To appear in the *International Journal of General Systems*
Vol. 00, No. 00, January 2013, 1–19

Decision-relative discernibility matrixes in the sense of entropies

WEI WEI†, JIYE LIANG†*, JUNHONG WANG†, YUHUA QIAN†
†*School of Computer and Information Technology, Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006, China*

(*Received 00 Month 201X; final version received 00 Month 201X*)

In rough set theory, attribute reduction is a basic issue, which aims to hold the discernibility of the attribute set. To obtain all of the reducts of an information system or a decision table, researchers have introduced many discernibility matrixes based reduction methods. However, the reducts in the sense of positive region can only be obtained by using the existing discernibility matrixes. In this paper, we introduce two discernibility matrixes in the sense of entropies (Shannon's entropy and complement entropy). By means of the two discernibility matrixes, we can achieve all of the reducts in the sense of Shannon's entropy and all of the reducts in the sense of complement entropy, respectively. Furthermore, we discover the relationships among the reducts in the sense of preserving positive region, Shannon's entropy and complement entorpy. The experimental studies show that by the proposed decision-relative discernibility matrices based reduction methods, all the reducts of a decision table in sense of entropies can be obtained.

Keywords: Rough sets; discernibility matrix; entropy; attribute reduction

1. Introduction

Rough set theory proposed by Pawlak (Pawlak, 1991; Pawlak and Skowron, 2007a,b) is a new soft computing tool for the analysis of a vague description of an object, and has become a popular mathematical framework for pattern recognition, image processing, data mining and knowledge discovery (Bazan *et al.*, 2003; Duntsch *et al.*, 1998; Guan and Bell, 1998; Nguyen, 2006; Wei *et al.*, 2012). Attribute reduction plays an essential role in analyzing an information table (Pawlak, 1991; Pawlak and Skowron, 2007b). An attribute reduct is a minimal subset of attributes sets that provides the same descriptive ability as the entire set of attributes (Yao *et al.*, 2006). As a consequence, to acquire more general and brief decision rules from inconsistent systems, attribute reduction is needed.

Recently, more attention has been focused on the area of attribute reduction, and many methods of attribute reduction have been developed in rough set theory (Hu and Cercone, 1995; Kryszkiewicz, 2001; Li *et al.*, 2004; Mi *et al.*, 2004; Nguyen and Slezak, 1999; Qian *et al.*, 2010a, 2009, 2008; Shao and Zhang, 2005; Skowron and Rauszer, 1992; Ślęzak, 1996; Wu *et al.*, 2005; Yang, 2006; Yamaguchi, 2009; Yao and Zhao, 2009; Ye and Chen, 2002; Ziarko, 1993). Among them, the discernibility matrix based attribute reduction is one of important methods for obtaining

*Corresponding author. Email: ljy@sxu.edu.cn.

reducts. By means of this method, we can get all the reducts of an information systems (Skowron and Rauszer, 1992). Hu et al. (Hu and Cercone, 1995) modified the discernibility matrix for computing the reducts of a decision table (called the decision-relative discernibility matrix). Ye and Chen (Ye and Chen, 2002) pointed out that the decision-relative discernibility matrix in (Hu and Cercone, 1995) is only used to obtain the reducts of the consistent decision tables, and presented a discernibility matrix suitable for the consistent and inconsistent decision tables. Furthermore, Yang (Yang, 2006), by dividing the universe into the consistent part and inconsistent part, proposed another decision-relative discernibility matrixes by which the consuming time for computing the reducts is significantly reduced.

For convenience of our further discussion, we also review three types of attribute reductions divided from the viewpoint of the discernibility power of reducts: (1) Positive region reduction. For a decision table reduced by this method, the positive region of a target decision with respect to the obtained reducts is same as with respect to the original attribute set (Hu and Cercone, 1995; Pawlak, 1991). (2) Shannon's entropy reduction. For a decision table reduced by this method, the Shannon's condition entropy of the decision attribute set with respect to the reduced condition attribute set is same as the one before attribute reduction. (Ślęzak, 2000, 2002; Wang, 2003). (3) Complement entropy reduction. For a decision table reduced by this method, the complement condition entropy of the decision attribute set with respect to the reduced condition attribute set is same as the one before attribute reduction (Liang *et al.*, 2004, 2002; Wei *et al.*, 2010, 2009).

All of the positive region reducts of a decision table can be obtained by means of the existing decision-relative discernibility matrixes. However, the reducts in the sense of entropies are also significant for decision making. To obtain all of the Shannon's entropy and the complement entropy reducts of a decision table, in this paper, we propose two new discernibility matrixes in the sense of entropies. Furthermore, we discover the inherent relationships among the reducts in the sense of positive region, Shannon's entropy and complement entropy.

The rest of the paper is organized as follows. Some preliminary concepts are reviewed in Section 2. In Section 3, two kinds of discernibility matrix based on entropies are introduced and the relationships among the different types of reducts are discovered. Section 4 demonstrates effectiveness of the proposed discernibility matrix by experiments. Section 5 concludes the paper.

2. Preliminaries

In this section, we review some basic concepts such as indiscernibility relation, discernibility matrix and attribute reduction.

An information table is a 4-tuple $S = \{U, A, V, f\}$ (for short $S = \{U, A)$), where U is a non-empty and finite set of objects, called a universe, and A is a non-empty and finite set of attributes, V_a is the domain of the attribute a, $V = \bigcup_{a\in A} V_a$ and $f : U \times A = V$ is a function ($f(x, a) \in V_a$ for each $a \in A$).

An indiscernibility relation $R_B = \{(x, y) \in U \times U \mid f(x, a) = f(y, a), \forall a \in B\}$ is determined by a attribute set $B \subseteq A$. $U/R_B = \{[x]_B \mid x \in U\}$ (just as U/B) indicates the partitions of U resulted from R_B, where $[x]_B$ denotes the equivalence class determined by x with respect to B, i.e., $[x]_B = \{y \in U \mid (x, y) \in R_B\}$.

Furthermore, for any $Y \subseteq U$, $(\overline{B}(Y), \underline{B}(Y))$ is defined as the rough set of Y with respect to B, where the lower approximation $\underline{B}(Y)$ and the upper approximation

$\overline{B}(Y)$ of Y are indicated by

$$\underline{B}(Y) = \{x|[x]_B \subseteq Y\},$$
$$\overline{B}(Y) = \{x|[x]_B \cap Y \neq \emptyset\}.$$

A partial relation $\preceq$ on $\{U/Q \mid Q \subseteq P\}$ is introduced as follows: $U/P \preceq U/Q$ if and only if, for every $P_i \in U/P$, there exists $Q_j \in U/Q$ such that $P_i \subseteq Q_j$. Thus, Q is coarser than P. And If we say Q is strictly coarser than P, then this relation is denoted by $U/P \prec U/Q$.

For an information system $S = (U, C \cup D)$, if C is a condition attribute set, D is a decision attribute set, $C \cap D = \emptyset$, then it is called a decision table. Moreover, this table is said to be consistent if $U/C \preceq U/D$, otherwise it's inconsistent. The objects in the condition classes which deduce certain decision rules constitute the consistent part of a decision table, and the objects in other condition classes constitute the inconsistent part of the decision table.

Pawlak (Pawlak, 1991) proposed the definition of a decision rule as follows: Z_{ij} : $des(X_i) \rightarrow des(Y_j)$, where $des(X_i)$ and $des(Y_j)$ denote the descriptions of the equivalence classes $X_i \in U/C$ and $Y_j \in U/D$. The certainty degree of a decision rule Z_{ij} is defined as: $\mu(Z_{ij}) = |X_i \cap Y_j|/|X_i|$, where $|\cdot|$ is the cardinality of a set.

Skowron and Rauszer (Skowron and Rauszer, 1992) introduced a matrix representation for storing the sets of attributes that discern pairs of objects, called a discernibility matrix.

