Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/080456

Information Entropy and Mutual Information-based Uncertainty Measures in Rough Set Theory

Lin Sun[1,2,3,*] *and Jiucheng Xu*[1,2]

[1] College of Computer and Information Engineering, Henan Normal University, 453007 Xinxiang, China
[2] Engineering Technology Research Center for Computing Intelligence and Data Mining, Henan Province, China
[3] International WIC Institute, Beijing University of Technology, 100124 Beijing, China

Received: 25 Aug. 2013, Revised: 27 Nov. 2013, Accepted: 28 Nov. 2013
Published online: 1 Jul. 2014

Abstract: As an extension of the classical set theory, rough set theory plays a crucial role in uncertainty measurement. In this paper, concepts of information entropy and mutual information-based uncertainty measures are presented in both complete and incomplete information/decision systems. Then, some important properties of these measures are investigated, relationships among them are established, and comparison analyses with several representative uncertainty measures are illustrated as well. Theoretical analysis indicates that these proposed uncertainty measures can be used to evaluate the uncertainty ability of different knowledge in complete/incomplete decision systems, and then these results can be helpful for understanding the essence of knowledge content and uncertainty measures in incomplete information/decision systems. Thus, these results have a wide variety of applications in rule evaluation and knowledge discovery in rough set theory.

Keywords: Rough set theory, information entropy, conditional information entropy, mutual information, uncertainty measure

1 Introduction

Rough set theory, developed by Pawlak [1] , has become a useful mathematic tool for dealing with vague and uncertain information in many areas, such as pattern recognition, feature selection, neural computing, decision support, data mining and knowledge discovery process from big data sets [2,3,4,5,6]. Uncertainty measure, as one of the most important issues in rough set theory, plays an important role in artificial intelligence and reasoning with uncertainty [7,8,9,10,11]. As follows, we briefly review some relevant literatures.

To evaluate uncertainty of a set, Pawlak presented several numerical measures in pure rough set theory, which are accuracy and roughness of a set and approximation accuracy of a rough classification [1]. The accuracy measure and the roughness measure are important numerical characterizations that quantify the imprecision of a rough set caused by its boundary region. Although these measures are effective, they have some restrictions [12,13,14]. Therefore, the applications of rough set theory in some fields are limited. For this reason, Xu et al. [15] gave an improved accuracy measure for rough sets, which calculated the imprecision of a set by using an excess entropy. However, this measure has a complex mathematical form. Recently, Yao [16] studied two definitions of approximations and associated measures based on equivalence relations. To evaluate uncertainty of a system, the entropy of a system was introduced by Shannon in [17]. It is a very useful mechanism for characterizing the information content in various modes and has been applied in many diverse fields [18,19,20,21]. The entropy and its variants were adapted to rough set theory in [22] and information interpretation of rough set theory was given in [10]. However, Shannon's entropy is not a fuzzy entropy, and cannot measure the fuzziness in rough set theory [23]. A new information entropy was proposed by Liang in [10], and then some important properties of this entropy were derived as well. Unlike the logarithmic behavior of Shannon's entropy, Liang's entropy can be used to measure the fuzziness of a rough set and a rough classification. Hu and Yu [24] redefined the joint entropy and conditional entropy based on Yager's work. He extended measures and then successfully used them to reduce hybrid attribution and measure uncertainty of

* Corresponding author e-mail: linsunok@gmail.com

fuzzy probability approximation spaces. Mi et al. [25] introduced an uncertainty measure for partition-based fuzzy-rough set model. Liang et al. [26] proposed a new method to measure the uncertainty of a set in an information system and the approximation accuracy of a rough classification in a decision table. Shi and Gong [27] defined rough entropy and granulation of covering, and then used them to characterize the uncertainty of covering for covering approximation space. Unlike most existing information entropies, Qian and Liang [7] proposed combination entropy to evaluate uncertainty of knowledge from an information system. The notion of information systems provides a convenient tool for the representation of objects in terms of their attribute values. According to whether or not there are missing data (null values), information systems can be classified into two categories: complete and incomplete [7]. However, all these studies were dedicated to evaluating uncertainty of a set in terms of the partition ability of knowledge. Since the equivalence classes are only regarded as the unit of information granule of a complete information system, these measures cannot be used to deal with an incomplete information system. Moreover, it is difficult to generalize the results of complete information systems to incomplete information systems [7]. In some cases, the uncertainty of a rough set cannot be well characterized by the existing measures. In this paper, we aim at solving this problem.

Classical rough set model is based on equivalence relation or partition, but this condition is difficult to be satisfied in many information systems [28]. Thus, the corresponding uncertainty measures above are not suitable for incomplete information systems. What's more, there are few studies on uncertainty measure issues in incomplete information/decision systems by pure rough set approach. To solve this issue, several interesting and meaningful extensions to equivalence relation have been proposed, such as tolerance relations [7,8,9,10,28, 29], covering rough sets [27,30], dominance-based rough sets [4,31,32], neighborhood operators [33,34], others [35]. However, the covering model is only suitable for information systems that contain features with multiple values, the dominance model is mainly suitable for knowledge acquisition in the incomplete decision system with preference-ordered domains of features, and the neighborhood model may be appropriate for dealing with numerical and categorical features by assigning different thresholds for different kinds of features [36]. These methods are usually considered as extensions of classical rough set theory. In fact, its extended models have been increasingly drawing people's attention. Based on the consideration, Qian et al. [37] proposed the conditional combination entropy, mutual information and defined a variety of combination entropy with maximal consistent block in incomplete information system. Dai and Xu [38] extended Pawlak's pure rough set uncertainty measures to incomplete information systems. However, these measures mainly focus on incomplete information systems rather than incomplete decision systems. Dai et al. [39] presented a kind of conditional entropy for incomplete decision systems. However, their conditional entropy is not monotonic, which makes it not so reasonable to evaluate the uncertainty in incomplete decision systems. Sun et al. [36] also investigated rough entropy-based uncertainty measures to evaluate the roughness and accuracy of knowledge in incomplete decision systems. So far, there are relatively few studies on uncertainty measures in incomplete decision systems. Thus, further studies on uncertainty measures for incomplete decision systems are necessary. Therefore, it is desirable to extend and hybridize these measures to deal with complete/incomplete data and solve many real world problems.

In this paper, the main objective is to construct information entropy and mutual information-based uncertainty measures for both complete and incomplete information/decision systems, and discuss their important properties and propositions by information theory approach. Then, relationships among these measures are investigated and comparison analyses with several representative uncertainty measures are illustrated. So far, however, the relationships have not been reported in an incomplete information/decision system, which would baffle further research and application of information entropy theory. Therefore, these proposed measures can provide important approaches to measuring the uncertainty ability of different knowledge in complete/incomplete decision systems, and then these results may be helpful for rule evaluation and knowledge discovery in complete/incomplete information systems. The rest of this paper is organized as follows. Some preliminary concepts are briefly reviewed in Section 2. In Section 3, the concepts of information entropy are introduced, and joint information entropy and conditional information entropy are presented to both complete and incomplete information/decision systems. Then comparison analyses of the proposed measures with several representative uncertainty measures are illustrated. Their important properties and propositions are induced, and then relationships among these measures are investigated as well. In Section 4, mutual information-based uncertainty measures are proposed to measure the uncertainty of both complete and incomplete information/decision systems, and several useful properties are derived. Finally, Section 5 concludes the paper.

2 Preliminaries

In rough set theory, an information system (IS) is a pair $IS = (U, A)$, where U is a non-empty finite set of objects, A is a non-empty finite set of attributes, and for any $a \in A$, there is a mapping a, a: $U \to V_a$, where V_a denotes the value domain of attribute a. With any subset of attributes $P \subseteq A$, there is a binary indistinguishable relation $IND(P)$ as follows: $IND(P) = \{(u,v) \in U \times U | \forall a \in P, a(u) = a(v)\}$.

For any $P \subseteq A$, the relation $IND(P)$ constitutes a partition of U, which is denoted by $U/IND(P)$ or just U/P. For any object $u \in U$, let $[u]_P$ denote the equivalence class of relation $IND(P)$, i.e., $[u]_P = \{v \in U|(u,v) \in IND(P)\}$. Each equivalence class $[u]_P$ is viewed as an information granule. Let IS be an information system. We define a partial relation $\preceq$ on the family $\{U/P|P \subseteq A\}$ as follows: $U/A \preceq U/P$ ($U/P \succeq U/A$) if and only if for every $X_i \in U/A$, there exists $Y_j \in U/P$ such that $X_i \subseteq Y_j$, where $U/A = \{X_1, X_2, \cdots, X_m\}$ and $U/P = \{Y_1, Y_2, \cdots, Y_n\}$ are partitions induced by A and P respectively. In this case, we say that P is coarser than A, or A is finer than P. If $U/A \preceq U/P$ and $U/P \succeq U/A$, we say that $P = Q$. If $U/A \preceq U/P$ and $U/A \neq U/P$, we say that P is strictly coarser than A (or A is strictly finer than P), denoted by $U/A \prec U/P$ (or $U/P \succ U/A$).

