Change Mechanism of a Decision Table's Decision Performance Caused by Attribute Reductions

Wei Wei Jiye Liang Yuhua Qian
Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, Shanxi, 030006, China
weiwei@sxu.edu.cn ljy@sxu.edu.cn jinchengqyh@126.com

Abstract

Attribute reduction is one of the important topics in the research on rough set theory, it plays an important role in machine learning and data mining. However, is the decision performance of a decision table changed after an attribute reduction? In this paper, we analyze the change of the decision performance through using the positive-region reduction and the Shannon entropy reduction. The change principles of three decision measures caused by attribute reductions are proposed. Experimental analysis is performed for illuminating the change mechanism of a decision table's decision performance.

***Keywords**: Decision evaluation; rough set theory; attribute reduction.*

1. Introduction

Rough set theory proposed by Z. Pawlak has become a popular mathematical framework for pattern recognition, image processing, feature selection, conflict analysis, decision support, data mining and knowledge discovery process from large data sets. As applications of rough set theory in decision problems, a number of reduct techniques have been proposed in the last twenty years [1, 3-8, 10, 11, 17-23]. In these reduct approaches, attribute reduction based on discernibility matrix and based on attribute dependency are classical ones.

A set of decision rules can be generated from a decision table by adopting any kind of reduction method mentioned above [16]. How to evaluate the decision performance of a decision rule is a very important issue in rough set theory. In [2], based on information entropy, Düntsch suggested some uncertainty measures of a decision rule and proposed three criteria for model selection. Moreover, several other measures such as certainty degree and support degree are often used to evaluate a decision rule [22]. However, all of these measures are only defined for a single decision rule and are not suitable for measuring the decision performance of a rule set. There are two more kinds of measures in the literature [12], which are approximation accuracy for decision classification and consistency degree for a decision table. Although these two measures, in some sense, could be regarded as measures for evaluating the decision performance of all decision rules generated from a decision table, they have some limitations. To overcome the shortcomings of the existing measures, in [15], three new measures are proposed for this objective, which are certainty measure (α), consistency measure (β), and support measure (γ).

In this paper, we mainly analyze the change of decision performance after performing any of attribute reduction based on the positive region and that based on the Shannon entropy. The rest of this paper is organized as follows. Some preliminary concepts are briefly recalled in Section 2. In Section 3, three measures (α, β and γ) are recalled for evaluating the decision performance of a set of rules. In Section 4, we analyze the difference of the decision performance between the reduced and the original decision table, and experimental analysis of each of the three measures are performed. Section 5 concludes this paper.

2. Preliminaries

In this section, we review some basic concepts such as indiscernibility relation, partition, partial relation of knowledge and decision tables.

Let $S = (U, A)$ be an information system, where U is a non-empty, finite set of objects and is called a universe and A is a non-empty, finite set of attributes. For each $a \in A$, a mapping $a : U \rightarrow V_a$ is determined by an information system, where V_a is the domain of a.

Each non-empty subset $B \subseteq A$ determines an indiscernibility relation in the following way,

$$R_B = \{(x, y) \in U \times U \mid a(x) = a(y), \forall a \in B\}.$$

The relation R_B partitions U into some equivalence classes given by $U/R_B = \{[x]_B \mid x \in U\}$, just U/B, where $[x]_B$ denotes the equivalence class determined by x with respect to B.

We define a partial relation $\preceq$ on the family $\{U/B \mid B \subseteq A\}$ as follows: $U/P \preceq U/Q$ (or $U/Q \succeq U/P$) if and only if, for every $P_i \in U/P$, there exists $Q_j \in U/Q$ such that $P_i \subseteq Q_j$, where $U/P = \{P_1, P_2, \cdots, P_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ are partitions induced by $P,\ Q \subseteq A$, respectively.

