Research Article

Uncertainty Analysis of Knowledge Reductions in Rough Sets

Ying Wang¹ and Nan Zhang²

¹ Department of Computer Science and Technology, Tongji University, 4800 Caoan Road, Shanghai 201804, China

² Department of Computer and Control Engineering, Yantai University, 32 Qingquan Road, Shandong 264005, China

Correspondence should be addressed to Ying Wang; wangy@tongji.edu.cn

Received 20 June 2014; Revised 3 August 2014; Accepted 3 August 2014; Published 27 August 2014

Academic Editor: Yunqiang Yin

Copyright © 2014 Y. Wang and N. Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Uncertainty analysis is a vital issue in intelligent information processing, especially in the age of big data. Rough set theory has attracted much attention to this field since it was proposed. Relative reduction is an important problem of rough set theory. Different relative reductions have been investigated for preserving some specific classification abilities in various applications. This paper examines the uncertainty analysis of five different relative reductions in four aspects, that is, reducts' relationship, boundary region granularity, rules variance, and uncertainty measure according to a constructed decision table.

1. Introduction

Uncertainty is associated with randomness, fuzziness, vagueness, roughness, and incomplete knowledge. The theories of probability, information, fuzzy set, evidence set, rough set, and so forth have been used for uncertainty analysis [1]. Rough set theory (RST) as a new comer and one of methods for the representation of uncertainty has gained an increasing attention from both the theoretical and the applied points of view. Uncertainty exists in realistic world inherently. There are many factors affecting the uncertainty of actual questions. Examining and analyzing the characteristics of uncertainty from various situations are vital in intelligent information processing.

Attribute reduction is an important problem of rough set theory. A reduct is a minimum subset of attributes that provides the same description or classification ability as the entire set of attributes [2]. A relative reduct in Pawlak rough set is a minimum subset of attributes that preserves the positive region of the classification unchangeable, and it can also be defined as a minimum subset of condition attributes that provides the same classification ability as the entire set of attributes in a decision table. Many objective functions for attribute reduction have been proposed and examined to find the reduct to preserve a specific property [3–10].

The relationship among various reductions has attracted interest of some researchers [3, 11–15]. This also inspires our interest to investigate the uncertainty variation under a reduction and the uncertainties in various relative reducts.

There are two types of uncertainty inherently in RST. The first type of uncertainty arises from the indiscernibility relation. It increases as the granularity of the partition becomes coarser. The second one results from the approximation regions of rough sets, since the lower approximation is the certain region and the upper approximation is the possible region. This gives rise to a direction of analyzing uncertainty in relative reductions.

In this paper, the uncertainty analysis of five relative reductions is investigated in four aspects. Firstly, the relationships among different reducts are listed according to some research results. Secondly, the boundary areas of five reductions are described in detail according to a constructed decision table. Thirdly, the quality analysis and variant process of classification rules generated from these five relative reductions are discussed. Lastly, the definition of uncertainty measure in classification is proposed.

In the remainder of this paper, some related notations are reviewed and uncertainty analysis of five relative reductions is discussed.

2. Basic Notions

Information table is used by Pawlak for raw data representation [2]. For classification tasks, we consider a special information table with a set of decision attributes. Such an information table is also called a decision table. In this part, only some related notations are reviewed.

Definition 1. A decision table is given as the following tuple:

$$S = (U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}), \quad (1)$$

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

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

A partition $\pi_D : U/D = \{D_1, D_2, \dots, D_{|U/D|}\}$ is used to denote the partition of the universe U defined by the decision attribute set D , and so is another partition π_B defined by a condition attribute set $B \subseteq C$. The equivalence classes induced by the partition π_B are the basic blocks to construct the Pawlak rough set approximations. For a decision class $D_i \in \pi_D$, the lower and upper approximations of D_i with respect to a partition π_B are defined by Pawlak [2]:

$$\begin{aligned} \underline{apr}_{\pi_B}(D_i) &= \{x \in U \mid [x]_B \subseteq D_i\} \\ \overline{apr}_{\pi_B}(D_i) &= \{x \in U \mid [x]_B \cap D_i \neq \emptyset\}. \end{aligned} \quad (2)$$

Set $BND_{\pi_B}(D_i) = \overline{apr}_{\pi_B}(D_i) - \underline{apr}_{\pi_B}(D_i)$ will be called the π_B -boundary of D_i .