In an information system, it may occur that some of the attribute values for an object are missing. These missing values can be represented by the set of all possible values for the attribute or equivalence by the domain of the attribute. To indicate such a situation, a distinguished value is usually assigned to those attributes. If V_a contains a null value for at least one attribute $a \in A$, then the $IS = (U, A)$ is called an incomplete information system (IIS), otherwise it is a complete information system (CIS). Further on, the symbol $*$ denotes the missing value. Let IIS be an incomplete information system, $P \subseteq A$ an attribute subset. The subset P determines a binary relation on U as follows: $SIM(P) = \{(u,v) \in U \times U | \forall a \in P, a(u) = a(v)$ or $a(u) = *$ or $a(v) = *\}$. In fact, $SIM(P)$ is a tolerance relation on U and the concepts of a tolerance relation have a wide variety of applications in classification. It shows that $SIM(P) = \cap_{a \in P} SIM(\{a\})$ easily. Let $S_P(u)$ denote the set $\{v \in U|(u,v) \in SIM(P)\}$. Generally, $S_P(u)$ denotes the maximal set of objects which are possibly indistinguishable by P with object u. Let $U/SIM(P)$ denote the family sets $\{S_P(u)|u \in U\}$, the classification or the knowledge induced by P. A member $S_P(u)$ from $U/SIM(P)$ will be called a tolerance class or a granule of information. It should be noted that the tolerance classes in $U/SIM(P)$ do not constitute a partition of U in general. They constitute a cover of U, i.e., $S_P(u) \neq \emptyset$, for every $u \in U$, and $\cup_{u \in U} S_P(u) = U$.

Let IIS be an incomplete information system with $P, Q \subseteq A$. We define a partial relation $\preceq$ on 2^A as follows: P is finer than Q (or Q is coarser than P), denoted by $P \preceq Q$ (or $Q \succeq P$), if and only if $S_P(u_i) \subseteq S_Q(u_i)$ for any $i \in \{1, 2, \cdots, |U|\}$. In fact, $P \prec Q \Leftrightarrow$ it follows that $S_P(u_i) \subseteq S_Q(u_i)$ for any $i \in \{1, 2, \cdots, |U|\}$, and there exists $j \in \{1, 2, \cdots, |U|\}$ such that $S_P(u_j) \subset S_Q(u_j)$.

An incomplete information system $IIS = (U, C \cup D)$ is called an incomplete decision system (IDS) if condition attributes and decision attributes are distinguished, where C is the condition attribute set and D is the decision attribute set with $C \cap D = \emptyset$. Thus, an incomplete decision system is a special case of an incomplete information system, which is generally expressed as $IDS = (U, C, D)$ in this paper. Also, a complete decision system (CDS) is a special case of a complete information system.

3 Information entropy-based uncertainty measures in rough set theory

As a measure of knowledge granularity, information entropy-based measures in incomplete information systems can reflect this difference in knowledge expression. Based on this thought, information entropy-based uncertainty measures for both complete and incomplete information/decision systems are presented by extending the definitions of the measures in complete information systems. Some important properties and propositions of these measures are investigated and relationships among them are discussed as well.

3.1 Information entropy and conditional information entropy in rough set theory

Let CIS be a complete information system and $U/A = \{X_1, X_2, \cdots, X_m\}$. The information entropy of knowledge A [29] is denoted by

$$E(A) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{|X_i^c|}{|U|} = \sum_{i=1}^{m} \frac{|X_i|}{|U|}(1 - \frac{|X_i|}{|U|}), \quad (1)$$

where X_i^c is the complement of X_i, i.e., $X_i^c = U - X_i$, $\frac{|X_i|}{|U|}$ represents the probability of equivalence class X_i within the universe U, and $\frac{|X_i^c|}{|U|}$ represents the probability of the complement set of X_i within the universe U.

Property 3.1. Let CIS be a complete information system and $P, Q \subseteq A$. If $U/P \preceq U/Q$, then $E(Q) \leq E(P)$.

Proof. Since $U/P \preceq U/Q$, it follows that $U/P \prec U/Q$ and $U/P = U/Q$. If $U/P \prec U/Q$, from Theorem 12 in [29], one has that $E(Q) < E(P)$. If $U/P = U/Q$, it is obvious that $E(Q) = E(P)$. Hence, if $U/P \preceq U/Q$, then $E(Q) \leq E(P)$ holds. This completes the proof.

Let IIS be an incomplete information system with $P \subseteq A$, and $U/SIM(P) = \{S_P(u_1), S_P(u_2), \cdots, S_P(u_{|U|})\}$. The information entropy of knowledge P [29] is defined as

$$IE(P) = \sum_{i=1}^{|U|} \frac{1}{|U|}(1 - \frac{|S_P(u_i)|}{|U|}). \quad (2)$$

Let IIS be an incomplete information system with $P, Q \subseteq A$. If there exists a one-to-one, onto function h: $U/SIM(P) \to U/SIM(Q)$ such that $|h(S_P(u_i))| = |S_P(u_i)|$ for any $i \in \{1, 2, \cdots, |U|\}$, then $IE(P) = IE(Q)$. It can be concluded that the above information entropy of knowledge is invariant with respect to different sets of tolerance classes of U that are size-isomorphic.

Property 3.2. Let IIS be an incomplete information system and $P,Q \subseteq A$. If $P \prec Q$, then $IE(Q) < IE(P)$.

Proof. Since $P \prec Q$, one has that $S_P(u_i) \subseteq S_Q(u_i)$, i.e., $|S_P(u_i)| \leq |S_Q(u_i)|$ for any $u_i \in U$, $S_P(u_i) \in U/SIM(P)$ and $S_Q(u_i) \in U/SIM(Q)$, and there exists $j \in \{1,2,\cdots,|U|\}$ such that $S_P(u_j) \subset S_Q(u_j)$, i.e., $|S_P(u_j)| < |S_Q(u_j)|$. Hence, we have that

$$\begin{aligned}
&|S_P(u_i)| \leq |S_Q(u_i)|, \forall u_i \in U \\
&\Rightarrow \sum_{i=1,i\neq j}^{|U|} \frac{1}{|U|}(1-\frac{|S_Q(u_i)|}{|U|}) \leq \sum_{i=1,i\neq j}^{|U|} \frac{1}{|U|}(1-\frac{|S_P(u_i)|}{|U|}) \\
&\Rightarrow \sum_{i=1,i\neq j}^{|U|} \frac{1}{|U|}(1-\frac{|S_Q(u_i)|}{|U|}) + \frac{1}{|U|}(1-\frac{|S_Q(u_j)|}{|U|}) \\
&< \sum_{i=1,i\neq j}^{|U|} \frac{1}{|U|}(1-\frac{|S_P(u_i)|}{|U|}) + \frac{1}{|U|}(1-\frac{|S_P(u_j)|}{|U|}) \\
&\Rightarrow \sum_{i=1}^{|U|} \frac{1}{|U|}(1-\frac{|S_Q(u_i)|}{|U|}) < \sum_{i=1}^{|U|} \frac{1}{|U|}(1-\frac{|S_P(u_i)|}{|U|}),
\end{aligned} \tag{3}$$

i.e., $IE(Q) < IE(P)$. This completes the proof.

Proposition 3.1. Let CIS be a complete information system with $P \subseteq A$. Information entropy of knowledge P degenerates into

$$IE(P) = \sum_{i=1}^{m} \frac{|X_i|}{|U|}(1-\frac{|X_i|}{|U|}). \tag{4}$$

Proof. Suppose that $U/P = \{X_1,X_2,\cdots,X_m\}$, $U/SIM(P) = \{S_P(u_1),S_P(u_2),\cdots,S_P(u_{|U|})\}$, and $X_i = \{u_{i1},u_{i2},\cdots,u_{is_i}\}$ $(i \in \{1,2,\cdots,m\})$, where $|X_i| = s_i$, and $\sum_{i=1}^{m} s_i = |U|$, and then the relationships among the elements in $U/SIM(P)$ and the elements in U/P are as follows: $X_i = S_P(u_{i1}) = S_P(u_{i2}) = \cdots = S_P(u_{is_i})$, i.e., $|X_i| = |S_P(u_{i1})| = |S_P(u_{i2})| = \cdots = |S_P(u_{is_i})|$. Thus, one has that $\frac{|X_i|}{|U|}(1-\frac{|X_i|}{|U|}) = \frac{1}{|U|}(1-\frac{|S_P(u_{i1})|}{|U|}) + \frac{1}{|U|}(1-\frac{|S_P(u_{i2})|}{|U|}) + \cdots + \frac{1}{|U|}(1-\frac{|S_P(u_{is_i})|}{|U|})$. Therefore, it can be easily obtained that

$$\begin{aligned}
IE(P) &= \sum_{i=1}^{m}(\frac{1}{|U|}(1-\frac{|S_P(u_{i1})|}{|U|}) + \frac{1}{|U|}(1-\frac{|S_P(u_{i2})|}{|U|}) \\
&\quad + \cdots + \frac{1}{|U|}(1-\frac{|S_P(u_{is_i})|}{|U|})) \\
&= \sum_{i=1}^{m} \frac{|X_i|}{|U|}(1-\frac{|X_i|}{|U|}).
\end{aligned} \tag{5}$$

Proposition 3.1 shows that the information entropy in complete information systems is a special form of the information entropy in incomplete information systems.