Let $S = (U, C \cup D)$ with $C \cap D = \emptyset$ be an information system, where C is called a condition attribute set, and D is called a decision attribute set, then S is called as a decision table. If $U/C \preceq U/D$, then S is said to be consistent, otherwise it is said to be inconsistent. And, the relative positive-region D with respect to C is defined as

$$POS_C(D) = \bigcup_{i=1}^{n} \underline{C}Y_i$$

Definition 1.[12] Let $S = (U, C \cup D)$ be a decision table, $B \subseteq C$. If B satisfies the following condition:
(1) $POS_C(D) = POS_B(D)$ and
(2) $\forall a \in C,\ POS_C(D) = POS_{C-\{a\}}(D)$,
then B is a positive-region reduct of D with respect to C.

In[18], by means of shannon entropy, the uncertainty of a decision table $S = (U, C \cup D)$ is depicted 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 $|X_i| \in U/C,\ |Y_j| \in U/D$.

Definition 2.[18] Let $S = (U, C \cup D)$ be a decision table, $B \subseteq C$. If B satisfies the following condition:
(1) $H(D|C) = H(D|B)$ and
(2) for $\forall a \in C,\ H(D|C) = H(D|C - \{a\})$,
then B is a Shannon entropy reduct of D with respect to C.

3. Decision rule and decision performance measurement in decision tables

In this section, we briefly recall certainty degree and support degree of a decision rule and the decision performance measurement of a decision table in rough set theory.

Let $S = (U, C \cup D)$ be a decision table, $X_i \in U/C$, $Y_j \in U/D$ and $X_i \cap Y_j \neq \emptyset$. By $des(X_i)$ and $des(Y_j)$, we denote the descriptions of the equivalence classes X_i and Y_j in the decision table S. A decision rule is formally defined as $Z_{ij} : des(X_i) \rightarrow des(Y_j)$.

Certainty degree and support degree of a decision rule Z_{ij} are defined as follows: $\mu(Z_{ij}) = |X_i \cap Y_j|/|X_i|$ and $s(Z_{ij}) = |X_i \cap Y_j|/|U|$, where $|\cdot|$ is the cardinality of a set. It is clear that the value of each of $\mu(Z_{ij})$ and $s(Z_{ij})$ of a decision rule Z_{ij} falls into the interval $[0, 1]$. However, $\mu(Z_{ij})$ and $s(Z_{ij})$ are only defined for a single decision rule and are not suitable for measuring the decision performance of a decision-rule set.

To measure the decision performance of a decision-rule set, in [15], three measures, certainty measure, consistency measure and support measure was proposed, they was defined as follow:

Definition 3.[15] Let $S = (U, C \cup D)$ be a decision table, and $RULE = \{Z_{ij} | Z_{ij} : des(X_i) \rightarrow des(Y_j), X_i \in U/C, Y_j \in U/D\}$, certainty measure α of S is defined as

$$\alpha(S) = \sum_{i=1}^{m}\sum_{j=1}^{n} s(Z_{ij})\mu(Z_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|X_i \cap Y_j|^2}{|U||X_i|},$$

consistency measure β of S is defined as

$$\beta(S) = \sum_{i=1}^{m} \frac{|X_i|}{|U|}[1 - \frac{4}{|X_i|}\sum_{j=1}^{N_i} |X_i \cap Y_j|\mu(Z_{ij})(1-\mu(Z_{ij}))],$$

support measure γ of S is defined as

$$\gamma(S) = \sum_{i=1}^{m}\sum_{j=1}^{n} s^2(Z_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|X_i \cap Y_j|^2}{|U|^2},$$

where $\mu(Z_{ij})$ and $s(Z_{ij})$ are the certainty degree and support degree of the rule Z_{ij}, respectively.

4. Change of decision performance induced by two kinds of reduction approaches

In this section, we analyze the change of decision performance of a decision table when some condition classes combine. Furthermore, the difference between the decision performance of an original decision table and that of each of the reduced decision tables is investigated, which are reduced by the positive-region reduction and the Shannon entropy reduction, respectively.

