International Journal of Uncertainty,
Fuzziness and Knowledge-Based Systems
Vol. 17, No. 6 (2009) 855–880

A NEW METHOD FOR MEASURING THE UNCERTAINTY IN INCOMPLETE INFORMATION SYSTEMS*

YUHUA QIAN†, JIYE LIANG‡ and FENG WANG§

Key Laboratory of Ministry of Education for Computational Intelligence and Chinese Information Processing, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006, China
†*jinchengqyh@126.com*
‡*ljy@sxu.edu.cn*
§*sxuwangfeng@126.com*

Received 5 March 2009
Revised 16 October 2009

Based on an intuitionistic knowledge content nature of information gain, the concept of combination entropy $CE(A)$ in incomplete information systems is first introduced, and some of its important properties are given. Then, the conditional combination entropy $CE(Q \mid P)$ and the mutual information $CE(P;Q)$ are defined. Unlike all existing measures for the uncertainty in incomplete information systems, the relationships among these three concepts can be established, which are formally expressed as $CE(Q \mid P) = CE(P \cup Q) - CE(P)$ and $CE(P;Q) = CE(P) - CE(P \mid Q)$. Furthermore, a variant $CE(C_A)$ of the combination entropy with maximal consistent block nature is introduced to measure the uncertainty of an incomplete information system in the view of maximal consistent block technique. Its monotonicity is the same as that of the combination entropy. Finally, the combination granulation $CG(A)$ and its variant $CG(C_A)$ with maximal consistent block nature are defined to measure discernibility ability of an incomplete information system, and the relationship between the combination entropy and the combination granulation is established as well. These results will be very helpful for understanding the essence of knowledge content and uncertainty measurement in incomplete information systems. Note that the combination entropy also can be further extended to measure the uncertainty in non-equivalence-based information systems.

Keywords: Incomplete information systems; combination entropy; combination granulation; maximal consistent block.

1. Introduction

Rough set theory, introduced by Pawlak,[1,2] is a relatively new soft computing tool for the analysis of a vague description of an object. The adjective vague, referring to the quality of information means inconsistency or ambiguity which follows from

*This is an extended version of the paper presented at the First International Conference on Rough Sets and Knowledge Technology, Chongqing, China.

information granulation. The rough sets philosophy is based on the assumption that with every object of the universe there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for object description. Objects having the same description are indiscernible (similar) with respect to the available information. The indiscernibility relation thus generated constitutes a mathematical basis of rough set theory; it induces a partition of the universe into blocks of indiscernible objects, called elementary sets, that can be used to build knowledge about a real or abstract world.[1–5] The use of the indiscernibility relation results in information granulation.

The notion of information systems (sometimes called data tables, attribute-value systems, knowledge representation systems, etc.), provides a convenient tool for the representation of objects in terms of their attribute values.[6–9] According to whether or not there are missing data (null values), information systems can be classified into two categories: complete and incomplete.[10]

The information entropy and the information granulation are two main approaches to measuring the uncertainty of a complete/incomplete information system. The entropy of a system as defined by Shannon gives a measure of the uncertainty about its actual structure.[11] It has been a useful mechanism for characterizing the information content in various modes and applications in many diverse fields. The information granulation is another measure for the uncertainty of an information system. In general, the information granulation represents the discernibility ability of an information/knowledge in information systems. The smaller the information granulation is, the stronger its discernibility ability is Refs. 12 and 13. Several authors (see, e.g. Refs. 14–17) have used Shannon's concept and its variants to measure the uncertainty in rough set theory. However, Shannon's entropy is not fuzzy entropy, and cannot measure the fuzziness in rough set theory. A new information entropy was proposed by Liang in Refs. 18 and 19, some important properties of this entropy were derived as well. Unlike the logarithmic behavior of Shannon's entropy, the gain function of Liang's entropy possesses the complement nature. Liang's entropy can be used to measure the fuzziness of rough set and rough classification. However, 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. Based on this consideration, Liang extended his information entropy to an incomplete information systems and established the relationship between the information entropy and the information granulation.[12] It is mentioned that the tolerance class of each object in the universe is regarded as the unit of information granule of Liang's entropy. To consider minimal information granule in an incomplete information system, Leung and Li[20] apply the concept of a maximal consistent block to formulate a new approximation to an object set with high level of accuracy. Guan[21] has extended the maximal consistent block technique to set-valued information systems. This method has been used for attribute reduction and rule acquisition in an incomplete information system. Mi *et al.*[22] gave a new fuzzy entropy and applied it for measuring the fuzziness of a fuzzy-rough set based partition. In the literature,[23] Qian and Liang proposed the concepts of combination entropy and combination granulation for measuring the uncertainty in complete information systems, which information gains possess intuitionistic knowledge content nature. Qian and Liang also investigated combination entropy and combination granulation in the context of incomplete information systems.[24] As we know, the relationship among the information entropy, the conditional entropy and the mutual information can satisfy an identical equation in a complete information system, i.e., the mutual information is equal to the difference that the information entropy is subtracted from the conditional entropy. For example, this identical equation is all satisfied by both classical Shannon's entropy and Liang's entropy in complete information systems.[11,13] However, so far this relationship has not been reported in an incomplete information system, which would baffle further research and application of information entropy theory.

This paper aims to present a new method for measuring the uncertainty of knowledge in any kind of non-equivalence-based information systems. For convenience, we only focus on incomplete information systems in this paper. In Sec. 2, some preliminary concepts such as complete information systems, incomplete information systems and maximal consistent blocks are briefly recalled. In Sec. 3, the concepts of the combination entropy, the conditional combination entropy and the mutual information of an incomplete information system are introduced, their some important properties are induced, and the relationships among them are established as well. In Sec. 3, based on maximal consistent block technique, a variant of the combination entropy is introduced to an incomplete information system, and several useful properties are given as well. In Sec. 4, the combination granulation and its variant are defined to measure the discernibility ability of an information/knowledge in incomplete information systems. The relationship between the combination entropy and the combination granulation is established. Finally, Sec. 5 concludes the paper.