Definition 1. *Let $S = (U, C)$ be an information system, the discernibility matrix of the information system S is $n \times n$ matrix, which is defined as $M^I = \{m_{ij}\}$, where $m_{ij} = \{c \in C| \ f(x_i, c) \neq f(x_j, c)\}$*

Based on the discernibility matrix, Skowron (Skowron and Rauszer, 1992) defined a discernibility function as follows.

Definition 2. *Let $S = (U, C)$ be an information system, the discernibility matrix $M^I = \{m_{ij}\}$, $m_{ij} = \{c \in C| \ f(x_i, c) \neq f(x_j, c)\}$. Then the discernibility function of a discernibility matrix is defined as*

$$f(M^I) = \bigwedge \left\{ \bigvee (m^I_{ij}) \mid \forall x_i, x_j \in U, m^I_{ij} \neq \emptyset \right\}.$$

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

The discernibility function $f(M^I)$ describes constraints which must hold to preserve discernibility between all pairs of discernible objects from S. It requires keeping at least one attribute from each non-empty entry of the discernibility matrix corresponding to any pair of discernible objects. It can be shown that for an information system $S = (U, A)$, the set of all prime implicants of $f(M^I)$ determines the set of all reducts of S.

To compute the reducts of a decision table, Hu and Cercone (Hu and Cercone, 1995) extend the definitions of discernibility matrix and discernibility function as follows.

Definition 3. *Let $S = (U, C \cup D)$ be a consistent decision table, $D = \{d\}$. Then* *the decision-relative discernibility matrix is defined as* $M^D = \{m'_{ij}\}$, *where*

$$m'_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d) \\ m_{ij},\ otherwise \end{cases}.$$

From the definition, we can see that if the values of two objects on the decision attribute is equal, then the two objects need not be differentiated with each other; Otherwise, the two objects need be distinguished by through their values on conditional attribute set.

Ye (Ye and Chen, 2002) pointed out that the above decision-relative discernibility matrix is only suitable for the consistent decision tables.

Definition 4. *Let* $S = (U, C \cup D)$ *be a decision table, then Ye's discernibility matrix of decision table* S *is defined as* $M^{Ye} = \{m^{Ye}_{ij}\}$, *where*

$$m^{Ye}_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ \min\{d(x_i), d(x_j)\} = 1 \\ \emptyset,\ otherwise \end{cases},$$

and $d(x_i) = |\{f(x_k, d), x_k \in [x_i]_C\}|$.

Ślęzak (Ślęzak, 2002) proposed another decision-relative discernibility matrix, which is constructed on the basis of the rough membership function. The matrix is defined as follows.

Definition 5. *Let* $S = (U, C \cup D)$ *be a decision table,* C *the condition attribute set and* D *the decision attribute set,* $D = \{d\}$. *Then the discernibility matrix is defined as* $M^{\mu} = \{m^{\mu}_{ij}\}$, *where*

$$m^{\mu}_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ \mu_{ik} \neq \mu_{jk} \\ \emptyset,\ otherwise \end{cases},$$

where $\mu_{ik} = \frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|}$, $\mu_{jk} = \frac{|[x_j]_C \cap Y_k|}{|[x_j]_C|}$, $[x_i]_C, [x_j]_C \in U/C$ *and* $Y_k \in U/D$.

Yang (Yang, 2006) modified Ye's discernibility matrix, and introduced a new discernibility matrix as follows.

Definition 6. *Let* $S = (U, C \cup D)$ *be a decision table,* C *the condition attribute set,* D *the decision attribute set,* $D = \{d\}$. *Then Yang's discernibility matrix is defined as* $M^{Yang} = \{m^{Yang}_{ij}\}$, *where*

$$m^{Yang}_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d)\ and\ x_i, x_j \in U_1 \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2 \\ \emptyset,\ otherwise \end{cases},$$

U_1 *is the consistent part of the decision table* S *and* U_2 *is the inconsistent part of the decision table* S.

According to Definition 2.3, two important results are obtained as follows: (1) If the two values of the objects in a decision table's consistent part is different, then the condition attributes on which their values are not equal should appear in the corresponding entry in Yang's discernibility matrix. This result indicates that the two objects in the consistent part can be distinguished by the reducts derived from Yang's discernibility matrix. (2) If one of the two objects is in the consistent part of a decision table and the other is in its inconsistent part, then the condition attributes on which their values are different is also in the corresponding entry of Yang's discernibility matrix. The fact implies that we can distinguish one object in a decision table's consistent part and the other object in inconsistent part by means of the reducts derived from the discernibility matrix.

Furthermore, based on the definition of Yang's discernibility matrix, the corresponding discernibility function is defined as:

$$f(M^{Yang}) = \bigwedge \left\{ \bigvee (m_{ij}^{Yang}) \mid \forall x, y \in U, m_{ij}^{Yang} \neq \emptyset \right\}.$$

By means of the similar idea in (Skowron and Rauszer, 1992), the set of all prime implicants of $f(M^{Yang})$ determines the set of all of the positive region reducts defined as follows.

Definition 7. *(Hu and Cercone, 1995) Let $S = (U, C \cup D)$ be a decision table and $B \subseteq C$. B is a positive region reduct of D with respect to C if B satisfies the following conditions:*

$$\begin{aligned} &(1) \quad POS_C(D) = POS_B(D);\ and \\ &(2) \quad for\ \forall b \in B, \quad POS_B(D) \neq POS_{B-\{b\}}(D). \end{aligned}$$

where $POS_C(D) = \bigcup_{i=1}^{n} \underline{C}Y_i$ and $Y_i \in U/D$.

To illustrate how to obtain the reducts by means of the Yang's discernibility matrix, we present the following example.

Example 1 Table 1 is a decision table about diagnosing rheum, in which $U = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}\}$ is the universe, $C = \{c_1(Headache)$, $c_2(Muscle\ pain)$, $c_3(Animal\ heat)$, $c_4(Cough)\}$ is the condition attribute set, and $D = \{d(Rheum)\}$ is the decision attribute set.

The discernibility matrix of Table 1 is given as follows:

$$M^{Yang} = \begin{pmatrix} \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_4\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_2, c_4\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_2, c_3\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_2\} \\ \{c_1, c_4\} & \{c_1, c_2, c_4\} & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_2, c_3\} & \emptyset & \{c_1, c_2\} & \emptyset \end{pmatrix},$$

Furthermore, the corresponding discernibility function is

$$\begin{aligned} f(M^{Yang}) &= \{c_1 \vee c_4\} \wedge \{c_1 \vee c_2 \vee c_4\} \wedge \{c_1 \vee c_2 \vee c_3\} \wedge \{c_1 \vee c_2\} \\ &= \{c_1 \vee c_4\} \wedge \{c_1 \vee c_2\} \\ &= \{c_1\} \vee \{c_2 \wedge c_4\}. \end{aligned}$$

Therefore, we obtain that $\{c_1\}$ and $\{c_2, c_4\}$ are all the positive region reducts of Table 1.

3. Decision-relative discernibility matrixes in the sense of entropies

In this section, we will introduce two types of new discernibility matrixes in the sense of entropies. They are the discernibility matrix in the sense of Shannon's entropy and complement entropy respectively.

To present the discernibility matrix in the sense of Shannon's entropy, we review the definition of Shannon's entropy.

Given a decision table $S = (U, C \cup D)$, C is the condition attribute set, D is the decision attribute set. Then, Shannon's condition entropy of D with respect to C is defined as

$$H(D|C) = -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|X_i \cap Y_j|}{|X_i|} \log_2 \frac{|X_i \cap Y_j|}{|X_i|},$$

where $X_i \in U/C$ and $Y_j \in U/D$.

Theorem 1. *(Gallager, 1968; Ślęzak, 2000, 2002) Let* $S = (U, C \cup D)$ *and* $S' = (U, B \cup D)$ *be two decision tables,* $B \subset C$, $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots, X_{v-1}, X_{v+1}, \cdots, X_m, X_u \cup X_v\}$ *and* $U/D = \{Y_1, Y_2, \cdots, Y_n\}$, *then*

$$H(D|B) \geq H(D|C),$$

especially, $H(D|B) = H(D|C)$, *if and only if* $\mu(Z_{uj}) = \mu(Z_{vj})$ *for* $j \leq n$, *where* $\mu(Z_{uj}) = \frac{|X_u \cap Y_j|}{|X_u|}$ *and* $\mu(Z_{vj}) = \frac{|X_v \cap Y_j|}{|X_v|}$.