4.1 Change mechanism of decision performance

Lemma 1. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, if $B \preceq C$, then

$$\alpha(S) \geq \alpha(S'),$$

and, if and only if $\mu(Z_{uj}) = \mu(Z_{vj}), \forall j \leq n$, then

$$\alpha(S) = \alpha(S').$$

where $\mu(Z_{uj})$ and $\mu(Z_{vj})$ represent the certainty degrees of the rules Z_{uj} and Z_{vj} in the table S respectively.

Proof. For simplicity, without any loss of the universe, let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots, X_{v-1}, X_{v+1}, \cdots, X_m, X_u \cup X_v\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

$$\begin{aligned}\alpha_\Delta &= \alpha(S') - \alpha(S) \\ &= \sum_{j=1}^{n} \frac{(|X_u \cap Y_j| + |X_v \cap Y_j|)^2}{|U|(|X_u| + |X_v|)} \\ &\quad - \sum_{j=1}^{n} \frac{|X_u \cap Y_j|^2}{|U||X_u|} - \sum_{j=1}^{n} \frac{|X_v \cap Y_j|^2}{|U||X_v|} \\ &= -\sum_{j=1}^{n} \frac{|X_u||X_v|(\mu(Z_{uj}) - \mu(Z_{vj}))^2}{|U|(|X_u| + |X_v|)} \le 0.\end{aligned}$$

Obviously, when $\mu(Z_{uj}) = \mu(Z_{vj})$, $\alpha_\Delta = 0$, i.e. $\alpha(S) = \alpha(S')$. This completes the proof. □

Lemma 2. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, if $B \preceq C$ and $\mu(Z_{uj}) = \mu(Z_{vj})$, $\forall j \le n$, then

$$\beta(S) = \beta(S').$$

where $\mu(Z_{uj})$ and $\mu(Z_{vj})$ represent the certainty degrees of the rules Z_{uj} and Z_{vj} in the table S respectively.

Proof. For simplicity, without any loss of the universe, let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots, X_{v-1}, X_{v+1}, \cdots, X'_m, X_u \cup X_v\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

$$\begin{aligned}\beta_\Delta &= \beta(S') - \beta(S) \\ &= \frac{1}{|U|}(|X_u| + |X_v| - 4\sum_{j=1}^{n} \frac{(|X_u \cap Y_j| + |X_v \cap Y_j|)^2}{|X_u| + |X_v|} \\ &\quad - \frac{(|X_u \cap Y_j| + |X_u \cap Y_j|)^3}{(|X_u| + |X_v|)^2}) \\ &\quad - \frac{1}{|U|}(|X_u| - 4\sum_{j=1}^{n} \frac{|X_u \cap Y_j|^2}{|X_u|} - \frac{|X_u \cap Y_j|^3}{|X_u|^2}) \\ &\quad - \frac{1}{|U|}(|X_v| - 4\sum_{j=1}^{n} \frac{|X_v \cap Y_j|^2}{|X_v|} - \frac{|X_v \cap Y_j|^3}{|X_v|^2}).\end{aligned}$$

Let $x = |X_u|$, $y = |X_v|$, $\delta_j = \frac{|X_u \cap Y_j|}{|X_u|}$ and $\sigma_j = \frac{|X_v \cap Y_j|}{|X_v|}$. It follows that

$$\begin{aligned}\beta_\Delta &= 4\sum_{j=1}^{n} \frac{(\delta_j x)^2}{x} - \frac{(\delta_j x)^3}{x^2} + 4\sum_{j=1}^{n} \frac{(\sigma_j y)^2}{y} - \frac{(\sigma_j y)^3}{y^2} \\ &\quad -4\sum_{j=1}^{n} \frac{(\delta_j x + \sigma_j y)^2}{x+y} - \frac{(\delta_j x + \sigma_j y)^3}{(x+y)^2}) \\ &= 4\sum_{j=1}^{n} \frac{xy(\delta_j - \sigma_j)^2((1 - 2\delta_j - \sigma_j)x)}{(x+y)^2} \\ &\quad + \frac{(1 - 2\sigma_j - \delta_j)y)}{(x+y)^2}.\end{aligned}$$