2. Incomplete Information Systems

In this section, we will review some basic concepts such as incomplete information systems, tolerance relation and maximal consistent block.

An information system is a pair $S = (U, A)$, where,

(1) U is a non-empty finite set of objects;

(2) A is a non-empty finite set of attributes;

(3) for every $a \in A$, there is a mapping a, $a : U \to V_a$, where V_a is called the value set of a.

Each subset of attributes $P \subseteq A$ determines a binary indistinguishable relation $IND(P)$ as follows

$$IND(P) = \{(u, v) \in U \times U | \ \forall a \in P, a(u) = a(v)\}.$$

Table 1. The incomplete information system about car.

Car	Price	Mileage	Size	Max-Speed
u_1	high	low	full	low
u_2	low	*	full	low
u_3	*	*	compact	low
u_4	high	*	full	high
u_5	*	*	full	high
u_6	low	high	full	*

It can be easily shown that $IND(P)$ is an equivalence relation on the set U.

For $P \subseteq A$, the relation $IND(P)$ constitutes a partition of U, which is denoted by $U/IND(P)$.

It may happen that some of the attribute values for an object are missing. For example, in medical information systems there may exist a group of patients for which it is impossible to perform all the required tests. 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, a so-called null value is usually assigned to those attributes.

If V_a contains a null value for at least one attribute $a \in A$, then S is called an incomplete information system, otherwise it is complete.[25,26] Further on, we will denote the null value by $*$.

Let $S = (U, A)$ be an information system, $P \subseteq A$ an attribute set. We define a binary relation on U as follows

$$SIM(P) = \{(u, v) \in U \times U |\ \forall a \in P, a(u) = a(v) \ \text{ or } \ a(u) = * \ \text{ or } \ a(v) = *\}.$$

In fact, $SIM(P)$ is a tolerance relation on U. The concept of a tolerance relation has a wide variety of applications in classification.[25,26]

It can be easily shown that $SIM(P) = \bigcap_{a \in P} SIM(\{a\})$.

Let $S_P(u)$ denote the set $\{v \in U | (u, v) \in SIM(P)\}$. $S_P(u)$ is the maximal set of objects which are possibly indistinguishable by P with 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 noticed that the tolerance classes in $U/SIM(P)$ do not constitute a partition of U in general. They constitute a covering of U, i.e., $S_P(u) \neq \emptyset$ for every $u \in U$, and $\bigcup_{u \in U} S_P(u) = U$.

Example 1. Consider descriptions of several cars in Table 1.

This is an incomplete information system, where $U = \{u_1, u_2, u_3, u_4, u_5\}$, and $A = \{a_1, a_2, a_3, a_4\}$ with a_1-Price, a_2-Mileage, a_3-Size, a_4-Max-Speed. By computing, it follows that

$$U/SIM(A) = \{S_A(u_1), S_A(u_2), S_A(u_3), S_A(u_4),\ S_A(u_5)\},$$

where $S_A(u_1) = \{u_1\}$, $S_A(u_2) = \{u_2, u_6\}$, $S_A(u_3) = \{u_3\}$, $S_A(u_4) = \{u_4, u_5\}$, $S_A(u_5) = \{u_4, u_5, u_6\}$, $S_A(u_6) = \{u_2, u_5, u_6\}$.

Now we define a partial order on the set of all classifications of U. Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$. We say that Q is coarser than P (or P is finer than Q), denoted by $P \preceq Q$, if and only if $S_P(u_i) \subseteq S_Q(u_i)$ for $\forall i \in \{1, 2, \ldots, |U|\}$. If $P \preceq Q$ and $P \neq Q$, we say that Q is strictly coarser than P (or P is strictly finer than Q) and denoted by $P \prec Q$.

In fact, $P \prec Q \Leftrightarrow$ for $\forall i \in \{1, 2, \ldots, |U|\}$, we have that $S_P(u_i) \subseteq S_Q(u_i)$, and $\exists j \in \{1, 2, \ldots, |U|\}$, such that $S_P(u_j) \subset S_Q(u_j)$.

However, tolerance classes are not the minimal units for describing knowledge or information in incomplete information system.[20]

Let $S = (U, A)$ be an information system, $P \subseteq A$ an attribute set and $X \subseteq U$ a subset of objects. We say X is consistent with respect to P if $(u, v) \in SIM(P)$ for any $u, v \in X$. If there does not exist a subset $Y \subseteq U$ such that $X \subset Y$, and Y is consistent with respect to P, then X is called a maximal consistent block of P. Obviously, in a maximal consistent block, all objects are not indiscernible with available information provided by a similarity relation.[20]

Henceforth, we denote the set of all maximal consistent blocks determined by $P \subseteq A$ as C_P, and the set of all maximal consistent blocks of P which includes some object $u \in U$ is denoted as $C_P(u)$. It is obvious that $X \in C_P$ if and only if $X = \bigcap_{u \in X} S_P(u)$.[20]

Example 2. Computing all maximal consistent blocks of A in Table 1.

By computing, from Example 1, we have that

$$C_A = \{\{u_1\}, \{u_2, u_6\}, \{u_3\}, \{u_4, u_5\}, \{u_5, u_6\}\},$$

C_A is the set of all maximal consistent blocks determined by A on U.

Here, we define another partial relation in incomplete information systems. Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$. $C_P = \{P^1, P^2, \ldots, P^m\}$, $C_Q = \{Q^1, Q^2, \ldots, Q^n\}$. We define the partial relation $\preceq'$ as follows

$P \preceq' Q \Leftrightarrow$ for every $P^i \in C_P$, there exists $Q^j \in C_Q$ such that $P^i \subseteq Q^j$.