From Theorem 1, it is easy to know that Shannon's conditional entropy is unchanged by combining two condition classes, if the certain degrees of the rules derived from them are equal. Otherwise, the combination of two condition classes will cause the change of Shannon's condition entropy.

Definition 8. *Let* $S = (U, C \cup D)$ *be a decision table,* C *condition attribute set,* D *decision attribute set,* $B \subseteq C$. *If* B *satisfies the following conditions:*

$$\begin{aligned} &(1) \quad H(D|C) = H(D|B) \, and \\ &(2) \quad for \, \forall b \in B, \quad H(D|B) \neq H(D|B - \{b\}), \end{aligned}$$

then B *is a relative reduct of* C *in the sense of Shannon's entropy.*

Remark: By Theorem 1 and Definition 8, we obtain that the entries in the discernibility matrix in the sense of Shannon's entropy should satisfy the following conditions: (1) If the decision values of two objects in the consistent part of a decision table are different, then the condition attributes on which the values of the two objects are unequal is in the corresponding entries of a decision table's discernibility matrix. (2) If one object is in the consistent part of a decision table and the other is in its inconsistent part, the condition attributes on which the values of the two objects are different is in the corresponding entries of a decision table's discernibility matrix. (3) If both of two objects are in the inconsistent part of a decision table and the certain degrees of the rules derived from the two objects are equal, then the condition attributes on which the values is not equal is in the corresponding entries. The first two conditions are the same as the ones of

Yang's discernibility matrix, and the last condition is a special one required by the property of Shannon's condition entropy.

From these above analyses, we introduce the discernibility matrix in the sense of Shannon's entropy is given as follows.

Definition 9. *Let $S = (U, C \cup D)$ be a decision table, C the condition attribute set, D the decision attribute set and $D = \{d\}$. The discernibility matrix in the sense of Shannon's entropy is defined as $M^S = \{m^S_{ij}\}$, where*

$$m^S_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d) \ \textit{and}\ x_i, x_j \in U_1 \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2 \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ \mu_{ik} \neq \mu_{jk}\ \textit{for}\ \forall Y_k \in U/\{d\},\ \textit{and}\ x_i, x_j \in U_2 \\ \qquad\qquad \emptyset,\ \textit{otherwise} \end{cases}$$

where $\mu_{ik} = \frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|}$, $\mu_{jk} = \frac{|[x_j]_C \cap Y_k|}{|[x_j]_C|}$, $[x_i]_C \in U/C$ *and* $[x_j]_C \in U/C$.

By means of the definition of the discernibility matrix based on Shannon's entropy, the corresponding discernibility function will be defined as follows.

Definition 10. *The discernibility function based on is defined as*

$$f(M^S) = \bigwedge \left\{ \bigvee (m^S_{ij}) \mid \forall x, y \in U, m^S_{ij} \neq \emptyset \right\}.$$

By means of the similar idea in (Skowron and Rauszer, 1992), for a decision table, the set of all prime implicants of $f(M^S)$ determines the set of all positive region reducts.

To illustrate how to obtain the reducts derived from the discernibility matrix in the sense of Shannon's entropy is given as the following example.

Example 2 The discernibility matrix of Table 1 in the sense of Shannon's entropy is

$$M^S = \begin{pmatrix} \emptyset & \emptyset & \emptyset & \emptyset & \{c_2, c_3, c_4\} & \{c_2, c_3, c_4\} & \{c_2, c_3, c_4\} & \emptyset & \emptyset & \{c_1, c_4\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \{c_3, c_4\} & \{c_3, c_4\} & \{c_3, c_4\} & \emptyset & \emptyset & \{c_1, c_2, c_4\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \{c_2, c_3, c_4\} & \{c_2, c_3, c_4\} & \{c_2, c_3, c_4\} & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \{c_3, c_4\} & \{c_3, c_4\} & \{c_3, c_4\} & \emptyset & \emptyset & \emptyset \\ \{c_2, c_3, c_4\} & \{c_3, c_4\} & \{c_2, c_3, c_4\} & \{c_3, c_4\} & \emptyset & \emptyset & \emptyset & \{c_2, c_3 & \{c_2, c_3\} & \emptyset \\ \{c_2, c_3, c_4\} & \{c_3, c_4\} & \{c_2, c_3, c_4\} & \{c_3, c_4\} & \emptyset & \emptyset & \emptyset & \{c_2, c_3 & \{c_2, c_3\} & \emptyset \\ \{c_2, c_3, c_4\} & \{c_3, c_4\} & \{c_2, c_3, c_4\} & \{c_3, c_4\} & \emptyset & \emptyset & \emptyset & \{c_2, c_3 & \{c_2, c_3\} & \{c_1, c_2, c_3\} \\ \emptyset & \emptyset & \emptyset & \emptyset & \{c_2, c_3\} & \{c_2, c_3\} & \{c_2, c_3\} & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \{c_2, c_3\} & \{c_2, c_3\} & \{c_2, c_3\} & \emptyset & \emptyset & \{c_1, c_2\} \\ \{c_1, c_4\} & \{c_1, c_2, c_4\} & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1, c_2, c_3\} & \emptyset & \{c_1, c_2\} & \emptyset \end{pmatrix}.$$

The discernibility function in the sense of Shannon's entropy is

$$\begin{aligned} f(M^S) &= \{c_2 \vee c_3 \vee c_4\} \wedge \{c_3 \vee c_4\} \wedge \{c_1 \vee c_4\} \wedge \{c_1 \vee c_2 \vee c_3\} \\ &\quad \wedge \{c_1 \vee c_2 \vee c_4\} \wedge \{c_1 \vee c_2\} \wedge \{c_2 \vee c_3\} \\ &= \{c_3 \vee c_4\} \wedge \{c_1 \vee c_4\} \wedge \{c_1 \vee c_2\} \wedge \{c_2 \vee c_3\} \\ &= \{c_1 \wedge c_3\} \vee \{c_2 \wedge c_4\}. \end{aligned}$$

Therefore, we have that $\{c_1, c_3\}$ and $\{c_2, c_4\}$ are all of the Shanon's entropy reducts of Table 1.

Furthermore, we introduce the following theorem to analyze the relationship between m^S_{ij} and m^μ_{ij}.

Theorem 2. *Let $S = (U, C \cup D)$ be a decision table, C the condition attribute set, $D = \{d\}$ the decision attribute set. If $M^S = \{m^S_{ij}\}$ is the discernibility matrix in the sense of Shannon's entropy and $M^\mu = \{m^\mu_{ij}\}$ is the the discernibility matrix in the sense of the rough membership function, then $m^S_{ij} = m^\mu_{ij}$.*

Proof. For proving the theorem, there are four cases should be investigated as follows.

(1) $x_i, x_j \in U_1$

In this case, x_i and x_j are in the consistent part of the decision table S. Therefore, for $\forall Y_k \in U_d$, we have that $\mu_{ik} = 0$ or 1 and $\mu_{jk} = 0$ or 1.

If $\mu_{ik} \neq \mu_{jk}$, then $\mu_{ik} = 0$ and $\mu_{jk} = 1$, or $\mu_{ik} = 1$ and $\mu_{jk} = 0$. Furthermore, $\mu_{ik} = 1$ and $\mu_{ik} = 0 \Rightarrow \frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|} = 1$ and $\frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|} = 0 \Rightarrow [x_i]_C = Y_k$ and $[x_j]_C \cap Y_k = \emptyset \Rightarrow f(x_i, d) \neq f(x_j, d)$. Similarly, $\mu_{ik} = 0$ and $\mu_{jk} = 1 \Rightarrow f(x_i, d) \neq f(x_j, d)$. Therefore, we have $\mu_{ik} \neq \mu_{jk} \Rightarrow f(x_i, d) \neq f(x_j, d)$.