Corollary 3.1. Let CIS be a complete information system with $P \subseteq A$. Then $IE(P) = E(P)$.

In the following, we investigate the information entropy of new knowledge composed of two given knowledge with the same universe. To do it in a much clearer way, we introduce the following lemmas.

Lemma 3.1. Let CIS be a complete information system and $P,Q \subseteq A$. Then $U/P \cap U/Q = U/(P\cup Q) = U/(Q\cup P)$.

Proof. It can be achieved by Theorem 3.1 in [40].

Lemma 3.2. Let IIS be an incomplete information system and $P,Q \subseteq A$. Then the following properties hold
(1) $SIM(P) \cap SIM(Q) = SIM(P \cup Q)$;
(2) $S_P(u) \cap S_Q(u) = S_{P\cup Q}(u)$ for any $u \in U$;
(3) $U/SIM(P) \cap U/SIM(Q) = U/SIM(P \cup Q)$;
(4) $\bigcup_{i=1}^{i=|U|} \bigcup_{j=1}^{j=|U|} \{S_P(u_i) \cap S_Q(u_j)\} = \bigcup_{i=1}^{i=|U|} \{S_P(u_i) \cap S_Q(u_i)\}$.

Proof. It can be achieved by Lemma 1 and Proposition 5 in [36].

Definition 3.1. Let CIS be a complete information system and $P,Q \subseteq A$, $U/P = \{X_1,X_2,\cdots,X_m\}$, $U/Q = \{Y_1,Y_2,\cdots,Y_n\}$. Joint information entropy of P and Q is defined as

$$E(P\cup Q) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}(1-\frac{|Y_j \cap X_i|}{|U|}). \tag{6}$$

Definition 3.2. Let IIS be an incomplete information system and $P,Q \subseteq A$, $U/SIM(P) = \{S_P(u_1),S_P(u_2),\cdots,S_P(u_{|U|})\}$, $U/SIM(Q) = \{S_Q(u_1),S_Q(u_2),\cdots,S_Q(u_{|U|})\}$. Joint information entropy of P and Q is defined as

$$IE(P\cup Q) = \sum_{i=1}^{|U|} \frac{1}{|U|}(1-\frac{|S_P(u_i) \cap S_Q(u_i)|}{|U|}). \tag{7}$$

Proposition 3.2. Let CIS be a complete information system and $P,Q \subseteq A$. The joint information entropy of P and Q degenerates into

$$IE(P\cup Q) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}(1-\frac{|Y_j \cap X_i|}{|U|}). \tag{8}$$

Proof. The proof is similar to that of Proposition 3.1.

Corollary 3.2. Let CIS be a complete information system and $P,Q \subseteq A$. Then $IE(P\cup Q) = E(P\cup Q)$.

Definition 3.3. Let CDS be a complete decision system with $P \subseteq C$, $U/P = \{X_1,X_2,\cdots,X_m\}$, and $U/D = \{Y_1,Y_2,\cdots,Y_n\}$. Conditional information entropy of P relative to D is defined as

$$E(D|P) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c - X_i^c|}{|U|^2}. \tag{9}$$

Property 3.3. Let CDS be a complete decision system with $P,Q \subseteq C$. If $U/P \prec U/Q$, then $E(D|P) \leq E(D|Q)$.

Proof. Suppose that $U/P = \{X_1,X_2,\cdots,X_m\}$, $U/Q = \{Z_1,Z_2,\cdots,Z_l\}$, and $U/D = \{Y_1,Y_2,\cdots,Y_n\}$. Since U/P

$\prec U/Q$, it follows that $m > l$, and then there exists a partition $\{I_1, I_2, \cdots, I_l\}$ of $\{1, 2, \cdots, m\}$ such that $Z_i = \cup\{X_k | k \in I_i, i = 1, 2, \cdots, l\}$. Hence, we can obtain that

$$\begin{aligned}
E(D|Q) &= \sum_{i=1}^{l}\sum_{j=1}^{n}\frac{|Y_j \cap Z_i||Z_i - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{l}\sum_{j=1}^{n}\frac{|Y_j \cap \bigcup_{k\in I_i} X_k||\bigcup_{k\in I_i} X_k - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{l}\sum_{j=1}^{n}\frac{(\sum_{k\in I_i}|Y_j \cap X_k|)|\bigcup_{k\in I_i} X_k - Y_j|}{|U|^2} \\
&\geq \sum_{k=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_k||X_k - Y_j|}{|U|^2} \\
&= \sum_{k=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_k||Y_j^c - X_k^c|}{|U|^2} = E(D|P), \qquad (10)
\end{aligned}$$

i.e., $E(D|P) \leq E(D|Q)$. This completes the proof.

Proposition 3.3. Let *CDS* be a complete decision system and $P \subseteq C$. $E(D|P) = 0$ if and only if $U/P \preceq U/D$.

Proof. $\Rightarrow$ Suppose that $E(D|P) = 0$, we need to prove $U/P \preceq U/D$. If $U/P \preceq U/D$ does not hold, then for any $Y_j \in U/D$, there exists some $X_i \in U/P$ such that $X_i \subseteq Y_j$ does not hold. Let $X_k \in U/P$, $Y_s \in U/D$, then $X_k \cap Y_s \neq \emptyset$, $X_k \cap Y_s \neq X_k$, and $1 \leq |X_k \cap Y_s| < |X_k|$. Hence, we have that

$$\begin{aligned}
&E(D|P) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - X_i \cap Y_j|}{|U|^2} \\
&= \sum_{i=1, i\neq k}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - X_i \cap Y_j|}{|U|^2} + \\
&\quad \sum_{j=1, j\neq s}^{n}\frac{|Y_j \cap X_k||X_k - X_k \cap Y_j|}{|U|^2} + \frac{|Y_s \cap X_k||X_k - X_k \cap Y_s|}{|U|^2} \\
&\geq \frac{|Y_s \cap X_k||X_k - X_k \cap Y_s|}{|U|^2} > 0, \qquad (11)
\end{aligned}$$

i.e., $E(D|P) > 0$. This yields a contradiction. Thus, $U/P \preceq U/D$ holds.

$\Leftarrow$ Suppose $U/P \preceq U/D$, then, for any $X_i \in U/P$, there exists some $Y_j \in U/D$ such that $X_i \subseteq Y_j$. It follows that $X_i \cap Y_j \neq \emptyset$ or $X_i \subseteq Y_j$ for any $X_i \in U/P$ and $Y_j \in U/D$. Then, $X_i \cap Y_j = X_i$, i.e., $X_i - X_i \cap Y_j = \emptyset$. Therefore, we can obtain that

$$\begin{aligned}
E(D|P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - X_i \cap Y_j|}{|U|^2} = 0. \qquad (12)
\end{aligned}$$

Proposition 3.3 illustrates that in the same universe, a knowledge cannot provide the system with any additional uncertainty (classification information) if it is coarser than the original knowledge in complete decision systems. Here, the following propositions will establish the relationships among the information entropy, the joint information entropy, and the conditional information entropy in a complete decision system.

Proposition 3.4. Let *CDS* be a complete decision system and $P \subseteq C$. Then $E(D|P) = E(P \cup D) - E(P)$.

Proof. It follows from Definitions 3.3 and 3.1 that

$$\begin{aligned}
&E(D|P) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||X_i - Y_j \cap X_i|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||(U - Y_j \cap X_i) - (U - X_i)|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||U - Y_j \cap X_i|}{|U|^2} - \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i||U - X_i|}{|U|^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}\frac{(|U| - |Y_j \cap X_i|)}{|U|} - \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}\frac{(|U| - |X_i|)}{|U|} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1 - \frac{|Y_j \cap X_i|}{|U|}) - \\
&\quad \sum_{i=1}^{m}(\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|})(1 - \frac{|X_i|}{|U|}) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1 - \frac{|Y_j \cap X_i|}{|U|}) - \sum_{i=1}^{m}\frac{|X_i|}{|U|}(1 - \frac{|X_i|}{|U|}) \\
&= E(P \cup D) - E(P). \qquad (13)
\end{aligned}$$

To further reveal some relationships between the condition attributes and the decision attribute in an incomplete decision system, we present following definitions and relative properties.