If $P \preceq' Q$ and $P \neq Q$, i.e., for some $P^{i_0} \in C_P$, there exists $Q^{j_0} \in C_Q$ such that $P^{i_0} \subset Q^{j_0}$, we denote it as $P \prec' Q$.

3. Combination Entropy in Incomplete Information Systems

In this section, the combination entropy in an incomplete information system is introduced to measure the uncertainty of a knowledge in incomplete information systems. Some of its properties are discussed as well.

Let $S = (U, A)$ be an incomplete information system. By $U/SIM(A) = \{S_A(u) \mid u \in U\}$, we denote the classification or the knowledge induced by A. Of particular interest is the discrete classification

$$U/SIM(A) = \omega = \{S_A(u) = \{u\} \mid u \in U\} \tag{1}$$

and the indiscrete classification

$$U/SIM(A) = \delta = \{S_A(u) = \{U\} \mid u \in U\}, \tag{2}$$

or just δ and ω if there is no confusion as to the domain set involved.

Definition 1. Let $S = (U, A)$ be an incomplete information system and $U/SIM(A) = \{S_A(u_1), S_A(u_2), \ldots, S_A(u_{|U|})\}$. The combination entropy of A is defined as

$$CE(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C_{|U|}^2 - C_{|S_A(u_i)|}^2}{C_{|U|}^2}, i \leq |U|, \tag{3}$$

where $\frac{C_{|U|}^2 - C_{|S_A(u_i)|}^2}{C_{|U|}^2}$ denotes the probability of pairs of the elements which are probably distinguishable from each other within the whole number of pairs of the elements on the universe U.

If $U/SIM(A) = \omega$, then the combination entropy of A achieves the maximum value $CE(A) = 1$;

If $U/SIM(A) = \delta$, then the combination entropy of A achieves the minimum value $CE(A) = 0$.

Obviously, for an incomplete information system $S = (U, A)$, we have that $0 \leq CE(A) \leq 1$.

Example 3. (Continued from Example 1) Computing the combination entropy of A in Table 1.

By computing, we have that

$$\begin{aligned} CE(A) &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C_{|U|}^2 - C_{|S_A(u_i)|}^2}{C_{|U|}^2} \\ &= \frac{1}{6}\left(\frac{15-0}{15} + \frac{15-1}{15} + \frac{15-0}{15} + \frac{15-1}{15} + \frac{15-3}{15} + \frac{15-3}{15}\right) \\ &= \frac{41}{45}\,. \end{aligned}$$

Remark. In fact, several authors have applied Shannon's entropy and its variants for measuring the uncertainty in the context of complete information systems.[14–17] However, these entropies can not be used in incomplete information systems. Furthermore, the literatures[27,28] investigated how to measure the uncertainty of a partition-based fuzzy rough set and that of a fuzzy information system, respectively. Whereas, these two measures only deal with the fuzziness and also can not be used for incomplete information systems. Note that the above combination entropy proposed can measure both the uncertainty in complete information systems and that in incomplete information systems, which may be better for calculating the uncertainty in the context of incomplete information systems.

Proposition 1. *Let $S = (U, A)$ be a complete information system and $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$. Then, the combination entropy of A degenerates into*

$$CE(A) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \left(1 - \frac{C_{|X_i|}^2}{C_{|U|}^2}\right),$$

i.e.,

$$CE(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \left(1 - \frac{C_{|S_A(u_i)|}^2}{C_{|U|}^2}) = \sum_{i=1}^{m} \frac{|X_i|}{|U|}(1 - \frac{C_{|X_i|}^2}{C_{|U|}^2}\right).$$

Proof. Let $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$, $X_i = \{u_{i1}, u_{i2}, \ldots, u_{is_i}\}$ $(i \leq m)$, where $|X_i| = s_i$, and $\sum\limits_{i=1}^{m} s_i = |U|$, then the relationships among the elements in $U/SIM(A)$ and the elements in $U/IND(A)$ are as follows

$$X_i = S_A(u_{i1}) = S_A(u_{i2}) = \cdots = S_A(u_{is_i}),$$

i.e.,

$$|X_i| = |S_A(u_{i1})| = |S_A(u_{i2})| = \cdots = |S_A(u_{is_i})|.$$

Hence, we have that

$$\begin{aligned}
CE(A) &= \sum_{i=1}^{m} \frac{|X_i|}{|U|}(1 - \frac{C_{|X_i|}^2}{C_{|U|}^2}) \\
&= 1 - \frac{1}{|U|} \sum_{i=1}^{m} |X_i| \times \frac{C_{|X_i|}^2}{C_{|U|}^2} \\
&= 1 - \frac{1}{|U|} \sum_{i=1}^{m} \frac{|S_A(u_{i1})| + |S_A(u_{i1})| + \cdots + |S_A(u_{is_i})|}{|X_i|} \frac{C_{|X_i|}^2}{C_{|U|}^2} \\
&= 1 - \frac{1}{|U|} \sum_{i=1}^{m} \frac{C_{|S_A(u_i)|}^2}{C_{|U|}^2} \\
&= \frac{1}{|U|} \sum_{i=1}^{|U|} (1 - \frac{C_{|S_A(u_i)|}^2}{C_{|U|}^2}).
\end{aligned}$$

This completes the proof. □

Remark. Proposition 1 states that the combination entropy in complete information systems is a special instance of the combination entropy in incomplete information systems. It means that the definition of combination entropy of complete information systems is a consistent extension to that of incomplete information systems.