Definition 2. The positive and boundary regions of the partition π_D with respect to a partition π_B are defined by Yao [16]:

$$\begin{aligned} POS_{\pi_B}(\pi_D) &= \bigcup_{1 \leq i \leq m} POS_{\pi_B}(D_i); \\ BND_{\pi_B}(\pi_D) &= \bigcup_{1 \leq i \leq m} BND_{\pi_B}(D_i). \end{aligned} \quad (3)$$

Theorem 3. In a decision table S , the positive and boundary regions of the partition π_D with respect to a partition π_B have the following properties [16].

- (1) If the decision table is consistent, then $POS_{\pi_B}(\pi_D) = U$ and $BND_{\pi_B}(\pi_D) = \emptyset$.
- (2) If the decision table is inconsistent, then $POS_{\pi_B}(\pi_D) \cup BND_{\pi_B}(\pi_D) = U$ and $BND_{\pi_B}(\pi_D) \neq \emptyset$.

Proposition 4. In a decision table S , if $BND_{\pi_B}(\pi_D) \neq \emptyset$, then S is inconsistent; otherwise, it is consistent.

Decision rule generation is an important issue in a decision table. The rough set approach offers all solutions to the problem of decision table simplification and many real applications have been found in various fields [17]. Rule sets may include certain and uncertain (or possible) rules. Theorem 3 can be written as Theorem 5 in the view of decision rules.

Theorem 5. In a decision table S , the rule sets have the following properties.

- (1) If the decision table is consistent, then any rule of the rule sets is a certain rule.
- (2) If the decision table is inconsistent, then the rule sets consist of certain rules and uncertain rules. And the number of uncertain rules is larger than zero.

Definition 6. In a decision table S , confidence of uncertainty rules of S with respect to a condition attribute set $B \subseteq C$ and an equivalence class $B_j \in \pi_B$ is defined as

$$\begin{aligned} \kappa(D_i \mid B_j) &= \frac{|(x, y) \in B_j \mid f_d(x) = f_d(y) = D_i|}{|(x, y) \in B_j|}, \\ i &= 1, 2, \dots, \left| \frac{U}{D} \mid B_j \right|, \end{aligned} \quad (4)$$

where $0 \leq \kappa \leq 1$ obviously. $\kappa(D_i \mid B_j)$ is written as κ_i when B_j is known. All the rules with respect to B_j in different D_i are called one pair of rules. The confidence of one pair of rules has the property that all the sum of κ_i is equal to 1. But the value range of i may be different with different subset B_j in partition π_B . A rule with premises (preconditions) P , conclusions (post conditions) Q , and confidence κ_i is denoted as $P \rightarrow \kappa_i Q$. If $\kappa_i = 1$, then the rule is certain; if $\kappa_i = 0$, then the rule is impossible; otherwise, it is possible or uncertain.

In order to express the degree of completeness and incompleteness of knowledge R about the nonempty set X , a pair of measures is defined by Pawlak [18]. They are the accuracy measure of $\alpha_R(X) = (\text{card } \underline{R}(X) / \text{card } \overline{R}(X))$ and R -roughness of $\rho_R(X) = 1 - \alpha_R(X) = 1 - (\text{card } \underline{R}(X) / \text{card } \overline{R}(X)) = (\text{card } R_{BN}(X) / \text{card } \overline{R}(X))$.

Pawlak also defines a measure to evaluate the quality of approximation of π_D by π_B [18], $\gamma_{\pi_B}(\pi_D) = (|POS_{\pi_B}(\pi_D)| / |U|)$. Similarly, $\rho_{\pi_B}(\pi_D) = 1 - \gamma_{\pi_B}(\pi_D) = (|BND_{\pi_B}(\pi_D)| / |U|)$; a variety of $\gamma_{\pi_B}(\pi_D)$ can be defined as and referred to as π_B -roughness of classification or approximation. It can express the degree of inexactness of knowledge π_B about the classification π_D .

