the same time.

Theorem 18 implies that as long as there is a simple observation and treatment for discernibility matrix, we will have cores and reducts in the inconsistent covering decision system. The following Corollary 7 can be founded in the literature [21, 25].

Corollary 8. Suppose $B \subseteq \pi$, then B is a relative reduct of π if and only if it is the minimal set satisfying $B \cap m_{ij} \neq \emptyset$ for $m_{ij} \neq \emptyset, i, j \leq n$.

5 Experimental analysis: a test application

Example 1. Here is a car which is to be considered for analysis. Suppose $U = \{x_1, \ldots, x_{10}\}$ to be a set of ten cars, and $E =$ {price; colour; quality; oil-consumption} to be a set of attributes. The values of 'price' are {*high; middle; low*}, the values of 'colour' are {*pretty; ordinary; poor*}, the values of 'quality' are {*good; bad*}, the values of 'oil-consumption' are {tiny; relative-tiny; reasonable; numerous}. We have four specialists

$E = \{A, B, C, D\}$ to evaluate the attributes of these cars. Moreover, their evaluation results are not the same when compared with one another. The evaluation results are listed below.
For attribute *price*:

$$A: high = \{x_1, x_2, x_4, x_6, x_7, x_8, x_9, x_{10}\},\ middle = \{x_3\},\ low = \{x_5\};$$

$$B: high = \{x_1, x_2, x_3, x_5, x_6, x_8, x_9, x_{10}\},\ middle = \{x_4\},\ low = \{x_6, x_7\};$$

$$C: high = \{x_1, x_2, x_3, x_4, x_8, x_9, x_{10}\},\ middle = \{x_8\},\ low = \{x_3, x_4, x_5, x_6, x_9\};$$

$$D: high = \{x_1, x_2, x_3, x_5, x_6, x_8, x_9, x_{10}\},\ middle = \{x_7\},\ low = \{x_5, x_6\};$$

For attribute *color*:

$$A: pretty = \{x_1, x_2, x_3, x_4, x_5\},\ ordinary = \{x_6, x_7, x_8, x_9\},\ low = \{x_{10}\};$$

$$B: pretty = \{x_1, x_2, x_3, x_4, x_5, x_6\},\ ordinary = \{x_7, x_8, x_9\},\ low = \{x_{10}\};$$

$$C: pretty = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\},\ ordinary = \{x_8, x_9\},\ low = \{x_{10}\};$$

$$D: pretty = \{x_1, x_2, x_3, x_4, x_5, x_7\},\ ordinary = \{x_6, x_8, x_9\},\ low = \{x_{10}\};$$

For attribute *quality*:

$$A: good = \{x_1, x_2, x_3, x_6, x_8, x_{19}\},\ bad = \{x_4, x_5, x_7, x_9\};$$

$$B: good = \{x_1, x_2, x_3, x_8, x_{10}\},\ bad = \{x_4, x_5, x_6, x_7, x_9\};$$

$$C: good = \{x_1, x_6, x_8, x_9, x_{10}\},\ bad = \{x_2, x_3, x_4, x_5, x_7\};$$

$$D: good = \{x_1, x_2, x_6, x_8, x_{10}\},\ bad = \{x_3, x_4, x_5, x_7, x_9\}$$

For attribute *oil-consumption:*

$$A: tiny = \{x_1, x_2\},\ relative - tiny = \{x_3, x_4, x_5\},\ reasonable = \{x_3, x_4, x_5\},\ numerous = \{x_7, x_9\};$$

$$B: tiny = \{x_1, x_3\},\ relative - tiny = \{x_2, x_4, x_5, x_6\},\ reasonable = \{x_7, x_8, x_{10}\},\ numerous = \{x_9\};$$

$$C: tiny = \{x_1, x_3\},\ relative - tiny = \{x_2, x_4, x_5, x_7\},\ reasonable = \{x_6, x_8, x_{10}\},\ numerous = \{x_9\};$$

$$D: tiny = \{x_1, x_2, x_6\},\ relative - tiny = \{x_3, x_4, x_5\},\ reasonable = \{x_8, x_9, x_{10}\},\ numerous = \{x_7, \};$$

We think that the evaluation of every index is has the same importance. Therefore, we get a cover rather than a partition for every car attribute, which implies a certain uncertainty caused by the interpretation of the data.