Proposition 2. *Let $S = (U, A)$ be an incomplete information system and $P, Q \subseteq A$ two subsets on A. If $P \prec Q$, then $CE(P) > CE(Q)$.*

Proof. Let $U/SIM(P) = \{S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|})\}$ and $U/SIM(Q) = \{S_Q(u_1), S_Q(u_2), \ldots, S_Q(u_{|U|})\}$. If $P \prec Q$, then for $\forall i \in \{1, 2, \ldots, |U|\}$, we have that $S_P(u_i) \subseteq S_Q(u_i)$ and there exists $j \in \{1, 2, \ldots, |U|\}$ such that $S_P(u_i) \subset S_Q(u_i)$, i.e., $|S_P(u_i)| < |S_Q(u_i)|$.

Hence, we have that

$$\begin{aligned}
|S_P(u_i)| < |S_Q(u_i)| &\Rightarrow C^2_{|S_P(u_i)|} < C^2_{|S_Q(u_i)|} \\
&\Rightarrow \sum_{i=1}^{|U|} C^2_{|S_P(u_i)|} < \sum_{i=1}^{|U|} C^2_{|S_Q(u_i)|} \\
&\Rightarrow 1 - \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_Q(u_i)|}}{C^2_{|U|^2}} < 1 - \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_P(u_i)|}}{C^2_{|U|^2}} \\
&\Rightarrow CE(Q) < CE(P)\,,
\end{aligned}$$

i.e.,

$$CE(P) > CE(Q)\,.$$

This completes the proof. □

Proposition 2 states that the combination entropy of knowledge increases as tolerance classes become smaller through finer classification.

Definition 2. Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. The combination entropy of $P \cup Q$ is defined as

$$CE(P \cup Q) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|U|} - C^2_{|S_P(u_i) \cap S_Q(u_i)|}}{C^2_{|U|}}\,. \tag{4}$$

Definition 2 denotes the combination entropy of the new information system composed of two given information systems with the same universe U.

From the above definition, the following properties can be obtained.

Proposition 3. *The following properties hold*

(1) $CE(P \cup Q) \geq CE(P)$;

(2) $CE(P \cup Q) \geq CE(Q)$.

Proof. They are straightforward. □

Corollary 1. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. If $P \preceq Q$, then $CE(P \cup Q) = CE(P)$.*

Proof. It is straightforward. □

Definition 3. Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. The conditional combination entropy of Q with respect to P is defined as

$$CE(Q \mid P) = \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i)\cap S_Q(u_i)|}}{C^2_{|U|}}\,. \tag{5}$$

Definition 4. Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. The mutual information between P and Q is defined as

$$CE(P;Q) = CE(P) + CE(Q) - CE(P \cup Q)\,. \tag{6}$$

The following proposition will establish the relationships among the combination entropy, the conditional combination entropy and the mutual information in incomplete information systems.

Proposition 4. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. Then the following properties hold*

(1) $CE(Q \mid P) = CE(P \cup Q) - CE(P)$,

(2) $CE(P;Q) = CE(P) - CE(P \mid Q) = CE(Q) - CE(Q \mid P)$.

Proof. (1) From Definition 3, we have that

$$\begin{aligned}
CE(Q \mid P) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i)\cap S_Q(u_i)|}}{C^2_{|U|}}\\
&= 1 - \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_P(u_i)\cap S_Q(u_i)|}}{C^2_{|U|}} - 1 + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_P(u_i)|}}{C^2_{|U|}}\\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|U|} - C^2_{|S_P(u_i)\cap S_Q(u_i)|}}{C^2_{|U|}} - \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|U|} - C^2_{|S_P(u_i)|}}{C^2_{|U|}}\\
&= CE(P \cup Q) - CE(P)\,.
\end{aligned}$$

(2) From Definition 4 and (1), we have that

$$\begin{aligned}
CE(P;Q) &= CE(P) + CE(Q) - CE(P \cup Q)\\
&= CE(Q) - (CE(P \cup Q) - CE(P))\\
&= CE(Q) - CE(Q \mid P)\,.
\end{aligned}$$

Like this proof, the equation $CE(P;Q) = CE(P) - CE(P \mid Q)$ can be proved. This completes the proof. □

It should be noted that these equations cannot be satisfied by some existing measures in incomplete information systems. These relationships will be helping for understanding the essence of the knowledge content and the uncertainty in an incomplete information system.

Proposition 5. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems. Then $CE(Q \mid P) = 0$ iff $P \preceq Q$.*

Proof. (1) Suppose $P \preceq Q$. Hence, for arbitrary $u_i \in U$, we have that $S_P(u_i) \subseteq S_Q(u_i)$, i.e., $S_P(u_i) \cap S_Q(u_i) = S_P(u_i)$.
Therefore, we have that

$$\begin{aligned} CE(Q \mid P) &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i) \cap S_Q(u_i)|}}{C^2_{|U|}} \\ &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i)|}}{C^2_{|U|}} \\ &= 0\,. \end{aligned}$$

(2) Suppose $CE(Q \mid P) = 0$, we need to prove $P \preceq Q$. If $P \preceq Q$ not holds, then there exists some $u_j \in U$ such that $S_P(u_j) \subseteq S_Q(u_j)$ not holds, i.e.,

$$1 \leq |S_P(u_j) \cap S_Q(u_j)| < |S_P(u_j)|\,.$$

Hence, we have that

$$\begin{aligned} CE(Q \mid P) &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i) \cap S_Q(u_i)|}}{C^2_{|U|}} \\ &= \frac{1}{|U|} \sum_{i=1, i \neq j}^{|U|} \frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i) \cap S_Q(u_i)|}}{C^2_{|U|}} \\ &\quad + \frac{1}{|U|} \frac{C^2_{|S_P(u_j)|} - C^2_{|S_P(u_j) \cap S_Q(u_j)|}}{C^2_{|U|}} \\ &\geq \frac{1}{|U|} \frac{C^2_{|S_P(u_j)|} - C^2_{|S_P(u_j) \cap S_Q(u_j)|}}{C^2_{|U|}} \\ &> \frac{1}{|U|} \frac{C^2_{|S_P(u_j)|} - C^2_{|S_P(u_j)|}}{C^2_{|U|}} \\ &= 0\,. \end{aligned}$$

This yields a contradiction. Thus, $P \preceq Q$.
This completes the proof. □

Proposition 5 states that in the same universe, a knowledge cannot provide the system any additional uncertainty (classification information) if it is coarser than the original knowledge.