If $f(x_i, d) \neq f(x_j, d)$, then $[x_i]_C = Y_k$ and $[x_j]_C \cap Y_k = \emptyset$, or $[x_i]_C \cap Y_k = \emptyset$ and $[x_j]_C = Y_k$. Furthermore, $[x_i]_C = Y_k$ and $[x_j]_C \cap Y_k = \emptyset \Rightarrow \frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|} = 1$ and $\frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|} = 0 \Rightarrow \mu_{ik} = 1$ and $\mu_{ik} = 0$. Similarly, $f(x_i, d) \neq f(x_j, d) \Rightarrow \mu_{ik} = 0$ and $\mu_{jk} = 1$. Therefore, we have $f(x_i, d) \neq f(x_j, d) \Rightarrow \mu_{ik} \neq \mu_{jk}$.

From the above analysis, we can conclude that m^S_{ij} is equal to m^μ_{ij} in this case.

(2) $x_i \in U_1, x_j \in U_2$

In this case, x_i is in the consistent part of the decision table S and x_j is in its inconsistent. Therefore, we have that $\mu_{ik} = 0$ or 1 and $0 < \mu_{jk} < 1$ for $\forall Y_k \in U_d$. It is obvious that μ_{ik} is unequal with μ_{jk}.

(3) $x_i, x_j \in U_2$ and $\mu_{ik} \neq \mu_{jk}$ for $\forall Y_k \in U/\{d\}$

In this case, x_i and x_j are in the inconsistent part of the decision table S. It is obvious that μ_{ik} is unequal with μ_{jk}.

(4) $x_i, x_j \in U_2$ and $\mu_{ik} = \mu_{jk}$ for $\forall Y_k \in U/\{d\}$

In this case, x_i and x_j are in the inconsistent part of the decision table S. It is obvious that μ_{ik} is equal with μ_{jk}.

In (1), (2) and (3), μ_{ik} is unequal with μ_{jk}, and $m^S_{ij} = m^\mu_{ij} = \{a \in C : f(x_i, a) \neq f(x_j, a)\}$, $\mu_{ik} \neq \mu_{jk}$ for $\forall Y_k \in U/\{d\}$. In (4), μ_{ik} is equal with μ_{jk}, and $m^S_{ij} = m^\mu_{ij} = \emptyset$. Therefore, we have that m^S_{ij} is equal with m^μ_{ij} in all of the cases. □

Theorem 3.5 states that the decision-relative discernibility matrix M^μ is in fact the same as the decision-relative discernibility matrix in the sense of Shannon's entropy.

In the following, we present a theorem to analyze the relationship between m^{Yang}_{ij} and m^S_{ij}.

Theorem 3. *Let $S = (U, C \cup D)$ be a decision table, C the condition attribute set, $D = \{d\}$ the decision attribute set. If m^{Yang}_{ij} and m^S_{ij} the entries in Yang's* *discernibility matrix and the discernibility matrix in the sense of Shannon's entropy respectively, then*

$$m_{ij}^{Yang} \subseteq m_{ij}^{S},$$

especially, if $m_{ij}^{Yang} \neq \emptyset$ *and* $m_{ij}^{S} \neq \emptyset$ *(or* $m_{ij}^{Yang} = \emptyset$ *and* $m_{ij}^{S} = \emptyset$*),*

$$m_{ij}^{Yang} = m_{ij}^{S}.$$

Proof. For proving the theorem, there are four cases should be analyzed as follows.

(1) $x_i, x_j \in U_1$
In this case, $m_{ij}^{Yang} = \{c \in C : f(x_i, c) \neq f(x_j, c) \text{ and } f(x_i, D) \neq f(x_j, D)\} = m_{ij}^{S}$, and m_{ij}^{Yang} and m_{ij}^{S} can be empty set.

(2) $x_i \in U_1$ and $x_j \in U_2$
In this case, $m_{ij}^{Yang} = \{c \in C : f(x_i, c) \neq f(x_j, c)\} = m_{ij}^{S}$, and m_{ij}^{Yang} and m_{ij}^{S} can also be empty set.

(3) $x_i, x_j \in U_2$ and $\mu_{ik} \neq \mu_{jk}$ for $\forall Y_k \in U/D$
In this case, $m_{ij}^{Yang} = \emptyset$, $m_{ij}^{S} = \{c \in C : f(x_i, c) \neq f(x_j, c)\}$, then $m_{ij}^{S} \supseteq m_{ij}^{Yang}$. Especially, m_{ij}^{S} can also be empty set,

(4) $x_i, x_j \in U_2$ and $\mu_{ik} = \mu_{jk}$ for $\forall Y_k \in U/D$
In this case, $m_{ij}^{Yang} = \emptyset$, $m_{ij}^{S} = \emptyset$. Therefore $m_{ij}^{S} = m_{ij}^{Yang}$. □

Theorem 3.6 states that the decision-relative discernibility matrix M^S contain more discernible information than the decision-relative discernibility matrix M^{Yang}. By the relation between M^S and M^{Yang}, we can deduce the following corollary.

Corollary 1. *Let* $S = (U, C \cup D)$ *be a decision table,* C *the condition attribute set,* D *the decision attribute set. If* $Core_D^P(C)$ *and* $Core_D^S(C)$ *are the core in the sense of positive region and the core in the sense of Shannon's entropy respectively, then*

$$Core_D^P(C) \subseteq Core_D^S(C).$$

It is easy to prove the corollary by means of Theorem 3, so we omit it. The same result of Corollary 1 has been in (**?**), which is proved from different perspectives.

Furthermore, we analyze the relationship between positive region reducts and Shannon's entropy reducts.

By Theorem3, without any loss of generality, we suppose that $m_{pq}^{S} = \{c_w\} \supset m_{pq}^{Yang} = \emptyset$, and $m_{ij}^{S} = m_{ij}^{Yang}$ for $\forall\ i \neq p,\ j \neq q$. The set of all positive region reducts is $\mathbf{RED}_D^P(C) = \{red_D^P(C)_1, red_D^P(C)_2, \cdots, red_D^P(C)_{|\mathbf{RED}_D^P(C)|}\}$, $c_1 \notin red_D^P(C)_i$, $1 \leq i \leq |\mathbf{RED}_D^P(C)|$. Thus, we have that

$$f(M^S) = \bigvee_{k\leq|\mathbf{RED}_D^P(C)|} (\wedge_{c_i\in m_{pq}^S} c_i) \wedge (\wedge_{c_j\in red_D^P(C)_k} c_j)$$
$$= \bigvee_{k\leq|\mathbf{RED}_D^P(C)|} (c_w \wedge (\wedge_{c_j\in red_D^P(C)_k} c_j)).$$

By the expression of $f(M^S)$, there exists $red_D^S(C)_u \in \mathbf{RED}_D^S(C)$ such that

$$c_w \wedge (\wedge_{c_j\in red_D^P(C)_v} c_j) \supseteq red_D^S(C)_u.$$

There are three cases should be analyzed as follows:

(1) $c_w \wedge (\wedge_{c_i\in red_D^P(C)_v} c_i) \supset red_D^S(C)_u$, $c_w \notin red_D^P(C)_v$ and $c_w \in red_D^S(C)_u$

In this case, there exists at least one attribute $b \in red_D^P(C)_v$ and $b \notin red_D^S(C)_u$. Therefore $red_D^P(C)_v \neq red_D^S(C)_u$, and $red_D^P(C)_v \cap red_D^S(C)_u \subset red_D^P(C)_v$

(2) $c_w \wedge (\wedge_{c_i\in red_D^P(C)_v} c_i) = red_D^S(C)_u$, $c_w \notin red_D^P(C)_v$ and $c_w \in red_D^S(C)_u$

In this case, $red_D^P(C)_v \subset red_D^S(C)_u$.

(3) $c_w \wedge (\wedge_{c_i\in red_D^P(C)_v} c_i) = red_D^S(C)_u$, $c_w \in red_D^P(C)_v$ and $c_w \in red_D^S(C)_u$

In this case, $red_D^P(C)_v = red_D^S(C)_u$.