$$price:\ A_1 = \{\{x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_9, x_{10}\}, \{x_3, x_4, x_6, x_7\}, \{x_3, x_4, x_5, x_6, x_7\},$$

$$colour:\ A_2 = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}, \{x_6, x_7, x_8, x_9\}, \{x_{10}\}\},$$

$$quality:\ A_3 = \{\{x_1, x_2, x_3, x_6, x_8, x_9, x_{10}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_9\}, \}$$

$$oil - consumption:\ A_4 = \{\{x_{1,2}, x_3, x_6\}, \{x_2, x_3, x_4, x_5, x_6, x_7\}, \{x_6, x_8, x_9, x_{10}\}, \{x_6, x_7, x_9\}\}$$

Final decision d is given as $U/d = \{$sale; further evaluation for sale; against sale$\}$; sale=$\{x_1, x_2, x_4, x_6\}$, further evaluation for sale=$\{x_4, x_5, x_7\}$, against sale=$\{x_1, x_2, x_4, x_6\}$. Suppose $\pi = \{A_i : i = 1, \dots, 4\}, \pi_{xi}$, abridges π_i, B_i means B_{xi} for short, then we can obtain:

$$\pi_1 = \{x_1, x_2, x_3, x_6\}, \pi_2 = \{x_3, x_2, x_6\}, \pi_3 = \{x_3, x_6\}$$

$$\pi_4 = \{x_3, x_4, x_6, x_7\}, \pi_5 = \{x_3, x_4, x_5, x_7\}, \pi_6 = \{x_6\}, \pi_7 = \{x_6, x_7\}$$

$$\pi_8 = \{x_6, x_8, x_9\}, \pi_9 = \{x_6, x_9\}, \pi_{10} = \{x_{10}\}.$$

The positive domain of d relative to π is:

$$POS_\pi(D) = \bigcup_{x \in U/D} \underline{\pi}(X) = \{x_1, x_2, x_3, x_6, x_{10}\} \tag{1}$$

Supposing $B = \pi - A_1$, then $B_1 = \{x_1, x_2, x_3, x_6\}$, $B_2 = B_3 = \{x_2, x_3, x_6\}$, $B_4 = B_5 = \{x_2, x_3, x_4, x_5, x_6, x_7\}$, $B_6 = \{x_6\}$, $B_7 = \{x_6, x_7\}$, $B_8 = \{x_6, x_8, x_9\}$, $B_9 = \{x_5, x_9\}$, $B_{10} = \{x_{10}\}$, the positive domain of d relative to B is:

$$POS_B(D) = \bigcup_{x \in U/D} \underline{B}(X) = \{x_1, x_2, x_3, x_6, x_{10}\} \tag{2}$$

As to (1) and (2), we could obtain $POS_B(D) = POS_\pi(D)$. By the Definition 12, A_i is a superfluous relative to d in π. Here we can see

$$\pi_{x1} \not\subset \pi_{x4}, \pi_{x4} \not\subset \pi_{x1} \Longrightarrow B_{x1} \not\subset B_{x4}, B_{x4} \not\subset B_{x1},$$

and

$$\pi_{x2} \not\subset \pi_{x4}, \pi_{x4} \not\subset \pi_{x2} \Longrightarrow B_{x2} \subset B_{x4}, \pi_{x3} \subset \pi_{x4} \Longrightarrow B_{x3} \subset B_{x4}.$$

The uppermost relation of x_2, x_4 to π changes after A_i is deSupposinged from π. A_1 is a superfluous element of π relative to D, so it is an exemplar of inconsistent covering decision system.
Suppose $U1 = \{x_1, x_2, x_3, x_6, x_{10}\}$ and $U2 = \{x_4, x_5, x_7, x_8, x_9\}$. By the Definition 12, we have the discernibility matrix of inconsistent covering decision system (U, π, D), which is follows (covers have been distinguished i instead of A_i, otherwise 0 instead of π):