For knowledge discovery based on information systems, one often meets a special type of information systems, called decision information systems, which can be used to extract decision rules. In a given decision information system, there is such an attribute D which is called a decision attribute. This decision attribute usually generates the corresponding partition or covering. In the following, to reveal the relationship between condition attributes and the decision attribute, we discuss two relative properties.

Proposition 6. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems and D be a decision attribute. If $P \preceq Q$, then $CE(P \mid D) \geq CE(Q \mid D)$.*

Proof. Since $P \preceq Q$, we have that $S_P(u_i) \subseteq S_Q(u_i)$ and $|S_P(u_i) \cap S_D(u_i)| \leq |S_Q(u_i) \cap S_D(u_i)|$ for arbitrary $u_i \in U$, $S_D(u_i) \in U/SIM(D)$. Therefore, we have that

$$\begin{aligned} CE(Q \mid D) &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_D(u_i)|} - C^2_{|S_D(u_i) \cap S_Q(u_i)|}}{C^2_{|U|}} \\ &\leq \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_D(u_i)|} - C^2_{|S_D(u_i) \cap S_P(u_i)|}}{C^2_{|U|}} \\ &= CE(P \mid D)\,. \end{aligned}$$

This completes the proof. □

Proposition 6 indicates that the finer a condition knowledge is, the more classification information it can provide to an apriori knowledge (target decision).

However, the reverse relation of this proposition cannot be established in general.

Example 4. Let the universe $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Assume that

$$\begin{aligned} U/SIM(P) = \{&\{1,3,4\}, \{2,5,6\}, \{1,3,4\}, \{1,3,4\}, \{2,5,6\}, \{2,5,6\}, \{7,8,9,10\}, \\ &\{7,8,9,10\}, \{7,8,9,10\}, \{7,8,9,10\}\}, \\ U/SIM(Q) = \{&\{1,5\}, \{2,3,4,6,7\}, \{2,3,4,6,7\}, \{2,3,4,6,7\}, \{1,5\}, \{2,3,4,6,7\}, \\ &\{2,3,4,6,7\}, \{8,9,10\}, \{8,9,10\}, \{8,9,10\}\}, \end{aligned}$$

and

$$\begin{aligned} U/SIM(D) = \{&\{1,3,5,8,9\}, \{2,4,6,7,10\}, \{1,3,5,8,9\}, \{2,4,6,7,10\}, \{1,3,5,8,9\}, \\ &\{2,4,6,7,10\}, \{2,4,6,7,10\}, \{1,3,5,8,9\}, \{1,3,5,8,9\}, \{2,4,6,7,10\}, \}\,. \end{aligned}$$

It is easily computed that

$$CE(P \mid D) = \frac{1}{450}(9+9+9+10+10+9+9+9+9+9) = \frac{46}{225},$$
$$CE(Q \mid D) = \frac{1}{450}(9+4+10+4+9+4+4+9+9+10) = \frac{36}{225},$$

i.e., $CE(P \mid D) \geq CE(Q \mid D)$.

However, $P \preceq Q$ cannot be satisfied in fact.

In practice, $CE(D|P)$ can be usually used to define the significance of an attribute set. In the following, we discuss the relationship between two conditional combination entropies in incomplete information systems.

Proposition 7. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems and D be a decision attribute. If $D \preceq P \preceq Q$, then $CE(D \mid P) \leq CE(D \mid Q)$.*

Proof. Since $D \preceq P \preceq Q$, we have that $S_P(u_i) \subseteq S_Q(u_i)$ and $|S_P(u_i)| \leq |S_Q(u_i)|$, and $S_P(u_i) \cap S_D(u_i) = S_Q(u_i) \cap S_D(u_i) = S_D(u_i)$, for arbitrary $u_i \in U$, $S_D(u_i) \in U/SIM(D)$. From the definition of the conditional combination entropy, we have that

$$\begin{aligned}
&CE(D \mid P) - CE(D \mid Q)\\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_P(u_i)|} - C^2_{|S_P(u_i)\cap S_D(u_i)|}}{C^2_{|U|}} - \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_Q(u_i)|} - C^2_{|S_Q(u_i)\cap S_D(u_i)|}}{C^2_{|U|}}\\
&= \frac{1}{|U|C^2_{|U|}}\sum_{i=1}^{|U|}((C^2_{|S_D(u_i)|} - C^2_{|S_D(u_i)|}) - (C^2_{|S_Q(u_i)|} - C^2_{|S_P(u_i)|}))\\
&= \frac{1}{|U|C^2_{|U|}}\sum_{i=1}^{|U|}(C^2_{|S_P(u_i)|} - C^2_{|S_Q(u_i)|})\\
&\leq 0\,.
\end{aligned}$$

This completes the proof. □

Proposition 7 shows that the coarser a condition knowledge is, the more classification information it can preserve with respect to a target decision. In other words, if $CE(D \mid P) \leq CE(D \mid Q)$, one says that the attribute set P is more significant than the attribute set Q with respect to the target decision D.