Obviously, $\delta_j = \sigma_j$, $\forall j \le n$, i.e., $\mu(Z_{uj}) = \mu(Z_{vj})$, $\beta_\Delta = 0$, $\beta(S) = \beta(S')$. This completes the proof. □

Lemma 3. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, if $B \preceq C$, then

$$\gamma(S') \ge \gamma(S).$$

Proof. For simplicity, without any loss of the universe, let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots, X_{v-1}, X_{v+1}, \cdots, X'_m, X_u \cup X_v\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

$$\begin{aligned}\gamma_\Delta &= \gamma(S') - \gamma(S) \\ &= \sum_{j=1}^{n} \frac{(|X_u \cap Y_j| + |X_v \cap Y_j|)^2}{|U|^2} \\ &\quad - \sum_{j=1}^{n} \frac{|X_u \cap Y_j|^2}{|U|^2} - \sum_{j=1}^{n} \frac{|X_v \cap Y_j|^2}{|U|^2} \\ &= \sum_{j=1}^{n} \frac{2|X_u \cap Y_j||X_v \cap Y_j|}{|U|^2} \ge 0.\end{aligned}$$

This completes the proof. □

4.2 Change of decision performance induced by positive-region reduction

Theorem 1. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables and B a positive-region reduct of C. In the consistent part of the decision table S, if some of the condition classes combine to a new condition class in S', then the certainty measure of arbitrary one of the original rules equals that of another rule among them.

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/D = \{Y_1, Y_2, \cdots, Y_n\}$ and $U/B = \{X'_1, X'_2, \cdots, X'_l\}, (l \le m)$, where $B \subseteq C$ and B is a positive region reduct of C.

For simplicity, without any loss of the universe, suppose that $X_1, X_2, \cdots, X_p$ is in the consistent part of the decision table S, $X'_1, X'_2, \cdots, X'_q$ is in the consistent part of the decision table S' and the $X_u, X_v (u, v \le p)$ become a new condition class $X'_w (w \le q)$, i.e. $X'_w = X_u \cup X_v$ and other condition classes are unchanged. The condition classes X'_w is in the inconsistent part of the table S' if $\exists j \le n$ such that $\mu(Z_{uj}) \ne \mu(Z_{vj})$. Clearly, it is contradiction with the assumption that B is a reduct of C. Therefore, $\mu(Z_{uj}) = \mu(Z_{vj})$, $\forall j \le n$. This completes the proof. □

Theorem 2. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, where S is consistent, $B \subseteq C$. If B is a positive-region reduct of C, then

$$\alpha(S) = \alpha(S').$$

By Theorem 1 and lemma 1, Theorem 2 is easy to be proved.

Theorem 3. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$. If B is a positive-region reduct of C, then

$$\alpha(S) \geq \alpha(S').$$

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1', X_2', \cdots, X_l'\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$. Furthermore, suppose that the consistent part of the table S consists of $X_1, X_2, \cdots, X_p$, the inconsistent part of the table consists of $X_{p+1}, X_{p+2}, \cdots, X_m$, the consistent part of the table S' consists of $X_1', X_2', \cdots, X_q'$, and the inconsistent part of the table S' consists of $X_{q+1}', X_{q+2}', \cdots, X_l'$.

Through using the positive-region reduction, the change of condition class has two cases, one is combination of condition classes in consistent part of a decision table, the other is combination of condition classes in inconsistent part of a decision table. These two cases are listed as follows:

1) the condition classes combined in the consistent part

Let the $X_u, X_v (u, v \leq p)$ become a new condition class X_w' after performing the positive-region reduction, i.e. $X_w' = X_u \cup X_v$, and the other condition classes are unchanged. From Theorem 1, it follows that $\mu(Z_{uj}) = \mu(Z_{vj}), j \leq n$. Moreover, according to Lemma 1, $\alpha(S) = \alpha(S')$.