This method to measure uncertainty is only associated with object numbers in the boundary region and the universe of discourse not the granularity of discourse. A modified measure, the rough entropy, is proposed by Beaubouef et al. [19], which combined the roughness of a set with approximate granularity:

$$Er(X) = -(\rho_R(X)) \left[\sum Q_i \log(P_i) \right]. \quad (5)$$

According to Beaubouef, the rough entropy of each decision class is the probabilities for each equivalence class belonging either wholly or in part to it. There is no ordering associated with individual class members. Therefore the probability of any one value of the class being named is the reciprocal of the number of elements in the class. If c_i is the cardinality of, or the number of elements in, equivalence class i and all members of a given equivalence class are equal, then $P_i = 1/c_i$ represents the probability of one of the values in class i . Q_i denotes the probability of equivalence class i within the universe. Q_i is computed by taking the number of elements in class i and dividing by the total number of elements in all equivalence classes combined.

The rough entropy $Er(X)$ indicates the uncertain percentage concerning granularity, but it is only defined for a set. By generalizing it to a partition, a definition of rough entropy for a classification is proposed in Section 3.4.

In order to support our discussion, the road traffic accident table is constructed by our group and is shown in Table 1. It is inconsistent obviously. The total number of data items is 15; among them only items 1 and 2 are consistent and others are inconsistent.

From Table 1, the partition π_D derived can be written as

$$\begin{aligned} D_1 &= \{x \in U \mid g(x) = N\} = \{x_1, x_5, x_6, x_9\}; \\ D_2 &= \{x \in U \mid g(x) = S\} \\ &= \{x_2, x_3, x_4, x_7, x_8, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}\}; \\ D_3 &= \{x \in U \mid g(x) = L\} = \{x_{15}\}. \end{aligned} \quad (6)$$

The partition π_C or equivalent classes $IND(C)$ derived by condition attribute set C can be written as

$$\begin{aligned} IND(C) &= \{\{x_1\}, \{x_2\}, \{x_3, x_4, x_5\}, \\ &\quad \{x_6, x_7, x_8\}, \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (7)$$

The positive and boundary regions of decision attribute with respect to condition attribute set C are as follows.

Positive region is

$$POS_{\pi_C}(\pi_D) = \{\{x_1\}, \{x_2\}\}. \quad (8)$$

Boundary region is

$$\begin{aligned} BND_{\pi_C}(\pi_D) &= \{\{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}, \\ &\quad \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (9)$$

The quality of classification or the degree of dependency of π_D on π_C defined by Pawlak in Table 1 is as follows: $\gamma_{\pi_C}(\pi_D) = 2/15$.

The π_C -roughness of classification is given below. Its value denotes the uncertain percentage of the discourse.

π_C -roughness is

$$\rho_{\pi_C}(\pi_D) = \frac{|BND_{\pi_C}(\pi_D)|}{|U|} = \frac{13}{15} = 0.867. \quad (10)$$

TABLE 1: Road traffic accident table.

Instant	Condition attributes						Decision attribute
	a	b	c	d	e	f	g
1	N	N	N	G	G	N	N
2	Y	N	N	G	G	N	S
3	Y	Y	N	G	G	N	S
4	Y	Y	N	G	G	N	S
5	Y	Y	N	G	G	N	N
6	Y	Y	N	G	B	Y	N
7	Y	Y	N	G	B	Y	S
8	Y	Y	N	G	B	Y	S
9	Y	Y	N	B	B	Y	N
10	Y	Y	N	B	B	Y	S
11	Y	Y	N	B	B	Y	S
12	Y	Y	N	B	B	Y	S
13	Y	Y	N	B	B	Y	S
14	Y	Y	Y	B	B	Y	S
15	Y	Y	Y	B	B	Y	L

Notes: ① a: drunk, b: tired, c: speedy, d: weather, e: road, f: mountains, and g: accident.

② Y: yes, N: no, G: good, B: bad, S: small, and L: large.

TABLE 2: The confidence of rules in Table 1.

Rules	Conditional attributes						Decision attribute	Validity
	a	b	c	d	e	f	g	κ
1	N	N	N	G	G	N	N	1
2	Y	N	N	G	G	N	S	1
3	Y	Y	N	G	G	N	N	1/3
4	Y	Y	N	G	G	N	S	2/3
5	Y	Y	N	G	B	Y	N	1/3
6	Y	Y	N	G	B	Y	S	2/3
7	Y	Y	N	B	B	Y	N	1/5
8	Y	Y	N	B	B	Y	S	4/5
9	Y	Y	Y	B	B	Y	S	1/2
10	Y	Y	Y	B	B	Y	L	1/2

Notes: κ -confidence, others are the same as in Table 1.