Corollary 2. *Let $S_1 = (U, P)$, $S_2 = (U, Q)$ be two incomplete information systems and D be a decision attribute. If $P \preceq Q \preceq D$, then $CE(D \mid P) = CE(D \mid Q) = 0$.*

From the above properties and discussions, one can know that the combination entropy can well characterize the uncertainty of knowledge in incomplete information systems. In the definition of combination entropy, $C^2_{|S_A(u_i)|}$ denotes the pairs of the elements which are probably indistinguishable from each other within the tolerance class $S_A(u_i)$. In fact, given any binary relation, one can induce a cover of the universe and determine a particular information system. Through using the idea of the combination entropy, we may use the combination entropy or its variants to measure the uncertainty of the information systems induced by a given binary relation. In other words, the combination entropy can not only characterize the uncertainty of an incomplete information system, but also measure those of some more kinds of information systems.

4. The Variant of Combination Entropy for Maximal Consistent Block Technique

Because the maximal consistent block technique can describe the minimal units for a knowledge or information in incomplete information systems,[20] in this section, a variant of the combination entropy with maximal consistent block nature is introduced to measure the uncertainty of an incomplete information system. It has very useful properties.

Let $S = (U, A)$ be an incomplete information system. By $C_A = \{A^1, A^2, \ldots, A^m\}$, we denote the maximal consistent blocks induced by A. Of particular interest is the discrete classification

$$C_A = \omega = \{u\} \mid u \in U\}$$

and the indiscrete classification

$$C_A = \delta = \{U\} \mid u \in U\},$$

or just δ and ω if there is no confusion as to the domain set involved.

Definition 5. Let $S = (U, A)$ be an incomplete information system, $P \subseteq A$, $C_P = \{P^1, P^2, \ldots, P^m\}$. The combination entropy of C_P is defined as

$$CE(C_P) = \frac{1}{m}\sum_{i=1}^{m} \frac{C^2_{|U|} - C^2_{|P^i|}}{C^2_{|U|}}. \tag{7}$$

where $\frac{C^2_{|U|} - C^2_{|P^i|}}{C^2_{|U|}}$ denotes the probability of pairs of the elements which are probably distinguishable each other within the whole number of pairs of the elements on the universe U in the view of maximal consistent block technique.

If $C_P = \omega$, then the combination entropy of A achieves the maximum value $CE(C_P) = 1$;

If $C_P = \delta$, then the combination entropy of A achieves the minimum value $CE(C_P) = 0$.

Obviously, for an incomplete information system $S = (U, A)$, we have that $0 \leq CE(C_A) \leq 1$.

Example 5. (Continued from Example 2) Computing the combination entropy with maximal consistent block nature of A in Table 1.

By computing, we have that

$$\begin{aligned} CE(C_A) &= \frac{1}{m}\sum_{i=1}^{m}\frac{C^2_{|U|} - C^2_{|A^i|}}{C^2_{|U|}} \\ &= \frac{1}{5}\left(\frac{15-0}{15} + \frac{15-1}{15} + \frac{15-0}{15} + \frac{15-1}{15} + \frac{15-0}{15}\right) \\ &= \frac{24}{25}\,. \end{aligned}$$

Proposition 8. *Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$ two subsets on A, $C_P = \{P^1, P^2, \ldots, P^m\}$, $C_Q = \{Q^1, Q^2, \ldots, Q^n\}$. If $P \prec' Q$, then $CE(C_P) > CE(C_Q)$.*

Proof. Thus $P \prec' Q$, for every $P^i \in C_P$, there exists $Q^j \in C_Q$ such that $P^i \subseteq Q^j$, and for some $P^{i_0} \in C_P$, there exists $Q^{j_0} \in C_Q$ such that $P^{i_0} \subset Q^{j_0}$ and $m > n$.

For $Q^{j_0} \in C_Q$, there exist $\{P^{i_1}, P^{i_2}, \ldots, P^{i_s}\}$ ($P^{i_k} \in C_P, k = \{1, 2, \ldots, s\}$) such that each $P^{i_k} \subseteq Q^{j_0}$, where $P^{i_0} \in \{P^{i_1}, P^{i_2}, \ldots, P^{i_s}\}$. Hence, $|P^{i_k}| \leq |Q^{j_0}|$, $|P^{i_0}| < |Q^{j_0}|$, and $C^2_{|P^{i_k}|} \leq C^2_{|Q^{j_0}|}$, $C^2_{|P^{i_0}|} < C^2_{|Q^{j_0}|}$. So $\frac{1}{s}\sum_{k=1}^{s} C^2_{|P^{i_k}|} < C^2_{|Q^{j_0}|}$.

Therefore, we have that

$$\begin{aligned} CE(C_Q) &= \frac{1}{n}\sum_{j=1}^{n}\frac{C^2_{|U|} - C^2_{|Q^j|}}{C^2_{|U|}} \\ &= 1 - \frac{1}{n}\left(\sum_{j=1, j\neq j_0}^{n}\frac{C^2_{|Q^j|}}{C^2_{|U|}} + \frac{C^2_{|Q^{j_0}|}}{C^2_{|U|}}\right) \\ &< 1 - \frac{1}{n}\left(\sum_{j=1, j\neq j_0}^{n}\frac{C^2_{|Q^j|}}{C^2_{|U|}} + \frac{1}{s}\sum_{k=1}^{s}\frac{C^2_{|P^{i_k}|}}{C^2_{|U|}}\right) \\ &\leq 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{C^2_{|P^i|}}{C^2_{|U|}} \\ &= CE(C_P)\,. \end{aligned}$$

It is clear that $CE(C_P) > CE(C_Q)$.
This completes the proof. $\square$

Proposition 8 shows that the combination entropy with maximal consistent block nature of knowledge increases as maximal consistent blocks become smaller through fine classification.

Proposition 9. *Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$ two subsets on A. If $P \preceq' Q$, then $CE(P) \geq CE(Q)$.*

Proof. Suppose that $U/SIM(P) = \{S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|})\}$, $U/SIM(Q) = \{S_Q(u_1), S_Q(u_2), \ldots, S_Q(u_{|U|})\}$, $C_P = \{P^1, P^2, \ldots, P^m\}$ and $C_Q = \{Q^1, Q^2, \ldots, Q^n\}$.