Definition 3.4. Let *IDS* be an incomplete decision system, $P \subseteq C$, $U/SIM(P) = \{S_P(u_1), S_P(u_2), \cdots, S_P(u_{|U|})\}$, and $U/SIM(D) = \{S_D(u_1), S_D(u_2), \cdots, S_D(u_{|U|})\}$. Conditional information entropy of P relative to D is defined as

$$IE(D|P) = \sum_{i=1}^{|U|}\frac{|S_P(u_i)| - |S_P(u_i) \cap S_D(u_i)|}{|U|^2}. \qquad (14)$$

Property 3.4. Let *IDS* be an incomplete decision system with $P, Q \subseteq C$. If $P \preceq Q$, $IE(D|P) \leq IE(D|Q)$.

Proof. Since $P \preceq Q$, it follows that $S_P(u_i) \subseteq S_Q(u_i)$, $|S_P(u_i)| \leq |S_Q(u_i)|$, and then $S_P(u_i) \cap S_D(u_i) \subseteq S_Q(u_i) \cap$

$S_D(u_i)$ for any $u_i \in U$, and $S_D(u_i) \in U/SIM(D)$. From Definition 3.4, we have that

$$\begin{aligned}
&IE(D|Q) - IE(D|P) \\
&= \sum_{i=1}^{|U|}\Big(\frac{|S_Q(u_i)| - |S_P(u_i)|}{|U|^2} - \\
&\qquad \frac{|S_Q(u_i) \cap S_D(u_i)| - |S_P(u_i) \cap S_D(u_i)|}{|U|^2}\Big) \\
&= \sum_{i=1}^{|U|} \frac{|S_Q(u_i)| - |S_P(u_i)| - |S_D(u_i) \cap (S_Q(u_i) - S_P(u_i))|}{|U|^2} \\
&\geq 0,
\end{aligned} \tag{15}$$

i.e., $IE(D|P) \leq IE(D|Q)$. Thus, it is obvious that $IE(D|P) = IE(D|Q)$ if and only if $\{S_Q(u_i) - S_P(u_i)\} \subseteq S_D(u_i)$ for any $u_i \in U$. This completes the proof.

Proposition 3.5. Let *IDS* be an incomplete decision system with $P \subseteq C$. Then $IE(D|P) = IE(P \cup D) - IE(P)$.

Proof. It follows immediately from Definition 3.4 that

$$\begin{aligned}
&IE(D|P) \\
&= \sum_{i=1}^{|U|} \frac{|S_P(u_i)|}{|U|^2} - \sum_{i=1}^{|U|} \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|^2} + \sum_{i=1}^{|U|} \frac{1}{|U|} - \sum_{i=1}^{|U|} \frac{1}{|U|} \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\Big(1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|}\Big) - \sum_{i=1}^{|U|} \frac{1}{|U|}\Big(1 - \frac{|S_P(u_i)|}{|U|}\Big) \\
&= IE(P \cup D) - IE(P).
\end{aligned} \tag{16}$$

Proposition 3.6. Let *IDS* be an incomplete decision system with $P \subseteq C$. $IE(D|P) = 0$ if and only if $P \preceq D$.

Proof. $\Rightarrow$ Suppose $IE(D|P) = 0$, we need to prove $P \preceq D$. If $P \preceq D$ does not hold, then there exists some $u_j \in U$ such that $S_P(u_j) \subseteq S_D(u_j)$ does not hold, i.e., $|S_P(u_j) \cap S_D(u_j)| < |S_P(u_j)|$. Hence, we can obtain that

$$\begin{aligned}
IE(D|P) &= \sum_{i=1, i \neq j}^{|U|} \frac{|S_P(u_i)| - |S_P(u_i) \cap S_D(u_i)|}{|U|^2} \\
&\quad + \frac{|S_P(u_j)| - |S_P(u_j) \cap S_D(u_j)|}{|U|^2} \\
&\geq \frac{|S_P(u_j)| - |S_P(u_j) \cap S_D(u_j)|}{|U|^2} \\
&> \frac{|S_P(u_j)| - |S_P(u_j)|}{|U|^2} = 0,
\end{aligned} \tag{17}$$

i.e., $IE(D|P) > 0$. This yields a contradiction. Thus, $P \preceq D$ holds.

$\Leftarrow$ Suppose $P \preceq D$, then, for any $u_i \in U$, it follows that $S_P(u_i) \subseteq S_D(u_i)$, i.e., $S_P(u_i) \cap S_D(u_i) = S_P(u_i)$. Therefore, we have that $IE(D|P) = \sum_{i=1}^{|U|} \frac{|S_P(u_i)| - |S_P(u_i)|}{|U|^2} = 0$. This completes the proof.

Proposition 3.6 states that any knowledge in the same universe cannot provide the system with any additional uncertainty if it is coarser than the original knowledge in incomplete decision systems.

Proposition 3.7. Let *CDS* be a complete decision system with $P \subseteq C$. The conditional information entropy of P relative to D degenerates into

$$IE(D|P) = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c - X_i^c|}{|U|^2}. \tag{18}$$

Proof. Let $X_i = \{u_{i1}, u_{i2}, ..., u_{is_i}\} \in U/P$ $(i \in \{1, 2, \cdots, m\})$, where $|X_i| = s_i$, $\sum_{i=1}^{m} s_i = |U|$, and $Y_j = \{u_{j1}, u_{j2}, \cdots, u_{jt_j}\} \in U/D$ $(j \in \{1, 2, \cdots, n\})$, where $|Y_j| = t_j$, $\sum_{j=1}^{n} t_j = |U|$. Similar to Proposition 3.1, the relationships among the elements in $U/SIM(P)$ and the elements in U/P are as follows: $X_i = S_P(u_{i1}) = S_P(u_{i2}) = \cdots = S_P(u_{is_i})$, i.e., $|X_i| = |S_P(u_{i1})| = |S_P(u_{i2})| = \cdots = |S_P(u_{is_i})|$, and the relationships among the elements in $U/SIM(D)$ and the elements in U/D are as follows: $Y_j = S_D(u_{j1}) = S_D(u_{j2}) = \cdots = S_D(u_{jt_j})$, i.e., $|Y_j| = |S_D(u_{j1})| = |S_D(u_{j2})| = \cdots = |S_D(u_{jt_j})|$. Therefore, we can obtain that

$$\begin{aligned}
IE(D|P) &= \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{u_k \in Y_j \cap X_i} \frac{|S_P(u_k)| - |S_P(u_k) \cap S_D(u_k)|}{|U|^2} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} |Y_j \cap X_i| \frac{|X_i| - |X_i \cap Y_j|}{|U|^2} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{|Y_j \cap X_i||X_i - Y_j|}{|U|^2} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{|Y_j \cap X_i||Y_j^c - X_i^c|}{|U|^2}.
\end{aligned} \tag{19}$$

Corollary 3.3. Let *CDS* be a complete decision system with $P \subseteq C$. Then $IE(D|P) = E(D|P)$.

Proposition 3.7 and Corollary 3.3 state that the conditional information entropy in complete decision systems is a special instance of the conditional information entropy in incomplete decision systems. In other words, the conditional information entropy under the equivalence relation is the extended formulation of the conditional information entropy under the tolerance relation. This means that the definition of conditional information entropy in a complete decision system is a consistent extension to that of conditional information entropy in an incomplete decision system. Hence, it follows that the conditional information entropy in incomplete decision systems is equivalent to the conditional information entropy in complete decision systems, and the conditional information entropy proposed above in incomplete decision systems is suitable for measuring the uncertainty in both incomplete and complete decision systems.

3.2 Comparison analysis with several representative uncertainty measures

Let *CDS* be a complete decision system and $P \subseteq C$, $U/P = \{X_1, X_2, \cdots, X_m\}$, $U/D = \{Y_1, Y_2, \cdots, Y_n\}$. Conditional information entropy of P relative to D [41] is denoted by

$$H_1(D|P) = -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|} \log_2 \frac{|Y_j \cap X_i|}{|X_i|}. \quad (20)$$

Let *IIS* be an incomplete information system with $P \subseteq A$. Information entropy of knowledge P [29] is denoted by

$$H_1'(P) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|S_P(u_i)|}{|U|}. \quad (21)$$

Definition 3.5. Let *IDS* be an incomplete decision system with $P \subseteq C$. Conditional information entropy of P relative to D is defined as

$$H_1'(D|P) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}. \quad (22)$$

Property 3.5. Let *IDS* be an incomplete decision system and $P \subseteq C$. Then $IE(D|P) \leq H_1'(D|P)$.