The quality of classification, $\gamma_{\pi_C}(\pi_D)$, denotes the percentage of objects to all objects in universe that are certainly classified to a decision class. We can use $\rho_{\pi_C}(\pi_D)$, namely, $1 - \gamma_{\pi_C}(\pi_D)$, to denote the percentage of uncertainly classified objects to all objects in universe and call it the quality of classification uncertainty. These two parameters of inconsistent decision table retain constant in reductions which preserve the positive region invariant.

The rules generated from Table 1 are listed in Table 2. These are primitive rules obtained from original data. The rule has the form $P \rightarrow \kappa Q$, denoting that the confidence of the classification rule is κ on the value of attributes set P . The first rule in line one from the top can be written as $(a = N, b = N, c = N, d = G, e = G, f = N) \rightarrow {}_1g = N$.

From all 15 data items, 10 rules are achieved as shown in Table 2. The first two of them are certainty rules, and other 8 possible rules are uncertainty rules and appear in four pairs.

TABLE 3: Five different reduct objections.

Year	Authors	Reduction objection	Reduction function
1982	Pawlak [2]	Positive region	$\text{POS}_B(D) = \text{POS}_C(D)$
1997	Miao [4]	Mutual information	$I(B : D) = I(C : D)$
2000	Slezak [5]	Distribution	$\mu_B(x) = \mu_C(x)$
2001	Kryszkiewicz [6]	General decision	$\tau_B(x) = \tau_C(x)$
2009	Miao et al. [3]	Boundary Partition	$\text{IND}(B/D) = \text{IND}(C/D)$

The four pairs are rules 3 and 4, rules 5 and 6, rules 7 and 8, and rules 9 and 10. The confidence degree of each rule is also listed in the table.

3. Uncertainty Analysis of Relative Reductions

There are various definitions of relative reductions in rough set theory [2–8, 11, 12]. All of them can be classified into three categories: region preservation, information preservation, and partition preservation.

Five relative reductions [2–6] chosen from the three categories will be investigated for uncertainty analysis in the following four aspects (Table 3). They are classical positive region reduction proposed by Pawlak [2], mutual information preservation reduction proposed by Miao [4], distribution reduction proposed by Slezak [5], general decision reduction proposed by Kryszkiewicz [6], and boundary partition reduction proposed by Miao et al. [3]. Related articles may be inquired for better explanation of these relative reductions.

3.1. The Relationships among Relative Reducts. A reduct is a minimum subset of attributes that provides the same description or classification ability as the entire set of attributes. The relationship among different relative reducts has been studied by many researchers. In order to describe the relationship among the five properties preservation reductions, some definition is given in the first place.

Definition 7. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$, its property ψ_1 preservation reduct is R_1 and property ψ_2 preservation reduct is R_2 .

- (1) If R_2 always preserves property ψ_1 simultaneously, then R_2 is defined as a ψ_1 reduct container.
- (2) If R_2 is a ψ_1 reduct container, then ψ_2 reduction is defined as a stronger one and ψ_1 reduction is a weaker one.

Because the reduct of property preservation is nonexclusive, we cannot say that the reduct of a weaker reduction is included in the reduct of a stronger one. We can only say that if ψ_2 reduction is stronger than ψ_1 reduction, then $\exists$ reducts R_2 and R_1 of properties ψ_2 and ψ_1 reduction, and $R_1 \subseteq R_2$.

According to Definition 7, some research results can be assembled here as the following theorems in inconsistent decision tables. The proof of these theorems can be seen in listing reference documents.

FIGURE 1: Relations of five relative reducts.

Theorem 8. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$, the boundary partition preserving reduct must be a distribution preserving reduct container [11].

Theorem 9. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$ and any subset B of attributes C , $B \subseteq C$, then $H(B \mid D) = H(C \mid D)$ if and only if $\forall x \in U, \exists \tau_B(x) = \tau_C(x)$ [13].

This means that a distribution reservation reduction is equivalent to mutual information preserving reduction. The distribution reservation reduction will be used in later of this paper as these two reductions.

Theorem 10. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$, the distribution preserving reduct must be a general decision preserving reduct container [12].

Theorem 11. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$, a general decision preserving reduct must be a positive preserving reduct; the boundary partition preserving reduct must be a general decision preserving reduct container [3].