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \{2,4\} & \{3,4\} & \{1,3,4\} & \{3,4\} & \{2,4\} & \{2,4\} \\ 0 & 0 & 0 & 0 & \{2,4\} & \{1\wedge 3, 2\wedge 3, 3, 4\} & \{1,3,4\} & \{2\wedge 3, 1\wedge 3, 3, 4\} & \{2,4\} & \{2,4\} \\ 0 & 0 & 0 & 0 & \{2,4\} & \{3,4\} & \{1,3,4\} & \{2\wedge 3, 3, 4\} & \{1,2,4\} & \{1,2,4\} \\ 0 & 0 & 0 & 0 & \{2,3,4\} & \{3,4\} & \{1,2,3,4\} & \{3,4\} & \{1,2,4\} & \{1,2,4\} \\ \{2,4\} & \{2,4\} & \{2,4\} & \{2\} & 0 & \{2,3,4\} & \{1,2,3,4\} & \{2,3,4\} & \{2\} & \{2\} \end{bmatrix}$$

and $f(U,\pi)(\bar{A}_1, \bar{A}_2, \bar{A}_3, \bar{A}_4) = \wedge\{\vee m_{ij} : 1 \le j < i \le 10, m_{ij} \ne \emptyset\} = (A_1 \vee A_3 \vee A_4) \wedge (A_2 \vee A_4) \wedge (A_3 \vee A_4) \wedge (A_2 \vee A_4) \wedge (A_1 \vee A_3 \vee A_4) \wedge ((A_1 \vee A_3) \vee (A_2 \vee A_3) \vee A_3 \vee A_4) \wedge (A_1 \vee A_2 \vee A_4) \wedge ((A_2 \wedge A_3) \vee A_3 \vee A_4) \wedge (A_1 \vee A_3 \vee A_4) \wedge (A_2 \vee A_3 \vee A_4) \wedge (A_3 \vee A_4) \wedge (A_1 \vee A_2 \vee A_3 \vee A_4) \wedge \{A_2\} = (A_2 \wedge A_4) \vee (A_2 \wedge A_3)$, so

$$Red(\pi) = \{\{A_3, A_4\}, \{A_2, A_3\}\},\ Core_D(\pi) = \{A_2\}.$$

It should be pointed out that if the covering decision system is consistent, then the method proposed in this section is equivalent to the one in Section 4. If is a partition, then the method adopted in this section is just the method for computing relative reducts of traditional rough sets in the literature [32] to ensure that we find the smallest reduction.
If these ten cars are trial samples, then we have two different kinds of evaluation references for other input samples: {colour; oil-consumption}, {colour; quality}. Clearly, the attributeis the key attribute for the evaluation of cars.
To illustrate the methods of space and computational complexity in the section, we will compare our methods with the methods of literature [21, 23, 24], such that if $n = |U_1| + |U_2|$ satisfies,
(1) The space complexity can be compared: without considering compression and storage of symmetric matrix, the elements of discernibility matrix in the section are $|U_1| \times n$, while the elements in the literature [21] are $n \times n$.
(2) The computational complexity can be compared: the computational complexity in literature [21, 23, 24] is $O(m \times n \times \log^n) + O(m \times n^2)$, while the method in this section is $O(m \times n \times \log^n) + O(m \times |U_1| \times n)$.
It is evident that the space and computational complexity in the section are lower than the literature [21].Therefore, the methods in this section could be used effectively not only to reduce the computational cost, but also in providing a new framework to certain extent based on the covering rough sets theory.

6 Conclusions

According to classical rough sets theory, attributes of decision systems consist of two parts namely conditional attributes and decision attributes. Every conditional attribute decide a partition in a complete decision system. The abstract information systems which come from reality problems are mostly incomplete decision system. Every conditional attribute in this decision system determines a cover of U. This paper mainly studies theories and methods of systems and discusses about related attribute reduction for covering decision information reduction algorithms on the basis of conditional information entropy. Moreover, attribute reduction of covering decision systems also have wide applications in the three-way, which indicates the importance of the direction of research currently.