Proof. It follows immediately from Definition 3.4 that

$$IE(D|P) = \sum_{i=1}^{|U|} \frac{1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}}{\frac{|U|^2}{|S_P(u_i)|}} = \sum_{i=1}^{|U|} \frac{1}{|U|} \left(1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}\right) \frac{|S_P(u_i)|}{|U|}. \quad (23)$$

Assume that a function $f_{u_i} = \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}$ for any $u_i \in U$, then we have that $IE(D|P) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|S_P(u_i)|}{|U|} (1 - f_{u_i})$. It follows similarly from Definition 3.5 that $H_1'(D|P) = -\frac{1}{|U|} \sum_{i=1}^{|U|} \log_2 f_{u_i}$. It is obvious that because $0 \leq f_{u_i} \leq 1$, one has that $1 - f_{u_i} \leq -\log_2 f_{u_i}$ and $\frac{|S_P(u_i)|}{|U|}(1 - f_{u_i}) \leq -\log_2 f_{u_i}$ for any $u_i \in U$. Thus, it can be obtained that $\sum_{i=1}^{|U|} \frac{|S_P(u_i)|}{|U|} (1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}) \leq -\sum_{i=1}^{|U|} \log_2 \frac{|S_P(u_i) \cap S_D(u_i)|}{|S_P(u_i)|}$, i.e., $IE(D|P) \leq H_1'(D|P)$. This completes the proof.

Proposition 3.8. Let *IDS* be an incomplete decision system with $P \subseteq C$. Then $H_1'(D|P) = H_1'(P \cup D) - H_1'(P)$.

Proof. It follows immediately from Definition 5 that

$$\begin{aligned} &H_1'(D|P) \\ &= -\sum_{i=1}^{|U|} \frac{1}{|U|} \left(\log_2 \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|} - \log_2 \frac{|S_P(u_i)|}{|U|}\right) \\ &= -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|} + \sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|S_P(u_i)|}{|U|} \\ &= H_1'(P \cup D) - H_1'(P). \end{aligned} \quad (24)$$

Proposition 3.9. Let *IDS* be an incomplete decision system, $P \subset Q \subseteq C$. $H_1'(D|Q) < H_1'(D|P)$ does not hold.

In the following, the performance of Proposition 3.9 in an incomplete decision system is shown through an illustrative example.

Example 3.1. Consider an incomplete decision system about several cars shown in Table 1, where $C = \{Price, Mileage, Size, Max-Speed\} = \{P, M, S, X\}$ and $D = \{Acceleration\}$.

Table 1: An incomplete decision system about cars

car	P	M	S	X	D
1	High	Low	Full	Low	Good
2	Low	*	Full	Low	Good
3	*	*	Compact	Low	Poor
4	High	*	Full	High	Good
5	*	*	Full	High	Excellent
6	Low	High	Full	*	Good

Assume that $U/SIM(C) = \{\{1\},\{2,6\},\{3\},\{4,5\},\{4,5,6\},\{2,5,6\}\}$, $U/SIM(\{S,X\}) = \{\{1,2,6\},\{1,2,6\},\{3\},\{4,5,6\},\{4,5,6\},\{1,2,4,5,6\}\}$, and $U/SIM(D) = \{\{1,2,4,6\},\{1,2,4,6\},\{3\},\{1,2,4,6\},\{5\},\{1,2,4,6\}\}$. Then, it is easily computed that $H_1'(D|C) = \frac{\log_2 3}{3}$ and $H_1'(D|\{S,X\}) = \frac{\log_2 3}{3} - \frac{\log_2 5}{6} - \frac{1}{2}$. We have $H_1'(D|C) - H_1'(D|\{S,X\}) = \frac{\log_2 5}{6} + \frac{1}{2}$, i.e., $H_1'(D|C) > H_1'(D|\{S,X\})$. As a result, this shows that $\{S,X\} \subset C = \{P,M,S,X\}$, then one has that $H_1'(D|\{P,M,S,X\}) > H_1'(D|\{S,X\})$.

Proposition 3.10. Let *CDS* be a complete decision system with $P \subseteq C$. The conditional information entropy of P relative to D degenerates into

$$H_1'(D|P) = -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|} \log_2 \frac{|Y_j \cap X_i|}{|X_i|}. \quad (25)$$

Proof. Similar to Proposition 3.7, we can obtain that

$$\begin{aligned} H_1'(D|P) &= -\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{u_k \in Y_j \cap X_i} \frac{1}{|U|} \log_2 \frac{|S_P(u_k) \cap S_D(u_k)|}{|S_P(u_k)|} \\ &= -\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \log_2 \frac{|Y_j \cap X_i|}{|X_i|} \\ &= -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|} \log_2 \frac{|Y_j \cap X_i|}{|X_i|}. \end{aligned} \quad (26)$$

Corollary 3.4. Let *CDS* be a complete decision system and $P \subseteq C$. Then $H_1'(D|P) = H_1(D|P)$.

Proposition 3.11. Let *CDS* be a complete decision system and $P \subseteq C$. Then $E(D|P) \leq H_1(D|P)$.

Proof. It can be achieved by Corollaries 3.3 and 3.4, and Property 3.5.

Let CDS be a complete decision system with $P \subseteq C$. Conditional information entropy of P relative to D [41] is denoted by

$$H_2(D|P) = -\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1+\log_2\frac{|Y_j \cap X_i|}{|X_i|} - \frac{|Y_j \cap X_i|}{|U|}). \quad (27)$$

Definition 3.6. Let IDS be an incomplete decision system with $P \subseteq C$. Conditional information entropy of P relative to D is defined as

$$H_2'(D|P) = -\sum_{i=1}^{|U|}\frac{1}{|U|}(1+\log_2\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|} - \frac{|S_P(u_i)\cap S_D(u_i)|}{|U|}). \quad (28)$$

Proposition 3.12. Let IDS be an incomplete decision system with $P \subseteq C$. $H_2'(D|P) = H_1'(D|P) - IE(D\cup P)$.

Proof. It follows from Definitions 3.6, 3.5 and 3.2 that

$$\begin{aligned} H_2'(D|P) &= -\sum_{i=1}^{|U|}\frac{1}{|U|}\log_2\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|} \\ &\quad -\sum_{i=1}^{|U|}\frac{1}{|U|}(1-\frac{|S_P(u_i)\cap S_D(u_i)|}{|U|}) \\ &= H_1'(D|P) - IE(D\cup P). \end{aligned} \quad (29)$$

Corollary 3.5. Let IDS be an incomplete decision system and $P \subseteq C$. Then $H_2'(D|P) \leq H_1'(D|P)$.

Proof. It can be achieved by Proposition 3.12 and $IE(D\cup P) \geq 0$.

Proposition 3.13. Let IDS be an incomplete decision system, $P \subset Q \subseteq C$. $H_2'(D|Q) < H_2'(D|P)$ does not hold.

The following example illustrates the performance of Proposition 3.13 in an incomplete decision system.

Example 3.2. (Continued from Example 3.1)

By computing, we have that $H_2'(D|C) = \frac{\log_2 3}{3} - \frac{7}{9}$ and $H_2'(D|\{S,X\}) = \frac{\log_2 3}{3} - \frac{\log_2 5}{6} - \frac{10}{9}$, then $H_2'(D|C) - H_2'(D|\{S,X\}) = \frac{\log_2 5}{6} + \frac{1}{3}$, i.e., $H_2'(D|C) > H_2'(D|\{S,X\})$. As a result, this shows that $\{S,X\} \subset C = \{P,M,S,X\}$, then one has that $H_2'(D|\{P,M,S,X\}) > H_2'(D|\{S,X\})$.

Proposition 3.14. Let CDS be a complete decision system with $P \subseteq C$. The conditional information entropy of P relative to D degenerates into

$$H_2'(D|P) = -\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1+\log_2\frac{|Y_j \cap X_i|}{|X_i|} - \frac{|Y_j \cap X_i|}{|U|}). \quad (30)$$

Proof. Suppose that CDS is a complete decision system, and then it follows from Proposition 3.12, and Corollaries 3.4 and 3.2 that

$$\begin{aligned} & H_2'(D|P) \\ &= H_1'(D|P) - IE(D\cup P) \\ &= H_1(D|P) - E(D\cup P) \\ &= -\sum_{i=1}^{m}\frac{|X_i|}{|U|}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|X_i|}\log_2\frac{|Y_j \cap X_i|}{|X_i|} \\ &\quad -\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1-\frac{|Y_j \cap X_i|}{|U|}) \\ &= -\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|U|}(1+\log_2\frac{|Y_j \cap X_i|}{|X_i|} - \frac{|Y_j \cap X_i|}{|U|}). \end{aligned} \quad (31)$$

Corollary 3.6. Let CDS be a complete decision system with $P \subseteq C$. Then $H_2'(D|P) = H_2(D|P)$.

Proposition 3.15. Let CDS be a complete decision system with $P \subseteq C$. Then $H_2(D|P) \leq H_1(D|P)$.