For good understanding, the relationships among five relative reductions are stated in Figure 1 in decision tables. The two-headed arrow expresses the equivalent relation and one-way arrow expresses the container relation. As shown in the figure, a distribution reservation reduct is equivalent to mutual information preserving reduct. A boundary partition reservation reduct must be container for any other four reducts. In other words, in the five relative reductions,

a boundary partition reservation reduction is the strongest; a positive region reduction is the weakest.

Five reducts of decision (Table 1) can be computed according to their reduction definitions. The computed results are as follows.

- (i) A positive region reservation reduct is

$$\text{RED}_{\text{POS}} = \{a, b\}. \quad (11)$$

- (ii) A general decision reservation reduct is

$$\text{RED}_{\text{GEN}} = \{a, b, c\}. \quad (12)$$

- (iii) A distribution or a mutual information reservation reduct is

$$\text{RED}_{\text{DIS}} = \text{RED}_{\text{MIF}} = \{a, b, c, d\}. \quad (13)$$

- (iv) A boundary partition reservation reduct and boundary region are

$$\text{RED}_{\text{BPA}} = \{a, b, c, d, e\}. \quad (14)$$

In the process of computing the above five reducts, the attribute adding guideline is adopted. So the reducts meet set inclusion relationship as the reduction becomes stronger. That is,

$$\text{RED}_{\text{POS}} \subseteq \text{RED}_{\text{GEN}} \subseteq \text{RED}_{\text{DIS}} \subseteq \text{RED}_{\text{BPA}}. \quad (15)$$

This expression is not always correct because of the nonexclusive reduct of a reduction.

3.2. The Boundary Areas of Five Reductions. Now, the problem will be investigated further from another view. From Theorem 3, we know that the universe of discourse is the sum of positive region and boundary region. Because all the five reductions discussed in this paper preserve the same positive region with the original table (see Section 3.1), the boundary regions of the five reductions are also the same from the view of object set. Now we examine the granularity in the boundary regions from granular computing. The computation results are as follows.

- (i) The reduct and boundary region of positive region reduction are

$$\text{RED}_{\text{POS}} = \{a, b\}, \quad \text{BND}_{\text{POS}} = \{x_3, x_4, \dots, x_{15}\}. \quad (16)$$

- (ii) The reduct and boundary region of general decision reduction are

$$\begin{aligned} \text{RED}_{\text{GEN}} &= \{a, b, c\}, \\ \text{BND}_{\text{GEN}} &= \{\{x_3, x_4, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (17)$$

- (iii) The reduct and boundary region of distribution reduction or mutual information reduction are

$$\begin{aligned} \text{RED}_{\text{DIS}} &= \{a, b, c, d\}, \\ \text{BND}_{\text{DIS}} &= \{\{x_3, x_4, \dots, x_8\}, \\ &\quad \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (18)$$

FIGURE 2: Boundary regions of five relative reductions.

- (iv) The reduct and boundary region of boundary partition reduction are

$$\begin{aligned} \text{RED}_{\text{BPA}} &= \{a, b, c, d, e\}, \\ \text{BND}_{\text{BPA}} &= \{\{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}, \\ &\quad \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (19)$$

The boundary regions of five relative reductions are illustrated in Figure 2. By observing the granularity of the boundary regions, we can get the following conclusion.

Proposition 12. If ψ_1 reduction is weaker than ψ_2 reduction and $R_1 \subseteq R_2$, then the knowledge on ψ_2 reduction is finer than the knowledge on ψ_1 reduction or the knowledge on ψ_1 reduction is coarser than the knowledge on ψ_2 reduction.

In RST, if R_1 and R_2 are two attribute sets, and $R_1 \subseteq R_2$, then we have $\text{IND}(R_1) \supseteq \text{IND}(R_2)$. So Proposition 12 can directly result from Pawlak's definition of knowledge granule in reference [18].

3.3. The Classification Rules in Five Relative Reductions. The quality of classification rules may be measured by rule confidence. Confidence is used to estimate the degree of validity of rules. It means the fraction of objects satisfying both the premises and the conclusions of rules in the set of all objects satisfying the premises of rules. In one rule pair, the higher the confidence of a rule is, the more valid the rule is.

All rules generated from five relative reducts and their confidences are shown in Table 4. This table is very complicated because so much information assembled in it. The upper section of the table contains all rules with premises (shown as Cond. attr.), conclusions (shown as Dec. attr.), and confidence. The lower section of the table expresses the inclusion relationship of upper cells and a specific reduction.