2) the condition classes combined in the inconsistent part

Let the $X_u, X_v (p < u \leq m, p < v \leq m)$ are combined to X_w' after the positive-region reduction, i.e. $X_w' = X_u \cup X_v$, and other condition classes are unchanged. From Lemma 1, it follows that $\alpha(S) = \alpha(S')$, if and only if $\mu(Z_{uj}) = \mu(Z_{vj}),\ \forall j \leq n$. Otherwise, $\alpha(S) > \alpha(S')$. This completes the proof. □

Theorem 4. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, where S is consistent, $B \subseteq C$. If B is a positive-region reduct of C, then

$$\beta(S) = \beta(S').$$

By Theorem 1 and lemma 2, Theorem 4 is easy to be proved.

Theorem 5. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, where S is consistent, $B \subseteq C$. If B is a positive-region reduct of C, then

$$\gamma(S) \leq \gamma(S'),$$

By Theorem 1 and lemma 3, Theorem 5 is easy to be proved.

Theorem 6. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$. If B is a positive-region reduct of C, then

$$\gamma(S) \leq \gamma(S'),$$

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1', X_2', \cdots, X_l'\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$. $X_1, X_2, \cdots, X_p$ is in the consistent part of the decision table S, $X_{p+1}, X_{p+2}, \cdots, X_m$ is in the inconsistent part of the decision table S, $X_1', X_2', \cdots, X_q'$ is in the consistent part of the decision table S', and $X_{q+1}', X_{q+2}', \cdots, X_l'$ is in the inconsistent part of the decision table S'.

After using the positive-region reduction, the change of condition classes can be two cases. One is a combination of condition classes in consistent part of a decision table, the other is combination of condition classes in inconsistent part of a decision table. These two cases are listed as follows:

1) the condition class combined in the consistent part

Suppose that the $X_u, X_v (u, v \leq p)$ become a new condition class X_w' after the positive-region reduction, i.e. $X_w' = X_u \cup X_v$, and other condition classes are unchanged. According to Lemma 3, $\gamma(S) \leq \gamma(S')$.

2) the condition class combined in the inconsistent part

Let the $X_u, X_v (p < u \leq m, p < v \leq m)$ combine to X_w' through using the positive-region reduction, i.e. $X_w' = X_u \cup X_v$, and other condition classes is unchanged. From Lemma 3, it follows that $\gamma(S) \leq \gamma(S')$. This completes the proof. □

For general decision tables, in the following, through experimental analysis, we illustrate the change of the decision performance after using the positive-region reduction. We have download the mushroom data set from UCI database. In order to verify their performance, we remove the incomplete part of from mushroom data set, and the rest 5644 objects are divided into eight parts with $1/8$ of all objects step. The first part is regard as the 1st data set, the combination of the first and the second part is regard as the 2nd data set, $\cdots$, the combination of all eight parts is regard as the 8th data set. The comparison of value of the three performance measures of original table and the one after using shannon entropy reduction is shown in Table 1.

4.3 Change of decision performance induced by Shannon entropy reduct

Theorem 7.[18] Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$, $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X_1, X_2, \cdots, X_{u-1}, X_{u+1}, \cdots, X_{v-1}, X_{v+1}, \cdots, X_m', X_u \cup X_v\}$, and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$, then

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

Table 1. the comparison of α, β, γ and those induced by a positive-region reduction