It follows from the definition of $\preceq'$ that for arbitrary $P^i \in C_P$, there exists $Q^j \in C_Q$ such that $P^i \subseteq Q^j$.

Next, we prove that $S_P(u) \subseteq S_Q(u)$ for $\forall u \in U$. We know that $S_P(u) = \bigcup\{X_k \in C_P \mid X_k \subseteq S_P(u)\} = \bigcup\{X_k \in C_P(u)\}$ $(k \leq m)$ and $S_Q(u) = \bigcup\{Y_t \in C_Q \mid Y_t \subseteq S_Q(u)\} = \bigcup\{Y_t \in C_Q(u)\}$ $(t \leq n)$ from Property 4 in Ref. 20. From the definition of maximal consistent block, we have that $u \in C_P(u)$, $u \in C_Q(u)$, $u \notin C_P - C_P(u)$ and $u \notin C_Q - C_Q(u)$. Hence, it follows from $P \preceq' Q$ that for arbitrary $X_k \in C_P(u)$, there exist $Y_t \in C_Q(u)$ such that $X_k \subseteq Y_t$. Thus, for arbitrary $u \in U$, we have that

$$\begin{aligned} S_P(u) &= \bigcup\{X_k \in C_P \mid X_k \subseteq S_P(u)\} = \bigcup_{k=1}^{m} X_k \\ &\subseteq \bigcup_{t=1}^{n} Y_t = \bigcup\{Y_t \in C_Q \mid Y_t \subseteq S_Q(u)\} \\ &= S_Q(u)\,, \end{aligned}$$

that is $|S_P(u)| \leq |S_Q(u)|$.
Hence, we have that

$$\begin{aligned} CE(P) &= \frac{1}{|U|}\sum_{i=1}^{|U|} \frac{C_{|U|}^2 - C_{|S_P(u_i)|}^2}{C_{|U|}^2} \\ &\geq \frac{1}{|U|}\sum_{i=1}^{|U|} \frac{C_{|U|}^2 - C_{|S_Q(u_i)|}^2}{C_{|U|}^2} \\ &= CE(Q)\,. \end{aligned}$$

This complete the proof. □

Proposition 9 states that the combination entropy of knowledge increases as maximal consistent blocks become smaller through fine classification.

Remark. From the proof above, it is easy to see that $S_P(u) \subseteq S_Q(u)$ for arbitrary $u \in U$ if $P \preceq' Q$, i.e., the partial relation $P \preceq Q$ can be induced by the partial relation $P \preceq' Q$. Hence, the partial relation $\preceq'$ is a special instance of the partial relation $\preceq$ in incomplete information systems in fact.

5. Combination Granulation in Incomplete Information Systems

In this section, the combination granulation and its variant with maximal consistent block nature in an incomplete information system are introduced. They have some very useful properties. The relationship between the combination entropy and the combination granulation in incomplete information systems is established.

Definition 6. Let $S = (U, A)$ be an incomplete information system, $U/SIM(A) = \{S_A(u_1), S_A(u_2), \ldots, S_A(u_{|U|})\}$. The combination granulation of A is defined as

$$CG(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}}\,, \tag{8}$$

where $\frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}}$ denotes the probability of pairs of the elements on tolerance class $S_A(u_i)$ within the whole number of pairs of the elements on the universe U.

If $U/SIM(A) = \delta$, then the combination granulation of A achieves the maximum value $CG(A) = 1$.

If $U/SIM(A) = \omega$, then the combination granulation of A achieves the minimum value $CG(A) = 0$.

Clearly, for an incomplete information system $S = (U, A)$, we have that $0 \leq CE(G) \leq 1$.

Proposition 10. *Let $S = (U, A)$ be a complete information system and $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$. Then knowledge granulation of A degenerates into*

$$CG(A) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C^2_{|X_i|}}{C^2_{|U|}}\,,$$

i.e.,

$$CG(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_A(U_i)|}}{C^2_{|U|}} = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C^2_{|X_i|}}{C^2_{|U|}}\,.$$

Proof. Let $U/SIM(A) = \{X_1, X_2, \ldots, X_m\}$, $X_i = \{u_{i1}, u_{i2}, \ldots, u_{is_i}\}$, where $|X_i| = s_i$. $\sum_{i=1}^{m} |s_i| = |U|$.

The relationship between the elements in $U/SIM(A)$ and the elements in $U/IND(A)$ can be described as follows

$$X_i = S_A(u_{i1}) = S_A(u_{i2}) = \cdots = S_A(u_{is_i})\,,$$

i.e.,

$$|X_i| = |S_A(u_{i1})| = |S_A(u_{i2})| = \cdots = |S_A(u_{is_i})|\,.$$

Therefore, we have that

$$\begin{aligned} CG(A) &= \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C^2_{|X_i|}}{C^2_{|U|}} \\ &= \frac{1}{|U|} \sum_{i=1}^{m} \frac{|S_A(u_{i1})| + |S_A(u_{i2})| + \cdots + |S_A(u_{is_i})|}{|X_i|} \frac{C^2_{|X_i|}}{C^2_{|U|}} \\ &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}} . \end{aligned}$$

This completes the proof. □

Remark. Proposition 10 states that the combination granulation in complete information systems is a special instance of the combination granulation in incomplete information systems. It means that the definition of combination granulation in complete information systems is a consistent extension to that of incomplete information systems.

Proposition 11. *Let $S = (U, A)$ be an incomplete information system and $P, Q \subseteq A$ two subsets on A. If $P \prec Q$, then $CG(P) < CG(Q)$.*

Proof. Let $U/SIM(P) = \{S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|})\}$ and $U/SIM(Q) = \{S_Q(u_1), S_Q(u_2), \ldots, S_Q(u_{|U|})\}$. If $P \prec Q$, then $S_P(u_i) \subseteq S_Q(u_i)$ ($i \in \{1, 2, \ldots, |U|\}$), and $\exists j \in \{1, 2, \ldots, |U|\}$ such that $S_P(u_i) \subset S_Q(u_j)$, i.e., $|S_P(u_j)| < |S_Q(u_j)|$.