Referring to the fourth row from the bottom, condition attributes aligning with cell, Pos reduct, are the reduct of positive region reduction, $\text{RED}_{\text{POS}} = \{a, b\}$. The confidences

TABLE 4: The rules description of five reductions.

		Cond. attr.			Dec. attr.		Confidence		
a	b	c	d	e	g	POS	GDR	DR	BPR
N	N	N	G	G	N	1	1	1	1
Y	N	N	G	G	S	1	1	1	1
		N		G	N	3/13	3/11	2/6	1/3
			G		S	9/13	8/11	4/6	2/3
		Y		B	S				1/3
					L				2/3
Y	Y	N		B	N			1/5	1/5
			B		S			4/5	4/5
		Y		B	S		1/2	1/2	1/2
					L	1/13	1/2	1/2	1/2
Pos reduct		5 rules and one pair of R.unc.							
General decision		6 rules and two pairs of R.unc.							
Distribution reduct		8 rules and three pairs of R.unc.							
Boundary partition reduct		10 rules and four pairs of R.unc.							

Notes: Cond. attr., Dec. attr., and R.unc. are abbreviations for condition attributes, decision attributes, and uncertain rules.

of five rules generated from positive reduct are listed in the nonzero cells in fourth column from right. Two of them are certain rules. Other three of them are possible rules and appear in one pair which has the same condition attributes but different decision attributes. They can be written in the form:

$$\begin{aligned}
 (a = N, b = N) &\longrightarrow {}_1g = N, \\
 (a = Y, b = N) &\longrightarrow {}_1g = S, \\
 (a = Y, b = Y) &\longrightarrow {}_{3/13}g = N, \\
 (a = Y, b = Y) &\longrightarrow {}_{9/13}g = S, \\
 (a = Y, b = Y) &\longrightarrow {}_{1/13}g = L.
 \end{aligned} \tag{20}$$

Similarly, condition attributes aligning with cell, boundary partition reduct, in the first row from the bottom are the reduct of boundary partition reduction, $RED_{BPA} = \{a, b, c, d, e\}$. The confidences of ten rules generated from boundary partition reduct are listed in the first column in cell confidence from right. They are the same in second and third rows from the bottom.

The change of uncertainty rules in different relative reductions follows a regular pattern. The pair numbers of possible rules are equal to the reduct attributes value numbers of inconsistent data in the table. The number of rules in one pair equals the number of decision values with the same reduct attributes value.

Uncertainty rules that resulted from different reductions can be compared under certain conditions. Because the reduct in rough sets is nonexclusive, only those uncertainty rules that resulted from reducts with set inclusion relation can be compared. This condition is satisfied for rules as shown in Table 4.

Considering the rules from two different reductions, pairs of uncertainty rules in stronger reduction may be integrated

into pairs of uncertainty rules in weaker one. In other words, the pair number of uncertainty rule in stronger reduction is larger than or equal to that in weaker one. The number of rules in one pair remains unchanged or gets larger in weaker reduction because of more decision values.

The rule confidence of new formed pair changes because some rules are combined into new pairs and the data item supporting the rules combined at the same time. As an example, when changing from GDR reduction to POS reduction, two pairs of possible rules become one pair. In combination with rule pairs of confidence (3/11, 8/11) and (1/2, 1/2), the reduct value numbers decrease from two to one and the decision values increase from two to three. Combining the data items of previous two pairs with the same decision value, new rules confidences are calculated. They are $9/13 = (8 + 1)/(11 + 2)$, $3/13 = (3 + 0)/(11 + 2)$, and $1/13 = (1 + 0)/(11 + 2)$ as shown in Table 4.

The rule confidence only makes sense in its pairs. Let us inspect the last row in the upper part; rule confidence based on the positive preservation is 1/13 and the general decision preservation is 1/2. It never means that the later rule is more reliable than the former. Rule confidence denotes that if an object meets the preconditions, then it has the probability to have the conclusion. The rule confidence in different pairs cannot be compared.

3.4. The Uncertainty Measure of Rough Sets. The uncertainty measure is an important subject in RST. Various definitions of uncertainty measure have been proposed by many researchers [16, 18–21]. All these measures can reflect the uncertainty of relative reductions to some extent.