Proof. It can be achieved by Corollaries 3.4, 3.5 and 3.6.

Let CDS be a complete decision system with $P \subseteq C$. Conditional information entropy of P relative to D [43] is denoted by

$$I_1(D|P) = \sum_{i=1}^{m}\frac{|X_i|}{|U|}\sum_{j=1}^{n}\frac{|Y_j \cap X_i|}{|X_i|}(1-\frac{|Y_j \cap X_i|}{|X_i|}). \quad (32)$$

Definition 3.7. Let IDS be an incomplete decision system with $P \subseteq C$. Conditional information entropy of P relative to D is defined as

$$I_1'(D|P) = \sum_{i=1}^{|U|}\frac{1}{|U|}(1-\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|}). \quad (33)$$

Property 3.6. Let IDS be an incomplete decision system and $P \subseteq C$. Then $IE(D|P) \leq I_1'(D|P) \leq H_1'(D|P)$.

Proof. It follows immediately from Property 3.5 that $IE(D|P) = \sum_{i=1}^{|U|}\frac{1}{|U|}(1-\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|})\frac{|S_P(u_i)|}{|U|}$. Then, from Definition 3.7, it is clear that $IE(D|P) = I_1'(D|P)\frac{|S_P(u_i)|}{|U|}$. Since $0 \leq \frac{|S_P(u_i)|}{|U|} \leq 1$, one has that $IE(D|P) \leq I_1'(D|P)$. Similar to Property 3.5, let a function $f_{u_i} = \frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|}$ for any $u_i \in U$. Thus, it is obvious that $0 \leq f_{u_i} \leq 1$. We have that $1 - f_{u_i} \leq -\log_2 f_{u_i}$. It can be obtained that $\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|}) \leq -\sum_{i=1}^{|U|}\log_2\frac{|S_P(u_i)\cap S_D(u_i)|}{|S_P(u_i)|}$, i.e., $I_1'(D|P) \leq H_1'(D|P)$. Hence, $IE(D|P) \leq I_1'(D|P) \leq H_1'(D|P)$ holds. This completes the proof.

Proposition 3.16. Let IDS be an incomplete decision system, $P \subset Q \subseteq C$. $I_1'(D|Q) < I_1'(D|P)$ does not hold.

In what follows, we show the performance of Proposition 3.16 in an incomplete decision system through an illustrative example.

Example 3.3. (Continued from Example 3.1)

By computing, we have that $I_1'(D|C) = \frac{1}{4}$ and $I_1'(D|\{S,X\}) = \frac{1}{5}$, i.e., $I_1'(D|C) > I_1'(D|\{S,X\})$. As a result, it shows that $\{S,X\} \subset C = \{P,M,S,X\}$, then one has that $I_1'(D|\{P,M,S,X\}) > I_1'(D|\{S,X\})$.

Proposition 3.17. Let $CDS = (U,C,D)$ be a complete decision system with $P \subseteq C$. The conditional information entropy of P relative to D degenerates into

$$I_1'(D|P) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}). \tag{34}$$

Proof. Similar to Proposition 3.7, it can be obtained that

$$\begin{aligned} I_1'(D|P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{u_k \in Y_j \cap X_i} \frac{1}{|U|}(1 - \frac{|S_P(u_k) \cap S_D(u_k)|}{|S_P(u_k)|}) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}) \\ &= \sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}). \end{aligned} \tag{35}$$

Corollary 3.7. Let CDS be a complete decision system and $P \subseteq C$. Then $I_1'(D|P) = I_1(D|P)$.

Proposition 3.18. Let CDS be a complete decision system and $P \subseteq C$. Then $E(D|P) \leq I_1(D|P) \leq H_1(D|P)$.

Proof. It can be achieved by Corollaries 3.3, 3.4 and 3.7, and Property 3.6.

Let CDS be a complete decision system with $P \subseteq C$. Conditional information entropy of P relative to D [44] is denoted by

$$I_2(D|P) = \sum_{i=1}^{m} \frac{|X_i|^2}{|U|^2} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}). \tag{36}$$

Definition 3.8. Let IDS be an incomplete decision system with $P \subseteq C$. Conditional information entropy of P relative to D is defined as

$$I_2'(D|P) = \sum_{i=1}^{|U|} \frac{|S_P(u_i)| - |S_P(u_i) \cap S_D(u_i)|}{|U|^2}. \tag{37}$$

Property 3.7. Let IDS be an incomplete decision system and $P \subseteq C$. Then $I_2'(D|P) = IE(D|P)$.

Proof. It can be achieved by Definitions 3.4 and 3.8.

Proposition 3.19. Let IDS be an incomplete decision system and $P \subseteq C$. Then $I_2'(D|P) \leq I_1'(D|P) \leq H_1'(D|P)$.

Proof. It can be achieved by Properties 3.6 and 3.7.

Proposition 3.20. Let IDS be an incomplete decision system and $P,Q \subseteq C$. If $P \preceq Q$, then $I_2'(D|Q) \leq I_2'(D|P)$.

Proof. It can be achieved by Properties 3.7 and 3.4.

Proposition 3.21. Let CDS be a complete decision system with $P \subseteq C$. The conditional information entropy of P relative to D degenerates into

$$I_2'(D|P) = \sum_{i=1}^{m} \frac{|X_i|^2}{|U|^2} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}). \tag{38}$$

Proof. Similar to Proposition 3.7, it can be obtained that

$$\begin{aligned} I_2'(D|P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{u_k \in Y_j \cap X_i} \frac{|S_P(u_k)| - |S_P(u_k) \cap S_D(u_k)|}{|U|^2} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} |Y_j \cap X_i| \frac{|X_i| - |Y_j \cap X_i|}{|U|^2} \\ &= \sum_{i=1}^{m} \frac{|X_i|^2}{|U|^2} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|} \frac{(|X_i| - |Y_j \cap X_i|)}{|X_i|} \\ &= \sum_{i=1}^{m} \frac{|X_i|^2}{|U|^2} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|X_i|}(1 - \frac{|Y_j \cap X_i|}{|X_i|}). \end{aligned} \tag{39}$$

Corollary 3.8. Let CDS be a complete decision system and $P \subseteq C$. Then $I_2'(D|P) = I_2(D|P)$.

Proposition 3.22. Let CDS be a complete decision system and $P \subseteq C$. Then $I_2(D|P) = E(D|P)$.

Proof. It can be achieved by Corollaries 3.3 and 3.8, and Property 3.7.

Proposition 3.23. Let CDS be a complete decision system and $P \subseteq C$. Then $I_2(D|P) \leq I_1(D|P) \leq H_1(D|P)$.

Proof. It can be achieved by Corollaries 3.4, 3.7 and 3.8, and Proposition 3.19.

4 Mutual information-based uncertainty measures in rough set theory

As we all know, mutual information can be considered as advanced statistics to rank the salient attributes. When applied in attribute reduction, mutual information plays a key role in measuring the relevance and redundancy among attributes [45]. The main advantages of mutual information are its robustness to noise and transformation. In the following, we investigate mutual information in both complete and incomplete information/decision systems, and discuss some relationships between mutual information and information entropy.

Definition 4.1. Let CIS be a complete information system and $P,Q \subseteq A$, $U/P = \{X_1, X_2, \cdots, X_m\}$, $U/Q = \{Y_1, Y_2, \cdots, Y_n\}$. Mutual information between Q and P is defined as

$$E(Q;P) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c \cap X_i^c|}{|U|^2}. \tag{40}$$

Definition 4.2. Let CDS be a complete decision system, $P \subseteq C$. Mutual information between D and P is defined as

$$E(D;P) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c \cap X_i^c|}{|U|^2}. \tag{41}$$

Here, the following propositions will establish the relationships among the information entropy, the conditional information entropy and the mutual information in a complete decision system.

Property 4.1. Let CDS be a complete decision system with $P \subseteq C$. $E(D;P) = E(D) - E(D|P) = E(P) - E(P|D)$.

Proof. It follows from Definitions 4.2 and 3.3 that

$$\begin{aligned}
&E(D)\\
&= \sum_{j=1}^{n} \frac{|Y_j|}{|U|}\frac{|Y_j^c|}{|U|}\\
&= \sum_{j=1}^{n} (\sum_{i=1}^{m} \frac{|Y_j \cap X_i|}{|U|})\frac{|Y_j^c|}{|U|}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}\frac{|Y_j^c|}{|U|}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}\frac{|(Y_j^c \cap X_i^c) \cup (Y_j^c - X_i^c)|}{|U|}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}\frac{|Y_j^c \cap X_i^c|}{|U|} + \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}\frac{|Y_j^c - X_i^c|}{|U|}\\
&= E(D;P) + E(D|P),
\end{aligned} \tag{42}$$

i.e., $E(D;P) = E(D) - E(D|P)$. Similarly, the equation $E(D;P) = E(P) - E(P|D)$ can be proved. This completes the proof.