Based on the above analysis, we can divide all positive region reducts into three types as follows.

(1) A positive region reduct is not the subset of each of shannon's entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{P1}(C)_v$.

(2) A positive region reduct is the proper subset of one of shannon's entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{P2}(C)$.

(3) A positive region reduct is the same as one of shannon's entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{P3}(C)$.

For these three types of positive region reducts, $\mathbf{RED}_D^P(C) = \mathbf{RED}_D^{P1}(C) \cup \mathbf{RED}_D^{P2}(C) \cup \mathbf{RED}_D^{P3}(C)$. $\mathbf{RED}_D^{P1}(C)$, $\mathbf{RED}_D^{P2}(C)$ and $\mathbf{RED}_D^{P3}(C)$ could be $\emptyset$ respectively. But they can not be empty sets at the same time. This classification of positive region reducts is help to discover the relationship between positive region reducts and Shannon's entorpy reducts.

In the following, we introduce the decision-relative discernibility matrix in the sense of complement entropy. So the complement condition entropy is first reviewed.

For a decision table $S = (U, C \cup D)$, complement entropy of D with respect to C is denoted as (Liang *et al.*, 2009, 2006)

$$E(D|C) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|X_i \cap Y_j|}{|X_i|}\frac{|Y_j^c - X_i^c|}{|X_i|},$$

where Y_j^c and X_i^c are the complements of Y_j and X_i, respectively.

Furthermore, we review the change mechanism of complement entropy with the partition of universe.

Theorem 4. *(Wei et al., 2010) Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots,$*

$X_{v-1}, X_{v+1}, \cdots, X_m, X_u \cup X_v\}$ *and* $U/D = \{Y_1, Y_2, \cdots, Y_n\}$, *then*

$$E(D|B) \geq E(D|C),$$

especially, if $\exists\ w \leq n$ *such that* $\mu(Z_{uw}) = \mu(Z_{vw}) = 1$ *and for* $\forall j \leq n$ *and* $j \neq w$ *such that* $\mu(Z_{uj}) = \mu(Z_{vj}) = 0$, *then*

$$E(D|B) = E(D|C).$$

From Theorem 4, we can see that complement condition entropy is unchanged while two condition classes in which the decision values of objects are same combine and they are in the consistent part of a decision table. Otherwise, the combination of two different equivalence classes make the complement condition entropy to become larger.

For convenience of our further discussion, we give the definition of complement condition entropy as follows.

Definition 11. *Let* $S = (U, C \cup D)$ *be a decision table,* $B \subseteq C$. *We call* B *a complement entropy reduct of* D *with respect to* C *if* B *satisfies the following conditions:*

$$\begin{aligned} &(1) \quad E(D|C) = E(D|B);\ and \\ &(2) \quad for\ \forall a \in B, \quad E(D|B) \neq E(D|B - \{a\}). \end{aligned}$$

Remark: By means of Definition 11 and Theorem 4, the entries in the discernibility matrix in the sense of complement entropy should satisfy the following conditions: (1) If the values of two objects in a decision table's consistent part of is different, then the condition attributes on which the values of these two objects are unequal should be in the corresponding entries in the discernibility matrix in the sense of complement entropy. (2) If one object in a table's consistent part and the other is in its inconsistent part, the condition attributes on which the values of these two objects are different should be in the corresponding entries in the discernibility matrix in the sense of complement entropy. (3) If both of two objects are in a decision table's inconsistent part, the condition attributes on which the values of the two objects are different should be in the corresponding entries in the discernibility. In these three conditions, the first two are the same as the conditions of Yang's discernibility matrix and discernibility matrix based on Shannon's entropy, and the last condition is a special one.

Based on the above analyses, we give the discernibility matrix in the sense of complement entropy as follows.

Definition 12. *Let* $S = (U, C \cup D)$ *be a decision table,* C *the condition attribute set,* D *the decision attribute set,* $D = \{d\}$. *Then, the discernibility matrix in the sense of complement entropy is defined as* $M^C = \{m_{ij}^C\}$, *where*

$$m_{ij}^C = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d)\ and\ x_i, x_j \in U_1 \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2 \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i, x_j \in U_2 \\ \emptyset,\ otherwise \end{cases} \tag{1}$$

Furthermore, the corresponding discernibility function based on complement entropy is defined as

$$f(M^C) = \bigwedge \left\{ \bigvee (m_{ij}^C) \mid \forall x, y \in U, m_{ij}^C \neq \emptyset \right\}. \tag{2}$$

By means of the similar idea in (Skowron and Rauszer, 1992), for a decision table, the set of all prime implicants of $f(M^C)$ determines the set of all complement entropy reducts. To illustrate how to obtain the reducts by means of the discernibility matrix in the sense of complement entropy, we present the following example.

Example 3 For Table 1, the discernibility matrix in the sense of complement entropy is

$$M^C = \begin{pmatrix}
\emptyset & \{c_2\} & \emptyset & \{c_2\} & \{c_2,c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_1,c_4\} \\
\{c_2\} & \emptyset & \{c_2\} & \emptyset & \{c_3,c_4\} & \{c_3,c_4\} & \{c_3,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_1,c_2,c_4\} \\
\emptyset & \{c_2\} & \emptyset & \{c_2\} & \{c_2,c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \emptyset \\
\{c_2\} & \emptyset & \{c_2\} & \emptyset & \{c_3,c_4\} & \{c_3,c_4\} & \{c_3,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \emptyset \\
c_2,\{c_3,c_4\} & \{c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_3,c_4\} & \emptyset & \emptyset & \emptyset & \{c_2,c_3\} & \{c_2,c_3\} & \emptyset \\
c_2,\{c_3,c_4\} & \{c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_3,c_4\} & \emptyset & \emptyset & \emptyset & \{c_2,c_3\} & \{c_2,c_3\} & \emptyset \\
c_2,\{c_3,c_4\} & \{c_3,c_4\} & \{c_2,c_3,c_4\} & \{c_3,c_4\} & \emptyset & \emptyset & \emptyset & \{c_2,c_3\} & \{c_2,c_3\} & \{c_1,c_2,c_3\} \\
\{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_3\} & \{c_2,c_3\} & \{c_2,c_3\} & \emptyset & \emptyset & \emptyset \\
\{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_4\} & \{c_2,c_3\} & \{c_2,c_3\} & \{c_2,c_3\} & \emptyset & \emptyset & \{c_1,c_2\} \\
\{c_1,c_4\} & \{c_1,c_2,c_4\} & \emptyset & \emptyset & \emptyset & \emptyset & \{c_1,c_2,c_3\} & \emptyset & \{c_1,c_2\} & \emptyset
\end{pmatrix},$$

The discernibility function derived from the above discernibility matrix is

$$\begin{aligned}
f(M^C) &= \{c_2\} \wedge \{c_2 \vee c_3 \vee c_4\} \wedge \{c_3 \vee c_4\} \wedge \{c_1 \vee c_4\} \wedge \{c_1 \vee c_2 \vee c_3\} \\
&\quad \wedge \{c_1 \vee c_2 \vee c_4\} \wedge \{c_1 \vee c_2\} \wedge \{c_2 \vee c_3\} \\
&= \{c_2\} \wedge \{c_3 \vee c_4\} \wedge \{c_1 \vee c_4\} \\
&= \{c_1 \wedge c_2 \wedge c_3\} \vee \{c_2 \wedge c_4\}
\end{aligned}$$

Therefore, we have that $\{c_1, c_2, c_3\}$ and $\{c_2, c_4\}$ are all the Shannon's entropy reducts of Table 1.

Theorem 5. *Let $S = (U, C \cup D)$ be a decision table, m_{ij}^S and m_{ij}^C are the entries in the discernibility matrices in the sense of Shannon's entropy and complement entropy respectively, then*

$$m_{ij}^S \subseteq m_{ij}^C$$

especially, if $m_{ij}^S \neq \emptyset$ and $m_{ij}^C \neq \emptyset$ (or $m_{ij}^S = \emptyset$ and $m_{ij}^C = \emptyset$),

$$m_{ij}^S = m_{ij}^C.$$

In similar way with the proof of Theorem 3, the theorem can be easily proved.