Upon the rough entropy definition of Theresa B in 1998 and π_B -roughness of classification/approximation in Section 2, we give the rough entropy Definition 13 for classification task. Both the roughness and the granularity are concerned in this measure.

Definition 13. Given a decision table $S = \{U, A = C \cup D, \{V_a \mid a \in A\}, \{I_a \mid a \in A\}\}$, a condition attribute set $B \subseteq C$, the rough entropy $Er_{\pi_B}(\pi_D)$ of a classification is calculated by

$$Er_{\pi_B}(\pi_D) = (\rho_{\pi_B}(\pi_D)) \frac{\sum_{j \in D_j \in U/D} \sum_i (Q_{ij} \log(P_i))}{|U/D| \log(1/|U|)}. \quad (21)$$

The term $\rho_{\pi_B}(\pi_D)$ denotes the roughness of the classification of π_D with respect to partition π_B governed by attribute set B . The second term is a fraction whose nominator is the summation of rough entropy of each decision class D_j , $j = 1, 2, \dots, |U/D|$, and denominator is normalized factor to ensure the value of fraction to be less than 1. So the rough entropy is less than the roughness of a classification.

On the basis of road traffic accident table (Table 1) and the reducts of five relative reductions, the computation results are as follows.

From Table 1, the partition π_D derived can be written as

$$\begin{aligned} D_1 &= \{x \in U \mid g(x) = N\} = \{x_1, x_5, x_6, x_9\}; \\ D_2 &= \{x \in U \mid g(x) = S\} \\ &= \{x_2, x_3, x_4, x_7, x_8, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}\}; \\ D_3 &= \{x \in U \mid g(x) = L\} = \{x_{15}\}. \end{aligned} \quad (22)$$

The partition π_C or equivalent classes $IND(C)$ derived by condition attribute set C can be written as

$$IND(C) = \{\{x_1\}, \{x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}, \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \quad (23)$$

The partition or equivalent classes derived by reducts of five relative reductions are as follows:

$$\begin{aligned} IND(RES_{POS}) &= \{\{x_1\}, \{x_2\}, \{x_3, x_4, \dots, x_{15}\}\}, \\ IND(RES_{GEN}) &= \{\{x_1\}, \{x_2\}, \{x_3, x_4, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}, \\ IND(RES_{DIS}) &= \{\{x_1\}, \{x_2\}, \{x_3, x_4, \dots, x_8\}, \\ &\quad \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}, \\ IND(RES_{BPA}) &= \{\{x_1\}, \{x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}, \\ &\quad \{x_9, x_{10}, \dots, x_{13}\}, \{x_{14}, x_{15}\}\}. \end{aligned} \quad (24)$$

In this example, $IND(RES_{BPA})$ is equal to $IND(C)$, so the rough entropy of boundary partition preservation reduction equals that of original data in Table 1. It is just a special case. If two reductions are equivalent, as for distribution preservation reduction and mutual information preservation reduction, all measures for them will be equal.

The values of rough entropy of five relative reductions are worked out and shown in Table 5. The rough entropy $Er_{\pi_B}(\pi_D)$ of a classification in Definition 13 can be used effectively to describe the uncertainty measure. The results meet Proposition 14 very well.

TABLE 5: The uncertainty measure of five relative reductions.

Reduction	Raw data	POS	GEN	DIS	BPA
Rough entropy	0.214	0.712	0.395	0.287	0.214

Proposition 14. In an inconsistent decision table, the uncertainty of classification will increase after a reduction. The weaker is the reduction; the larger is its rough entropy (uncertainty of classification).

When a problem is resolved at a reduced attributes set after a reduction, its computational complexity decreases and uncertainty increases. This uncertainty variance can be observed from the following four aspects. The first one is the reduct, a reduced attribute set. The second one is the granularity of the universe of discourse, including positive and boundary regions. The third one is the number of rule pairs. The last is reflected in the measure, rough entropy of a classification.

4. Conclusion

A relative reduct is a minimum set of attributes that keeps a particular classification property. Different property preservation reductions bring about different reducts. In this paper, the uncertainty analysis of five relative reductions in an inconsistent decision table is discussed for classification tasks from four aspects. In the former three points, the relationships among reduct sets, boundary regions, and rules produced by five relative reductions are compared, respectively. In the last section rough entropy, a measure of uncertainty for a decision table is proposed for classification task. The conclusion is obtained that the weaker reduction has more uncertainty than the stronger one in an inconsistent decision table.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (nos. 61075056 and 61273304). This work is also partially supported by the Natural Science Foundation of Shandong Province (no. ZR2013FQ020) and Research Foundation of Yantai University (no. JS12B28).