Corollary 4.1. Let CDS be a complete decision system with $P \subseteq C$. Then $E(D;P) = E(D) + E(P) - E(P \cup D)$.

Proof. It can be achieved by Property 4.1 and Proposition 3.4.

Proposition 4.1. Let CDS be a complete decision system and P, $Q \subseteq C$. If $U/P \prec U/Q$, then $E(D;Q) \leq E(D;P)$.

Proof. It follows from Property 4.1 that $E(D;P) = E(D)$ $-E(D|P)$ and $E(D;Q) = E(D)-$ $E(D|Q)$. Then, from Property 3.3, one has that $-E(D|Q) \leq -E(D|P)$. Thus, it is obvious that $E(D)-$ $E(D|Q) \leq E(D) - E(D|P)$, i.e., $E(D;Q) \leq E(D;P)$. This completes the proof.

Definition 4.3. Let IIS be an incomplete information system and $P, Q \subseteq A$. Mutual information between Q and P is defined as

$$\begin{aligned}
&IE(Q;P)\\
&= \sum_{i=1}^{|U|} \frac{|U| - |S_P(u_i)| - |S_Q(u_i)| + |S_P(u_i) \cap S_Q(u_i)|}{|U|^2}.
\end{aligned} \tag{43}$$

Definition 4.4. Let IDS be an incomplete decision system, $P \subseteq C$. Mutual information between D and P is defined as

$$\begin{aligned}
&IE(D;P)\\
&= \sum_{i=1}^{|U|} \frac{|U| - |S_P(u_i)| - |S_D(u_i)| + |S_P(u_i) \cap S_D(u_i)|}{|U|^2}.
\end{aligned} \tag{44}$$

Proposition 4.2. Let IDS be an incomplete decision system, $P \subseteq C$. $IE(D;P) = IE(P) + IE(D) - IE(P \cup D)$.

Proof. It follows immediately from Definition 4.4 that

$$\begin{aligned}
IE(D;P) &= \sum_{i=1}^{|U|} \frac{1}{|U|}((1 - \frac{|S_P(u_i)|}{|U|}) + (1 - \frac{|S_D(u_i)|}{|U|}) -\\
&\qquad (1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|}))\\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}(1 - \frac{|S_P(u_i)|}{|U|}) + \sum_{i=1}^{|U|} \frac{1}{|U|}(1 - \frac{|S_D(u_i)|}{|U|})\\
&\quad - \sum_{i=1}^{|U|} \frac{1}{|U|}(1 - \frac{|S_P(u_i) \cap S_D(u_i)|}{|U|})\\
&= IE(P) + IE(D) - IE(P \cup D).
\end{aligned} \tag{45}$$

Corollary 4.2. Let IDS be an incomplete decision system, $P \subseteq C$. $IE(D;P) = IE(P) - IE(P|D) = IE(D) - IE(D|P)$.

Proof. From Propositions 3.5 and 4.1, we have that $IE(D;P) = IE(P) + IE(D) - IE(P \cup D) = IE(D)-$ $(IE(P \cup D) - IE(P)) = IE(D) - IE(D|P)$. Similarly, the equation $IE(D;P) = IE(P) - IE(P|D)$ can be proved. This completes the proof.

It should be noted that these equations cannot be satisfied by some existing measures in incomplete decision systems. Furthermore, these relationships will be helpful for understanding the essence of the knowledge content and the uncertainty in both incomplete and complete decision systems.

Proposition 4.3. Let CDS be a complete decision system with $P \subseteq C$. The mutual information between D and P degenerates into

$$IE(D;P) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c \cap X_i^c|}{|U|^2}. \tag{46}$$

Proof. From Corollary 4.2, we have that $IE(D;P) = IE(D) - IE(D|P)$. It follows from Propositions 3.1 and 3.7 that $IE(D) = \sum_{j=1}^{n} \frac{|Y_j|}{|U|}(1 - \frac{|Y_j|}{|U|})$ $= \sum_{j=1}^{n} \frac{|Y_j|}{|U|}\frac{|Y_j^c|}{|U|}$ and $IE(D|P) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i||Y_j^c - X_i^c|}{|U|^2}$. From Corollary 4.2, we have $IE(D;P) = IE(D)-$ $IE(D|P) = \sum_{j=1}^{n} \frac{|Y_j|}{|U|}\frac{|Y_j^c|}{|U|} - \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|}\frac{|Y_j^c - X_i^c|}{|U|}$. Thus, it is obvious from Property 4.1 that $\sum_{j=1}^{n} \frac{|Y_j|}{|U|}\frac{|Y_j^c|}{|U|} =$

$\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \frac{|Y_j^c - X_i^c|}{|U|} + \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \frac{|Y_j^c \cap X_i^c|}{|U|}$, i.e., $\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \frac{|Y_j^c \cap X_i^c|}{|U|} = \sum_{j=1}^{n} \frac{|Y_j|}{|U|} \frac{|Y_j^c|}{|U|} - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \frac{|Y_j^c - X_i^c|}{|U|} = IE(D;P)$. This completes the proof.

Corollary 4.3. Let CDS be a complete decision system with $P \subseteq C$. Then $IE(D;P) = E(D;P)$.

Proposition 4.3 and Corollary 4.3 state that the mutual information in complete decision systems is a special instance of the mutual information in incomplete decision systems. Until now, most of the existing uncertainty measures cannot be used in incomplete decision systems. According to the properties mentioned and the corresponding discussions above, it is known that the conditional information entropy and the mutual information proposed above can well characterize the uncertainty of knowledge in incomplete decision systems. Hence, it can be shown that these measures which are proposed provide important approaches to measuring the uncertainty ability of different knowledge in incomplete information/decision systems. But so far, these uncertainty measures and the relationships among them above have not been reported in incomplete information/ decision systems. In fact, given any binary relation, one can induce a cover of the universe and determine a particular information system. Furthermore, through using the idea of the information theory, we may use the information entropy, the mutual information and their variants to measure the uncertainty of the information systems induced by a given binary relation. In other words, these proposed uncertainty measure approaches can not only characterize the uncertainty of an incomplete information system, but also measure those of some more kinds of information systems.

Remark. Unlike most of the existing measures for the uncertainty in both complete and incomplete information/decision systems, the relationships among these concepts (the information entropy, the conditional information entropy, and the mutual information) can be established, which are formally expressed as $IE(P) = E(P)$ and $IE(P \cup Q) = E(P \cup Q)$ in a CIS with $P,Q \subseteq A$, $E(D|P) = E(P \cup D) - E(P)$, $IE(D|P) = E(D|P)$, $H_1'(D|P) = H_1(D|P)$, $E(D|P) \leq H_1(D|P)$, $H_2'(D|P) = H_2(D|P)$, $H_2(D|P) \leq H_1(D|P)$, $I_1'(D|P) = I_1(D|P)$, $E(D|P) \leq I_1(D|P) \leq H_1(D|P)$, $I_2'(D|P) = I_2(D|P)$, $I_2(D|P) = E(D|P)$, $I_2(D|P) \leq I_1(D|P) \leq H_1(D|P)$, $E(D;P) = E(D) - E(D|P) = E(P) - E(P|D)$, $E(D;P) = E(D) + E(P) - E(P \cup D)$, and $IE(D;P) = E(D;P)$ in a CDS with $P \subseteq C$, and $IE(D|P) = IE(P \cup D) - IE(P)$, $IE(D|P) \leq H_1'(D|P)$, $H_1'(D|P) = H_1'(P \cup D) - H_1'(P)$, $H_2'(D|P) = H_1'(D|P) - IE(D \cup P)$, $H_2'(D|P) \leq H_1'(D|P)$, $IE(D|P) \leq I_1'(D|P) \leq H_1'(D|P)$, $I_2'(D|P) = IE(D|P)$, $I_2'(D|P) \leq I_1'(D|P) \leq H_1'(D|P)$, $IE(D;P) = IE(P) + IE(D) - IE(P \cup D)$, $IE(D;P) = IE(P) - IE(P|D) = IE(D) - IE(D|P)$ in an IDS with $P \subseteq C$. These relationships are very significant for reasonably applying an uncertainty measure in both complete and incomplete information/ decision systems. However, most of the existing entropies and their extensions in incomplete information/decision systems can not establish the relationships above. Therefore, these uncertainty measures mentioned above may be much better for measuring the knowledge content of incomplete information/ decision systems.