Corollary 2. *Let $S = (U, C \cup D)$ be a decision table. If $Core_D^S(C)$ and $Core_D^C(C)$ are the cores in the sense of Shannon's entropy and in the sense of complement entropy respectively, then*

$$Core_D^S(C) \subseteq Core_D^C(C)$$

It is easily to prove the above corollary by Theorem 5, so we omit it.

From Corollary

$$Core_D^P(C) \subseteq Core_D^S(C) \subseteq Core_D^C(C)$$

By the results of Theorem5, based on the similar analysis with the classification of positive region reducts, we can divide all Shannon's entropy reducts into three types as follows.

(1) The shannon's entropy reduct that is not the subset of each of complement entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{S1}(C)$.

(2) The shannon's entropy reduct that is the proper subset of one of complement entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{S2}(C)$.

(3) The shannon's entropy reduct that is the same as one of complement entropy reducts. For convenience, the set of this type of reducts is denoted as $\mathbf{RED}_D^{S3}(C)$.

For these three types of Shannon's entropy reducts, $\mathbf{RED}_D^{S}(C) = \mathbf{RED}_D^{S1}(C) \cup \mathbf{RED}_D^{S2}(C) \cup \mathbf{RED}_D^{S3}(C)$. $\mathbf{RED}_D^{S1}(C)$, $\mathbf{RED}_D^{S2}(C)$ and $\mathbf{RED}_D^{S3}(C)$ could be $\emptyset$ respectively. But they can not be empty sets at the same time. This classification of Shannon's entropy is help to discover the inherent relationship between Shannon's entorpy reducts and complement entorpy reducts.

4. Experiment analysis

There are two representative strategies which can obtain all reducts of a decision table. They are the strategy traversing all powerset of condition attribute set (is short for traversing strategy) and the strategy based on discernibility matrix (is short for discernibility matrix strategy), respectively. It is easy to know the former is more time consuming than the later, and the same reducts can be obtained (Skowron and Rauszer, 1992). For convenience of development of this paper, the algorithms computing positive region reducts, Shannon's entropy reducts and complement entropy reducts based on traversing strategy are denoted as Algorithm P1, Algorithm S1, Algorithm L1, respectively. And the algorithms obtaining these reducts based on discernibility matrix are denoted as Algorithm P2, Algorithm S2, Algorithm L2, respectively. Since these algorithms are common, we omit their description.

In the sequence, in order to compare the reducts obtained by algorithms P1, S1 and C1 with P2, S2 and C2, we employ the UCI dataset Spect, and construct ten datasets by randomly selecting 200 objects from this dataset, which are denoted as Dataset 1, Dataset 2, ..., and Dataset 10. These algorithms are run on a personal computer with Windows XP and Intel(R) Core(TM)2 Quad CPU Q9400, 2.66GHz and 3.37GB memory. The software being used is Matlab 7.0.

From the experimental results, we can see that the reducts obtained by using the algorithms based on traversing strategy are the same as based on discernibility matrix strategy. For the length limitation of this paper, we do not list all reducts of 10 datasets obtained by these algorithms (P1, S1, C1, P2, S2 and C2), and only give the numbers of these reducts in Table 2. From Table 2, we can find that the numbers of positive region reducts are usually larger than the ones of Shannon's entropy reducts, and the numbers of Shannon's entropy reducts are usually larger than the ones of complement entropy reducts. However, the principle do not always holds, for example Dataset 10.

To verify the relationships among these types of reducts, we list all of reducts of Dataset 2 in Tables 3-5 for the representative of ten experimental datasets. From Table 3, we can see that $\mathbf{RED}_D^{P1}(C) = \{p_1, p_4, p_5, p_6, p_7, p_9, p_{11}, p_{21}\}$, $\mathbf{RED}_D^{P2}(C) = \{p_2, p_3, p_8, p_{10}, p_{12}, p_{13}, p_{14}, p_{15}, p_{16}, p_{17}, p_{18}, p_{19}, p_{20}, p_{22}, p_{23}, p_{24}, p_{25}\}$, $\mathbf{RED}_D^{P3}(C) = \{p_{26}, p_{27}, p_{28}, p_{29}, p_{30}, p_{31}, p_{32}, p_{33}\}$. From the results of Table 4, we can see that $\mathbf{RED}_D^{S1}(C) = \emptyset$, $\mathbf{RED}_D^{S2}(C) = \{s_1, s_4, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}\}$, $\mathbf{RED}_D^{S3}(C) = \{s_2, s_3, s_5, s_6, s_{13}, s_{14}, s_{15}, s_{16}, s_{17}, s_{18}, s_{19}, s_{20}, s_{21}, s_{22}, s_{23}, s_{24}\}$. In Table 5, all of complement entropy reducts are given.

5. Conclusions

In this paper, we introduce the discernibility matrixes in the sense of Shannon's entropy and complement entropy. By means of the discernibility functions derived from them, we can obain all shannon's entropy reducts and complement entropy reducts. Furthermore, through analyzing the reducts in the sense of different discernibility power, we revealed the relations among different reducts. Finally, the numerical experiments show the effective of these two proposed matrixes in the sense of entropies, and verify the relationships among these different reducts. These results are helpful to guide the selection of reducts for some practical applications, and the granularity selection in multi-granular rough sets.

Acknowledgements:

This work was supported by National Natural Science Fund of China (No. 71031006, 61202018, 60903110), the Natural Science Foundation of Shanxi Province (No.2010021017-3, 2009021017-1).

References

Bazan, J., Peters, J.F., Skowron,A., Nguyen, H.S., Szczuka, M., 2003. Rough set approach to pattern extraction from classifiers, *Electronic Notes in Theoretical Computer Science*, 82, 1–10.

Düntsch, I., Gediga, G., 1998. Uncertainty measures of rough set prediction, *Artificial Intelligence*, 106, 109–137.

Guan, J.W., Bell, D.A., 1998. Rough computational methods for information systems, *Artificial Intelligence*, 105, 77–103.

Gallager, R.G., 1968. Information Theory and Reliable Communication, *John Wiley Sons New York.*

Hu, Q.H., Xie, Z.X., Yu, D.R., 2007. Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, *Pattern Recognition*, 40, 3509–3521.

Hu, X.H., Cercone, N., 1995. Learning in relational databases: a rough set approach. *International Journal of Computational Intelligence*, 11 (2), 323–338.

Kryszkiewicz, M., 2001. Comparative study of alternative type of knowledge reduction in inconsistent systems, *International Journal of Intelligent Systems*, 16, 105–120.

Li, D.Y., Zhang, B., Leung, Y., 2004. On knowledge reduction in inconsistent decision information systems, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(5), 651–672.

Liang, J.Y., Wang, J.H., Qian, Y.H., 2009. A new measure of uncertainty based on knowledge granulation for rough sets, *Information Sciences*, 179, 458–470.

Liang, J.Y., Shi, Z.Z., Li, D.Y., Wireman, M.J., 2006. The information entropy, rough entropy and knowledge granulation in incomplete information systems, *International Journal of General Systems*, 34(1), 641–654.

Liang, J.Y., Shi, Z.Z., Li, D.Y., 2004. The information entropy, rough entropy and knowledge granulation in rough set theory, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12 (1), 37–46.

Liang, J.Y., Chin, K.S., Dang, C.Y., Richard, C.M. YAM, 2002. A new method for measuring uncertainty and fuzziness in rough set theory, *International Journal of General Systems*, 31 (4), 331–342.

Mi, J.S., Wu, W.Z., Zhang, W.X., 2004. Approaches to knowledge reduction based on variable precision rough set model. *Information Sciences*, 159 (3-4), 255–272.

Nguyen, H.S., Slezak, D., 1999. Approximate reducts and association rules correspondence and complexity results, *In*: N. Zhong, A. Skowron, S. Oshuga, Eds. *Proceeding of RSFDGrc′99*, Yamaguchi, Japan, LNAI 1711, 137–145.