References

- [1] D. Q. Miao, D. Y. Li, Y. Y. Yao et al., *Uncertainty and Granular Computing*, Science Press, Beijing, China, 2011, (Chinese).
- [2] Z. Pawlak, "Rough sets," *International Journal of Computer & Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [3] D. Q. Miao, Y. Zhao, Y. Y. Yao, H. X. Li, and F. F. Xu, "Relative reducts in consistent and inconsistent decision tables of

- the Pawlak rough set model," *Information Sciences*, vol. 179, no. 24, pp. 4140–4150, 2009.
- [4] D. Q. Miao, *The research of rough set theory and application in machine learning [Doctorate]*, Institute of Automation of Chinese Academy of Sciences, Beijing, China, 1997.
 - [5] D. Slezak, "Normalized decision functions and measures for inconsistent decision tables analysis," *Fundamenta Informaticae*, vol. 44, no. 3, pp. 291–319, 2000.
 - [6] M. Kryszkiewicz, "Comparative study of alternative types of knowledge reduction in inconsistent systems," *International Journal of Intelligent Systems*, vol. 16, pp. 105–120, 2001.
 - [7] Q. H. Hu, Z. X. Xie, and D. R. Yu, "Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation," *Pattern Recognition Letters*, vol. 40, no. 12, pp. 3509–3521, 2007.
 - [8] Q. H. Hu, D. R. Yu, and Z. X. Xie, "Information-preserving hybrid data reduction based on fuzzy-rough techniques," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 414–423, 2006.
 - [9] X. Hu and N. Cercone, "Learning in relational databases: a rough set approach," *Computational Intelligence*, vol. 11, no. 2, pp. 323–338, 1995.
 - [10] D. Ye and Z. Chen, "A new type of attribute reduction for inconsistent decision tables and its computation," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 18, no. 2, pp. 209–222, 2010.
 - [11] W. F. Du and K. Y. Qin, "Analysis of standard and relation of several attribute reducts in inconsistent decision table," *Chinese Journal of Electronic*, vol. 39, no. 6, pp. 1336–1340, 2011.
 - [12] W. Zhang, J. Mi, and W. Wu, "Knowledge reductions in inconsistent information systems," *Chinese Journal of Computers*, vol. 26, no. 1, pp. 12–18, 2003.
 - [13] Z. Xu, B. Yang, and W. Song, "Comparative study of different attribute reduction based on decision table," *Chinese Journal of Electronics*, vol. 15, no. 4A, pp. 953–956, 2006.
 - [14] J. Zhou, D. Miao, W. Pedrycz, and H. Zhang, "Analysis of alternative objective functions for attribute reduction in complete decision tables," *Soft Computing*, vol. 15, no. 8, pp. 1601–1616, 2011.
 - [15] G. Y. Wang, J. Zhao, J. An, and Y. Wu, "A comparative study of algebra viewpoint and information viewpoint in attribute reduction," *Fundamenta Informaticae*, vol. 68, no. 3, pp. 289–301, 2005.
 - [16] Y. Y. Yao, "Decision-theoretic rough set models," in *Proceedings of the 2nd International Conference on Rough Sets and Knowledge Technology*, vol. 4481 of *Lecture Notes in Artificial Intelligence*, p. 112, 2007.
 - [17] A. Skowron, "Boolean reasoning for decision rules generation," in *Proceedings of the International Symposium on Methodologies for Intelligent Systems*, pp. 295–305, 1993.
 - [18] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic, Boston, Mass, USA, 1991.
 - [19] T. Beaubouef, F. E. Petry, and G. Arora, "Information-theoretic measures of uncertainty for rough sets and rough relational databases," *Information Sciences*, vol. 109, no. 1–4, pp. 185–195, 1998.
 - [20] I. Düntsch and G. Gediga, "Uncertainty measures of rough set prediction," *Artificial Intelligence*, vol. 106, no. 1, pp. 109–137, 1998.
 - [21] J. Liang and Z. Shi, "The information entropy, rough entropy and knowledge granulation in rough set theory," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 12, no. 1, pp. 37–46, 2004.