5 Conclusions

In many real-world tasks, data available are incomplete, but decisions must be made with the incomplete data for the time being. Most decision systems are incomplete for various subjective and objective reasons. Therefore, it is necessary to develop a theory which can handle incomplete data. In this article, we introduce concepts of information entropy and mutual information-based uncertainty measures, and then some important properties of them are discussed. From these properties, it can be shown that these proposed uncertainty measures can be used to evaluate the uncertainty ability of different knowledge in complete/incomplete decision systems. Furthermore, the relationships among these proposed measures are established, and then compared with several representative uncertainty measures, these proposed uncertainty measures may be much better measures to evaluate the uncertainty of the knowledge content of incomplete information/decision systems. However, this paper is mainly oriented towards theory and methodology instead of specific applications. Thus, these results will have a wide variety of applications in rule evaluation and knowledge discovery in rough set theory. Future work will therefore focus on how to use (some of) these approaches or improved versions of them to solve specific problems, such as measuring the knowledge content, constructing a decision tree, and building a heuristic function in a heuristic reduct algorithm for incomplete data, and acquiring rules or extracting features from engineering data.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Nos. 60873104, 61040037, 61370169), the Key Project of Science and Technology Department of Henan Province (No. 112102210194), the Science and Technology Research Key Project of Educational Department of Henan Province (Nos. 12A520027, 13A520529), the Education Fund for Youth Key Teachers of Henan Normal University, and the Postgraduate Science and Technology Program of Beijing University of Technology (No. ykj-2012-6765).
The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] Z. Pawlak, Rough sets, International Journal of Information and Computer Science, **11**, 341-356 (1982).

[2] L. Sun, N. Zhong and J.C. Xu, Granularity-based user-centric multi-strategies and application in knowledge retrieval, Journal of Computers, **8**, 2101-2109 (2013).

[3] L. Sun and J. C. Xu, A granular computing approach to gene selection, Bio-Medical Materials and Engineering, **24**, 1307-1314 (2014).

[4] S. Y. Li, T. R. Li and D. Liu, Incremental updating approximations in dominance-based rough sets approach under the variation of the attribute set, Knowledge-Based Systems, **40**, 17-26 (2013).

[5] L. Sun, J. C. Xu, C. Wang, T. H. Xu and J. Y. Ren, Granular computing-based granular structure model and its application in knowledge retrieval, Information Technology Journal, **11**, 1714-1721 (2012).

[6] A. Mellit,H. Eleuch, M. Benghanem, C. Elaoun and A. Massi Pavan, Energy Conversion and Management, **51**, 771–782 (2010).

[7] Y. H. Qian and J. Y. Liang, Combination entropy and combination granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, **16**, 179-193 (2008).

[8] J. C. Xu and L. Sun, Knowledge entropy and feature selection in incomplete decision systems, Applied Mathematics and Information Sciences, **6**, 829-837 (2013).

[9] L. Sun, J. C. Xu, S. Q. Li, X. Z. Cao and Y. P. Gao, New approach for feature selection by using information entropy, Journal of Information and Computational Science, **8**, 1525-1532 (2011).

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

[11] L. Sun, J. C. Xu, Z. A. Xue and L. J. Zhang, Rough entropy-based feature selection and its application, Journal of Information and Computational Science, **8**, 1525-1532 (2011).

[12] L. Sun, J. C. Xu, Y. W. Hu and L. N. Du, Granular space-based feature selection and its applications, Journal of Software, **8**, 817-826 (2013).

[13] L. Sun, J. C. Xu, J. Y. Ren, T. H. Xu and Q. Q. Zhang, Granularity partition-based feature selection and its application in decision systems, Journal of Information and Computational Science, **9**, 3487-3500 (2012).

[14] L. Sun, J. C. Xu, Z. A. Xue and J. Y. Ren, Decision degree-based decision tree technology for rule extraction, Journal of Computers, **7**, 1769-1779 (2012).

[15] B. W. Xu, Y. M. Zhou and H. M. Lu, An improved accuracy measure for rough sets, Journal of Computer and System Sciences, **71**, 163-173 (2005).

[16] Y. Y. Yao, Notes on rough set approximations and associated measures, Journal of Zhejiang Ocean University, **29**, 399-410 (2010).

[17] C. E. Shannon, The mathematical theory of communication, The Bell System Technical Journal, **27**, 373-423 (1948).

[18] Claude Gravel, Quantum Physics Letters, **1**, 87–96 (2012).

[19] K. Berrada, H Eleuch and Y. Hassouni, J. Phys. B: At. Mol. Opt. Phys., **44**, 145503 (2011).

[20] Kamal Berrada, Abderrahim Chafik, Hichem Eleuch and Yassine Hassouni, Quantum Inf. Process., **9**, 13–26 (2010).

[21] K. Berrada , M. El Baz, H. Eleuch and Y. Hassouni, International Journal of Modern Physics C,**21**, 291–305 (2010).

[22] M. J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems, **28**, 283-297 (1999).

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

[24] Q. H. Hu and D. R. Yu, Entropies of fuzzy indiscernibility relation and its operations, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, **12**, 575-589 (2004).

[25] J. S. Mi, Y. Leung and W. Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, International Journal of General Systems, **34**, 77-90 (2005).

[26] J. Y. Liang, J. H. Wang and Y. H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Information Sciences, **179**, 458-470 (2009).

[27] Z. H. Shi and Z. T. Gong, The further investigation of covering-based rough sets: Uncertainty characterization, similarity measure and generalized models, Information Sciences, **180**, 3745-3763 (2010).

[28] Z. Q. Meng and Z. Z. Shi, A fast approach to attribute reduction in incomplete decision systems with tolerance relation-based rough sets, Information Sciences, **179**, 2774-2793 (2009).

[29] J. Y. Liang and Y. H. Qian, Information granules and entropy theory in information systems, Science in China Series F: Information Sciences, **51**, 1427-1444 (2008).

[30] W. Zhu and F. Y. Wang, Reduction and axiomization of covering generalized rough sets, Information Sciences, **152**, 217-230 (2003).

[31] X. B. Yang, J. Xie, X. N. Song and J. Y. Yang, Credible rules in incomplete decision system based on descriptors, Knowledge-Based Systems, **22**, 8-17 (2009).

[32] M. W. Shao and W. X. Zhang, Dominance relation and rules in an incomplete ordered information system, International Journal of Intelligent Systems, **20**, 13-27 (2005).

[33] Y. Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences, **101**, 239-259 (1998).

[34] L. Sun and J. C. Xu, Gene selection using mutual information based uncertainty measures for tumor classication, Bio-Medical Materials and Engineering, **24**, 763-770 (2014).

[35] K. Thangavel, A. Pethalakshmi, Dimensionality reduction based on rough set theory: A review, Applied Soft Computing, **9**, 1-12 (2009).

[36] L. Sun, J. C. Xu and Y. Tian, Feature selection using rough entropy-based uncertainty measures in incomplete decision systems, Knowledge-Based Systems, **36**, 206-216 (2012).

[37] Y. Qian, J. Liang and F. Wang, A new method for measuring the uncertainty in incomplete information systems, International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, **17**, 855-880 (2009).

[38] J. H. Dai and Q. Xu, Approximations and uncertainty measures in incomplete information systems, Information Sciences, **198**, 62-80 (2012).

[39] J. H. Dai, W. T. Wang, Q. Xu and H. W. Tian, Conditional entropy for incomplete decision systems and its application in data mining, International Journal of General Systems, **41**, 713-728 (2012).
[40] J. W. Guan and D. A. Bell, Rough computational methods for information systems, International Journal of Artificial Intelligences, **105**, 77-103 (1998).
[41] G. Y. Wang, Y. Y. Yao and H. Yu, A survey on rough set theory and its application, Chinese Journal of Computers, **32**, 1229-1246 (2009).
[42] L. Sun, J. C. Xu and Y. Y. Ma, Rules extraction method of decision tree based on new conditional entropy, Computer Applications, **27**, 884-887 (2007).
[43] J. Qian, F. Y. Yee, X. P. Meng and D. Y. Liu, Attribute reduction algorithm based on new conditional information quantity, Systems Engineering and Electronics, **29**, 2154-2157 (2007).
[44] L. Qi and Y. S. Liu, Rapid reduction algorithm based on the conditional information quantity, Transactions of Beijing Institute of Technology, **27**, 604-608 (2007).
[45] S. Foithong, O. Pinngern and B. Attachoo, Feature subset selection wrapper based on mutual information and rough sets, Expert Systems with Applications, **39**, 574-584 (2012).

Lin Sun works at College of Computer & Information Engineering, Henan Normal University. He is currently a Ph.D. Candidate in School of Electronic Information and Control Engineering, Beijing University of Technology. He received his B.S. and M.S. degree in Computer Science and Technology, Henan Normal University in 2003 and 2007, respectively. His main research interests include granular computing, rough set, and data mining.

Jiucheng Xu is currently a Professor at College of Computer & Information Engineering, Henan Normal University. He received his B.S. degree in Mathematics, Henan Normal University in 1986, the M.S. degree and the Ph.D. degree in Computer Science and Technology, Xi'an Jiaotong University in 1995 and 2004, respectively. His main research interests include granular computing, rough set, data mining, and intelligent information processing.