Nguyen, H.S., 2006. Approximate Boolean reasoning: foundations and applications in data mining, *Lecture Notes in Computer Science*, 3100, 334–506.

Pawlak, Z., 1991. Rough Sets: Theoretical Aspects of Reasoning about Data, *London: Kluwer Academic Publisher.*

Pawlak, Z., Skowron, A., 2007a. Rough sets and Boolean reasoning, *Information Sciences*, 177, 41–73.

Pawlak, Z., Skowron, A., 2007b. Rudiments of rough sets, *Information Sciences*, 177, 3–27.

Qian, Y.H., Liang, J.Y., Pedrycz, W., Dang, C.Y., 2010a. Positive approximation: an accelerator for attribute reduction in rough set theory, *Artificial Intelligence*, 174, 597–618.

Qian, Y.H., Liang, J.Y., Yao, Y.Y., Dang, C.Y., 2010b. MGRS: a mulit-granulation rough set, *Information Sciences*, 180, 949–970.

Qian, Y.H., Dang, C.Y., Liang, J.Y., Tang, D.W., 2009. Set-valued ordered information systems, *Information Sciences*, 179, 2809–2832.

Qian, Y.H., Liang, J.Y., Dang, C.Y., 2008. Interval ordered information systems, *Computer and Mathematics with Applications*, 56, 1994–2009.

Shao, M.W., Zhang, W.X., 2005. Dominance relation and rules in an incomplete ordered information system, *International Journal of Intelligent Systems*, 20, 13–27.

Skowron, A., Rauszer, C., 1992. The discernibility matrices and functions in information systems. *Intelligent Decision Support: Handbook of Applications and Advances of Rough Set Theory*, 331-362.

Ślęzak, D., 1996. Approximate reducts in decision tables, *In*: *Proceeding of IPMU′96*, Granada, Spain, 3, 1159–1164.

Ślęzak, D., Foundations of entropy-based Bayesian networks: theoretical results & rough set based extraction from data, 2000. *in:* IPMU'00. *Proceedings of the 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, vol. 1, Madrid, Spain, 248–255.

Ślęzak, D., 2002. Approximate entropy reducts, *Fundamenta Informaticae*, 53 (3-4), 365–390.

Wang, G.Y., Rough reduction in algebra view and information view, *International Journal of Intelligent Systems*, 18, 679–688.

Wei, W., Liang, J.Y., Qian, Y.H., 2012. A comparative study of rough sets to hybrid data, *Information Sciences*, 190(1), 1–16.

Wei, W., Liang, J.Y., Qian, Y.H., Wang, F., Dang, C.Y., 2010. Comparative study of decision performance of decision tables induced by attribute reductions, *International Journal of General Systems*, 39 (8), 813–838.

Wei, W., Liang, J.Y., Qian, Y.H., Wang, F., 2009. An attribute reduction approach and its accelerated version for hybrid data, *The 8th IEEE International Conference on Cognitive Informatics*, 167–173

Wu, W.Z., Zhang, M., Li, H.Z., Mi, J.S., 2005. Knowledge reduction in random information systems via Dempster-Shafer theory of evidence, *Information Sciences*, 174, 143–164.

Yang, M., 2006. An incremental updating algorithm of the computation of a core based on the improved discernibility matrix, *Chinese Journal of Computers*, 29 (3), 407–413.

Yamaguchi, D., 2009. Attribute dependency functions considering data efficiency, *International Journal of Approximate Reasoning*, 51 (1), 89–98.

Yao, Y.Y., Zhao, Y., 2009. Discernibility matrix simplification for constructing attribute reducts, *Information Sciences*, 179, 867–882.

Yao, Y.Y., Zhao, Y., Wang, J., Han, S.Q., 2006. A model of machine learning based on user preference of attributes, *in: Proceedings of International Conference on Rough Sets and Current Trends in Computing*, 587–596.

Ye, D.Y., Chen, Z.J., 2002. A new discernibility matrix and the computation of a core, *Acta Electronica Sinica*, 30 (7), 1086–1088.

Ziarko, W., 1993. Variable precision rough set model, *Journal of Computer and System Science*, 46, 39–59.

Table 1. A decision table about diagnosing rheum

Patients	Headache	Muscle pain	Animal heat	Cough	Rheum
e_1	Yes	Yes	Normal	No	No
e_2	Yes	Yes	High	No	No
e_3	Yes	Yes	Normal	No	Yes
e_4	Yes	Yes	high	No	Yes
e_5	Yes	No	High	Yes	Yes
e_6	Yes	No	High	Yes	Yes
e_7	Yes	No	High	Yes	No
e_8	Yes	Yes	Very high	Yes	Yes
e_9	Yes	Yes	Very high	Yes	No
e_{10}	No	Yes	Normal	Yes	Yes

Table 2. The comparison of the numbers of reducts through different algorithms on 10 datasets

Datasets	Algorithm P1	Algorithm P2	Algorithm S1	Algorithm S2	Algorithm C1	Algorithm C2
1	2	2	2	2	2	2
2	33	33	24	24	24	24
3	16	16	12	12	12	12
4	6	6	4	4	4	4
5	4	4	4	4	4	4
6	12	12	12	12	12	12
7	12	12	12	12	8	8
8	6	6	6	6	6	6
9	4	4	4	4	4	4
10	22	22	14	14	16	16

Table 3. Comparison of the reducts obtained by using Algorithm P2 and Algorithm S2 on Dataset 2

No.	Positive region reducts	Relationship
p_1	1,2,3,4,5,6,8,10,14,16,19,21,22	$p_1 \cup \{13, 20\} \supset s_1$
p_2	1,2,3,4,6,8,10,12,14,16,19,21,22	$p_2 \cup \{13\} = s_2$
p_3	1,2,3,4,8,10,11,14,16,19,21,22	$p_3 \cup \{13, 20\} = s_8$
p_4	1,2,3,4,6,7,8,10,14,16,19,21,22	$p_4 \cup \{13, 20\} \supset s_{14}$
p_5	1,3,4,6,7,8,9,10,14,16,19,21,22	$p_5 \cup \{13, 20\} \supset s_{14}$
p_6	1,3,4,6,7,8,10,13,14,16,19,21,22	$p_6 \cup \{20\} \supset s_{14}$
p_7	1,3,4,6,7,8,10,14,16,17,19,21,22	$p_7 \cup \{13, 20\} \supset s_{14}$
p_8	1,3,4,6,7,8,10,14,16,19,20,21,22	$p_8 \cup \{13\} = s_{14}$
p_9	1,3,4,7,8,9,10,11,14,16,19,21,22	$p_9 \cup \{13, 20\} \supset s_{20}$
p_{10}	1,3,4,7,8,10,11,13,14,16,19,21,22	$p_{10} \cup \{20\} = s_{20}$
p_{11}	1,3,4,7,8,10,11,14,16,17,19,21,22	$p_{11} \cup \{13, 20\} \supset s_{20}$
p_{12}	1,3,4,7,8,10,11,14,16,19,20,21,22	$p_{12} \cup \{13\} = s_{20}$
p_{13}	1,2,3,4,5,6,8,14,16,19,20,21,22	$p_{13} \cup \{13\} = s_1$
p_{14}	1,2,3,4,6,8,12,14,16,17,19,20,21,22	$p_{14} \cup \{13\} = s_3$
p_{15}	1,2,3,4,5,8,11,14,16,19,20,21,22	$p_{15} \cup \{13\} = s_7$
p_{16}	1,2,3,4,8,11,14,16,17,19,20,21,22	$p_{16} \cup \{13\} = s_9$
p_{17}	1,3,4,5,6,7,8,14,16,19,20,21,22	$p_{17} \cup \{13\} = s_{13}$
p_{18}	1,3,4,6,7,8,14,16,17,19,20,21,22	$p_{18} \cup \{13\} = s_{15}$
p_{19}	1,3,4,5,7,8,11,14,16,19,20,21,22	$p_{19} \cup \{13\} = s_{19}$
p_{20}	1,3,4,7,8,11,14,16,17,19,20,21,22	$p_{20} \cup \{13\} = s_{21}$
p_{21}	1,2,3,5,6,8,9,10,13,14,16,19,21,22	$p_{21} \cup \{20\} \supset s_4$
p_{22}	1,2,3,6,8,9,10,12,13,14,16,19,21,22	$p_{22} \cup \{20\} = s_5$
p_{23}	1,2,3,8,9,10,11,13,14,16,19,21,22	$p_{23} \cup \{20\} = s_{11}$
p_{24}	1,3,6,7,8,9,10,13,14,16,19,21,22	$p_{24} \cup \{20\} = s_{17}$
p_{25}	1,3,7,8,9,10,11,13,14,16,19,21,22	$p_{25} \cup \{20\} = s_{23}$
p_{26}	1,2,3,5,6,8,9,13,14,16,19,20,21,22	$p_{26} = s_4$
p_{27}	1,2,3,6,8,9,12,13,14,16,17,19,20,21,22	$p_{27} = s_6$
p_{28}	1,2,3,5,8,9,11,13,14,16,19,20,21,22	$p_{28} = s_{10}$
p_{29}	1,2,3,8,9,11,13,14,16,17,19,20,21,22	$p_{29} = s_{12}$
p_{30}	1,3,5,6,7,8,9,13,14,16,19,20,21,22	$p_{30} = s_{16}$
p_{31}	1,3,6,7,8,9,13,14,16,17,19,20,21,22	$p_{31} = s_{18}$
p_{32}	1,3,5,7,8,9,11,13,14,16,19,20,21,22	$p_{32} = s_{22}$
p_{33}	1,3,7,8,9,11,13,14,16,17,19,20,21,22	$p_{33} = s_{24}$