Measure	data set							
	1st	2nd	3rd	4th	5th	6th	7th	8th
α	0.782979	0.740426	0.778251	0.829433	0.851631	0.846809	0.773455	0.728561
α'	0.782979	0.740426	0.776359	0.821868	0.850496	0.845863	0.773455	0.718842
β	0.565957	0.480851	0.556501	0.666430	0.705910	0.711899	0.619588	0.559178
β'	0.565957	0.480851	0.558392	0.658865	0.707045	0.712845	0.619588	0.568897
γ	0.001418	0.000709	0.000473	0.000355	0.000284	0.000236	0.000203	0.000177
γ'	0.001459	0.001013	0.001286	0.015477	0.000918	0.000725	0.000441	0.019752

and if and only if $\frac{|X_u \cap Y_j|}{|X_u|} = \frac{|X_v \cap Y_j|}{|X_v|}$, $\forall j < n$, i.e. $\mu(Z_{uj}) = \mu(Z_{vj})$, then

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

Theorem 8. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$. If B is a Shannon entropy reduct of C, then

$$\alpha(S) = \alpha(S').$$

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X'_1, X'_2, \cdots, X'_l\}$, and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

Through using shannon entropy reduction, without any lose of the universe, we suppose that X_u and $X_v(u, v \leq m)$ combine to $X'_w(w \leq l)$. From Theorem 7, it follows that $\mu(Z_{uj}) = \mu(Z_{vj})$. Moreover, by Lemma 1, $\alpha(S) = \alpha(S')$. This completes the proof. □

Theorem 9. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$. If B is a Shannon entropy reduct of C, then

$$\beta(S) = \beta(S').$$

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X'_1, X'_2, \cdots, X'_l\}$, and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

For simplicity, without any lose of the universe, we suppose that after using shannon entropy reduction, X_u and $X_v(u, v \leq m)$ combine to $X'_w(w \leq l)$. In terms of Theorem 7, it follows that $\mu(Z_{uj}) = \mu(Z_{vj})$. Furthermore, by Lemma 2, $\beta(S) = \beta(S')$. This completes the proof. □

Theorem 10. Let $S = (U, C \cup D)$ and $S' = (U, B \cup D)$ be two decision tables, $B \subseteq C$. If B is a Shannon entropy reduct of C, then

$$\gamma(S) \leq \gamma(S').$$

Proof. Let $U/C = \{X_1, X_2, \cdots, X_m\}$, $U/B = \{X'_1, X'_2, \cdots, X'_l\}$ and $U/D = \{Y_1, Y_2, \cdots, Y_n\}$.

For simplicity, without any lose of the universe, we suppose that the $X_u, X_v(p < u \leq m, p < v \leq m)$ combine to X'_w after the Shannon entropy reduct, i.e. $X'_w = X_u \cup X_v$, and other condition classes is unchanged. By Lemma 3, $\gamma(S) \leq \gamma(S')$. This completes the proof. □

In the following, we illustrate the change of the decision performance after using the shannon entropy reduction. We use the same data sets in the table 1. The comparison of value of the three performance measures of original decision table and the one after using shannon entropy reduction is shown in Table 2.

From Table 1 and Table 2, it is easy to draw the following conclusions:

1) through using a positive-region reduction, the certainty measure α is not bigger than the original certainty measure, the support measure γ is not smaller than the original support measure, and the change of the consistency measure β is uncertain; and

2) after performing a Shannon entropy reduction, each of the certainty measure α and the consistency measure β is the same as each of those induced by a original decision table, and the support measure γ is not smaller than the original support measure.

5. Conclusion

In this paper, we have analyzed the change mechanism of the decision performance after performing the positive-region reduction and Shannon entropy reduction, and have obtained some of their important properties. These three measures may be changed through using a positive-region reduction. However, the certainty measure and the consistent measure are unchanged after using a Shannon entropy reduction, and the support measure is usually increased. These results may be helpful for determining which of the positive-region reduction and the attribute reduction based on Shannon entropy is preferred for a practical decision problem in the context of complete decision tables.