References

[1] Z. Pawlak. Rough sets. Computer and Information Sciences 11(1982) 341-356.
[2] H. Yu, D.C. Yang. Decision Table Reduction based on Conditional Information Entropy 07(2002) 759-756.
[3] J.N. Mordeson, Rough set theory applied to (fuzzy) ideal theory, Fuzzy Sets and Systems 121(2001) 315-324.
[4] G. Cattaneo. Abstract approximate spaces for rough theories[M], in: Polkowski,Skowron (Eds.), Rough Sets in Knowledge Discovery1: Methodology and Applications, Physicaverlaf, Heidelberg, 1998, pp. 59-98.
[5] H.S. Nguyen, D. Slezak. Approximation reducts and association rules correspondence and complexity results, in: N. Zhong, A. Skowron, S. Oshuga (Eds.), Proceedings of RSFDGrC'99[C]. Japan: LNAI1711, 1999.
[6] M. Yang. Approximate Reduction Based on Conditional information Entropy in Decision Tables. Acta Electronica Sinica 11(2007) 2156-2160.
[7] Z.W. Mo. Some Notes on Entropy and Rough Set Theory. Journal of Sichuan Normal University (Natural Science) 03(2002) 257-260.
[8] P. Yang. Approximate reduction based on conditional information entropy over vertically partitioned multi-decision table. Control and Decision 10(2008) 1103-1108.
[9] Y. Wang. A Knowledge Reduction Method Based on Information Entropy and Discernibility Matrix. Journal of Sichuan College of Education 3(2008) 07-10.
[10] Z. Pawlak. Rough sets: Theoreticsl Aspects of Reasoning About Data[M]. Dordrecht: Kluwer Academic Publishing, 1991.
[11] Z. Pawlak, Andrzej Skowron. Rudiments of rough sets, Information Sicences 177(2006) 3-27.
[12] Z. Pawlak, Andrzej Skowron. Rough sets and Boolean reasoning. Information Sciences 177(2006) 28-40.
[13] W. Ziarko. Variable precision rough set model. Computer and System Sciences 46(1993) 39-59.
[14] W.-X. Zhang, W.-Z. Wu, J.-Y. Liang, D.-Y. Li. Theory and method of rough sets[J]. Beijing: Science Press, 2001.
[15] B. Huang, X. He, X.Z. Zhou. Rough Entropy Based on Generalized Rough sets Covering Reduction. Journal of software 15(2004) 215-218.
[16] D.G. Chen, C.Z. Wang Chang, Q.H. Hu. A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets. Information Sciences 177(20-07) 3500-3518.
[17] Z. Bonikowski, E. Bryniarski, U. Wybraniec, Extensions and intensions in the rough set theory, Information Sciences 107(1998) 149-167.
[18] C.D. Yang. Research on decision systems reduction with rough sets. Harbin Engineering University 2011, PP:1-120.
[19] H. Guo. Research on knowledge reduction based on rough set theory for the inconsistent decision systems 2016, pp:1-150.
[20] Y.P. Shi, Y. Huang, C.Z. Wang, Q. He. Attribute reduction based on the Boolean matrix. Granular Computing 4(2019) 313-322.
[21] F. Li, Y.Q. Yin. Approaches to knowledge reduction of reduction of covering decision systems based on information theory. Information Sciences 179(2009) 694-1704.
[22] D.L. Ma. The construction of decision tree based on the covering rough set theory[J]. Beijing: North China Electric University 2016:1-41.
[23] M.P. Chen, et al. Multi-label attribute reduction algorithm based on neighborhood rough set. Journal of Minnan Normal University (Natural Science) 31(2018) 1-11.
[24] Y.L. Zhang, et al. Attribute reduction of covering decision information system based on evidence theory. Pattern recognition and artificial intelligence 31(2018) 797-808.
[25] X. Zhang, et al. Algorithms of rule acquisition and confidence-preserved attribute reduction in covering decision systems. Journal of Shangdong University (Natural Science) 53(2018)120-126.
[26] M.R. Forster, Key concepts in model selection: performance and generalizabilitu, J. Math. Psychol. 44(2000) 205-231.

[27] P.W. Woodward, J.C. Naylor, An application of Bayesian methods in SPC, Statisitician 42(1993) 461-469
[28] H. Yu, Z.G. Liu, G. Y. Wang. An automatic method to determine the number of clusters using decision-theoretic rough sets, Int. J. Approx. Reason 55(2014) 101-115.
[29] B. Zhou, Y.Y. Yao. J.G. Luo, Cost-sensitive three-way email spam filtering, J. Intell. Inf. syst 42(20-14) 19-45.
[30] Z.H. Jiang, X.B. Yang, H.L. Yu, D. Liu, P.X. Wang, Y.H. Qia. Accelerator for multi-granularity attribute reduction. Knowledge-Based Systems 04(2019) 145-158.
[31] Y.L. Zhang, J.J. Li. Relative reduction of covering decision systems. Journal of engineering mathematics 05(2009) 930-935.
[32] Lech Polkowski, Shusaku Tsumoto, Tsau Y. Lin. Rough set methods and applicaftions: newdevelopments in knowledge discovery in information systems[M]. NewYork: Physical-Verlag, 2000.