Table 4. Comparison of the reducts obtained by using Algorithm S2 and Algorithm C2 on Dataset 2

No.	Shannon's entropy reducts	Relationship
s_1	1,2,3,4,5,6,8,13,14,16,19,20,21,22	$s_1 \cup \{12\} = c_{13}$
s_2	1,2,3,4,6,8,10,12,13,14,16,19,20,21,22	$s_2 = c_{14}$
s_3	1,2,3,4,6,8,12,13,14,16,17,19,20,21,22	$s_3 = c_{15}$
s_4	1,2,3,5,6,8,9,13,14,16,19,20,21,22	$s_4 \cup \{12\} = c_{16}$
s_5	1,2,3,6,8,9,10,12,13,14,16,19,20,21,22	$s_5 = c_{17}$
s_6	1,2,3,6,8,9,12,13,14,16,17,19,20,21,22	$s_6 = c_{18}$
s_7	1,2,3,4,5,8,11,13,14,16,19,20,21,22	$s_7 \cup \{12\} = c_{18}$
s_8	1,2,3,4,8,10,11,13,14,16,19,20,21,22	$s_8 \cup \{12\} = c_{20}$
s_9	1,2,3,4,8,11,13,14,16,17,19,20,21,22	$s_9 \cup \{12\} = c_{21}$
s_{10}	1,2,3,5,8,9,11,13,14,16,19,20,21,22	$s_{10} \cup \{12\} = c_{22}$
s_{11}	1,2,3,8,9,10,11,13,14,16,19,20,21,22	$s_{11} \cup \{12\} = c_{23}$
s_{12}	1,2,3,8,9,11,13,14,16,17,19,20,21,22	$s_{12} \cup \{12\} = c_{24}$
s_{13}	1,3,4,5,6,7,8,13,14,16,19,20,21,22	$s_{13} = c_1$
s_{14}	1,3,4,6,7,8,10,13,14,16,19,20,21,22	$s_{14} = c_2$
s_{15}	1,3,4,6,7,8,13,14,16,17,19,20,21,22	$s_{15} = c_3$
s_{16}	1,3,5,6,7,8,9,13,14,16,19,20,21,22	$s_{16} = c_4$
s_{17}	1,3,6,7,8,9,10,13,14,16,19,20,21,22	$s_{17} = c_5$
s_{18}	1,3,6,7,8,9,13,14,16,17,19,20,21,22	$s_{18} = c_6$
s_{19}	1,3,4,5,7,8,11,13,14,16,19,20,21,22	$s_{19} = c_7$
s_{20}	1,3,4,7,8,10,11,13,14,16,19,20,21,22	$s_{20} = c_8$
s_{21}	1,3,4,7,8,11,13,14,16,17,19,20,21,22	$s_{21} = c_9$
s_{22}	1,3,5,7,8,9,11,13,14,16,19,20,21,22	$s_{22} = c_{10}$
s_{23}	1,3,7,8,9,10,11,13,14,16,19,20,21,22	$s_{23} = c_{11}$
s_{24}	1,3,7,8,9,11,13,14,16,17,19,20,21,22	$s_{24} = c_{12}$

Table 5. the reducts obtained by using Algorithm C2 on Dataset 2

No.	complement entropy reducts
c_1	1,3,4,5,6,7,8,13,14,16,19,20,21,22
c_2	1,3,4,6,7,8,10,13,14,16,19,20,21,22
c_3	1,3,4,6,7,8,13,14,16,17,19,20,21,22
c_4	1,3,5,6,7,8,9,13,14,16,19,20,21,22
c_5	1,3,6,7,8,9,10,13,14,16,19,20,21,22
c_6	1,3,6,7,8,9,13,14,16,17,19,20,21,22
c_7	1,3,4,5,7,8,11,13,14,16,19,20,21,22
c_8	1,3,4,7,8,10,11,13,14,16,19,20,21,22
c_9	1,3,4,7,8,11,13,14,16,17,19,20,21,22
c_{10}	1,3,5,7,8,9,11,13,14,16,19,20,21,22
c_{11}	1,3,7,8,9,10,11,13,14,16,19,20,21,22
c_{12}	1,3,7,8,9,11,13,14,16,17,19,20,21,22
c_{13}	1,2,3,4,5,6,8,12,13,14,16,19,20,21,22
c_{14}	1,2,3,4,6,8,10,12,13,14,16,19,20,21,22
c_{15}	1,2,3,4,6,8,12,13,14,16,17,19,20,21,22
c_{16}	1,2,3,5,6,8,9,12,13,14,16,19,20,21,22
c_{17}	1,2,3,6,8,9,10,12,13,14,16,19,20,21,22
c_{18}	1,2,3,6,8,9,12,13,14,16,17,19,20,21,22
c_{19}	1,2,3,4,5,8,11,12,13,14,16,19,20,21,22
c_{20}	1,2,3,4,8,10,11,12,13,14,16,19,20,21,22
c_{21}	1,2,3,4,8,11,12,13,14,16,17,19,20,21,22
c_{22}	1,2,3,5,8,9,11,12,13,14,16,19,20,21,22
c_{23}	1,2,3,8,9,10,11,12,13,14,16,19,20,21,22
c_{24}	1,2,3,8,9,11,12,13,14,16,17,19,20,21,22

Wei Wei is a PhD candidate from School of Computer and Information Technology at Shanxi University, China. He received the MSc degree in Computers with applications at Shanxi University (2007). His research interests include data mining and knowledge discovery, rough set theory and granular computing.

Jiye Liang is a professor of School of Computer and Information Technology and Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education at Shanxi University. He received the PhD degree in Information Science from Xi'an Jiaotong University. He also has a MS degree in computational mathematics from Xi'an Jiaotong University. His research interests include artificial intelligence, granular computing, data mining and knowledge discovery.

Junhong Wang is a lecturer of School of Computer and Information Technology and Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education at Shanxi University. She received the PhD degree in Computer with applications at Shanxi University. Her research interests include Formal concept analysis, rough set theory, data mining and knowledge discovery.

Yuhua Qian is a PhD candidate from School of Computer and Information Technology at Shanxi University, China. He received the MS degree in Computers with applications at Shanxi University. His research interests include data mining and knowledge discovery, rough set theory, granular computing and artificial intelligence.