Table 2. the comparison of α, β, γ and those induced by a Shannon entropy reduction

Measure	data set							
	1st	2nd	3rd	4th	5th	6th	7th	8th
α	0.782979	0.740426	0.778251	0.829433	0.851631	0.846809	0.773455	0.728561
α'	0.782979	0.740426	0.778251	0.829433	0.851631	0.846809	0.773455	0.728561
β	0.565957	0.480851	0.556501	0.666430	0.705910	0.711899	0.619588	0.559178
β'	0.565957	0.480851	0.556501	0.666430	0.705910	0.711899	0.619588	0.559178
γ	0.001418	0.000709	0.000473	0.000355	0.000284	0.000236	0.000203	0.000177
γ'	0.001459	0.001013	0.001278	0.017674	0.001044	0.000813	0.000441	0.081572

Acknowledgements:

This work was supported by the high technology research and development program of China (No. 2007AA01Z165), the national natural science foundation of China (No. 60773133), the foundation of doctoral program research of ministry of education of China (No. 20050108004), key project of science and technology research of the ministry of education of China (No. 206017) and the natural science foundation of Shanxi province (No. 2008011038).

References

[1] M. Beynon. Reducts within the variable precision rough sets model: a further investigation. *European Journal of Operational Research*, 134(3):592–605, 2001.

[2] I. Düntsch and G. Gediaga. Uncertainty measures of rough set prediction. *Artificial Intelligence*, 106(1):109–137, 1998.

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

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

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

[6] J. Y. Liang and D. Y. Li. *Uncertainty and Knowledge Acquisition in Information Systems*. Science Press, Beijing, 2005.

[7] J. Y. Liang and Z. B. Xu. The algorithm on knowledge reduction in incomplete information systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 24(1):95–103, 2002.

[8] T. Y. Lin. Introduction to special issues on data mining and granular computing. *International Journal of Approximate Reasoning*, 40:1–2, 2005.

[9] Q. Liu, H. S. Zhu, and L. Liu. The granulations based on meaning of rough logical formulas and its lock resolution. *Granular Computing, GrC2007*, pages 44–49, 2007.

[10] J. S. Mi, W. Z. Wu, and W. X. Zhang. Comparative studies of knowledge reductions in inconsistent systems. *Fuzzy Systems and Mathematics*, 17(3):54–60, 2003.

[11] H. S. Nguyen. Rough sets and association rule generation. *Fundamenta Informaticae*, 40(4):383–405, 1999.

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

[13] W. Pedrycz. Relational and directional in the construction of information granules. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 32(5):605–614, 2002.

[14] Y. H. Qian, J. Y. Liang, C. Y. Dang, H. Y. Zhang, and J. M. Ma. On the evaluation of the decision performance of an incomplete decision table. *Data and Knowledge Engineering*, 65(3):373–400, 2008.

[15] Y. H. Qian, J. Y. Liang, D. Y. Li, H. Y. Zhang, and C. Y. Dang. Measures for evaluating the decision performance of a decision table in rough set theory. *Information Science*, 178:181–202, 2008.

[16] A. Skowron and C. Rauszer. *The discernibility matrices and functions in information systems, In: R Slowiński(Eds),Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory*. Kluwer Academic Publisher, Dordrecht, 1992.

[17] D. Slezak. Approximate entropy reducts. *Fundamenta Informaticae*, 53(3):365–387, 2002.

[18] G. Y. Wang. Rough reduction in algebra view and information view. *International Journal of Intelligent Systems*, 18:679–688, 2003.

[19] J. T. Yao and M. Zhang. Feature selection with adjustable criteria. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Proceedings of RSFDGrC05*, LNAI 3641:204–213, 2005.

[20] Y. Y. Yao. Information granulation and rough set approximation. *Internatinal Journal of Intelligent Systems*, 16(1):87–104, 2001.

[21] Y. Y. Yao. Decision-theoretic rough set models. *Rough Sets and Knowledge Technology, RSKT 2007*, LNAI 4481:1–12, 2007.

[22] W. X. Zhang, W. Z. Wu, J. Y. Liang, and D. Y. Li. Theory and method of rough sets. 2001.

[23] W. Ziarko. Variable precision rough set model. *Journal of Computer and System Science*, 46(1):39–59, 1993.