Hence, it follows that

$$\begin{aligned} |S_P(u_j)| < |S_Q(u_j)| &\Rightarrow C^2_{|S_P(u_j)|} < C^2_{|S_Q(u_j)|} \\ &\Rightarrow \sum_{i=1}^{|U|} C^2_{|S_P(u_i)|} < \sum_{i=1}^{|U|} C^2_{|S_Q(u_i)|} \\ &\Rightarrow CG(P) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_P(u_i)|}}{C^2_{|U|}} < \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_Q(u_i)|}}{C^2_{|U|}} = CG(Q) , \end{aligned}$$

i.e.,

$$CG(P) < CG(Q) .$$

This completes the proof. □

Proposition 11 states that the combination granulation of knowledge decreases as tolerance classes become smaller through finer classification.

Here, we will establish the relationship between the combination entropy and the combination granulation in incomplete information systems as follows.

Proposition 12. *Let $S = (U, A)$ be an incomplete information system and $U/SIM(A) = \{S_A(u_1), S_A(u_2), \ldots, S_A(u_{|U|})\}$. Then, the relationship between the combination entropy $CE(A)$ and combination granulation $CG(A)$ is as follows*

$$CE(A) + CG(A) = 1. \tag{9}$$

Proof. Let $S = (U, A)$ be an incomplete information system, $U/SIM(A) = \{S_A(u_1), S_A(u_2), \ldots, S_A(u_{|U|})\}$.

By Definition 1 and Definition 2, we have that

$$\begin{aligned} CE(A) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C_{|U|}^2 - C_{|S_A(u_i)|}^2}{C_{|U|}^2} \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\left(1 - \frac{C_{|S_A(u_i)|}^2}{C_{|U|}^2}\right) \\ &= 1 - \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C_{|S_A(u_i)|}^2}{C_{|U|}^2} \\ &= 1 - CG(A)\,. \end{aligned}$$

Obviously, $CE(A) + CG(A) = 1$.
This completes the proof. □

Remark. Proposition 12 shows the relationship between the combination entropy and the combination granulation is strict complement relationship, i.e., they possess the same capability on depicting the uncertainty of an incomplete information system.

Example 6. For Table 1, $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$, $A = \{Price, Mileage, Size, Max - Speed\}$ and $U/SIM(A) = \{\{u_1\}, \{u_2, u_6\}, \{u_3\}, \{u_4, u_5\}, \{u_4, u_5, u_6\}, \{u_2, u_5, u_6\}\}$. By computing, it follows that

$$\begin{aligned} CE(A) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C_{|U|}^2 - C_{|S_A(u_i)|}^2}{C_{|U|}^2} \\ &= \frac{1}{6}\left(1 + \frac{14}{15} + 1 + + \frac{14}{15} + \frac{12}{15} + \frac{12}{15}\right) \\ &= \frac{41}{45} \end{aligned}$$

and

$$\begin{aligned} CG(A) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}} \\ &= \frac{1}{6}\left(0+\frac{1}{15}+0+\frac{1}{15}+\frac{3}{15}+\frac{3}{15}\right) \\ &= \frac{4}{45}\,. \end{aligned}$$

It is obvious that $CE(A)+CG(A)=1$, i.e., the sum of the combination entropy and combination granulation of A is constant 1.

Definition 7. Let $S=(U,A)$ be an incomplete information system, $P\subseteq A$, $C_P=\{P^1,P^2,\dots,P^m\}$. The combination granulation of C_P is defined as

$$CG(C_P)=\frac{1}{m}\sum_{i=1}^{m}\frac{C^2_{|P^i|}}{C^2_{|U|}}. \tag{10}$$

where $\frac{C^2_{|P^i|}}{C^2_{|U|}}$ denotes the probability of pairs of the elements in the maximal consistent block P^i within the whole number of pairs of the elements on the universe U.

If $C_P=\delta$, then the combination granulation of C_P achieves the maximum value $CG(C_P)=1$.

If $C_P=\omega$, then the combination granulation of C_P achieves the minimum value $CG(C_P)=0$.

Clearly, for an incomplete information system $S=(U,A)$, we have that $0\le CE(C_A)\le 1$.

Proposition 13. *Let $S=(U,A)$ be an incomplete information system, $P,Q\subseteq A$ two subsets on A, $C_P=\{P^1,P^2,\dots,P^m\}$ and $C_Q=\{Q^1,Q^2,\dots,Q^n\}$. If $P\prec' Q$, then $CG(C_P)<CG(C_Q)$.*

Proof. Thus $P\prec' Q$, for every $P^i\in C_P$, there exists $Q^j\in C_Q$ such that $P^i\subseteq Q^j$, and for some $P^{i_0}\in C_P$, there exists $Q^{j_0}\in C_Q$ such that $P^{i_0}\subset Q^{j_0}$ and $m>n$.

For $Q^{j_0}\in C_Q$, there exist $\{P^{i_1},P^{i_2},\dots,P^{i_s}\}$ ($P^{i_k}\in C_P, k=\{1,2,\dots,s\}$) such that each $P^{i_k}\subseteq Q^{j_0}$, where $P^{i_0}\in\{P^{i_1},P^{i_2},\dots,P^{i_s}\}$. Hence, $|P^{i_k}|\le|Q^{j_0}|$, $|P^{i_0}|<|Q^{j_0}|$, and $C^2_{|P^{i_k}|}\le C^2_{|Q^{j_0}|}$, $C^2_{|P^{i_0}|}<C^2_{|Q^{j_0}|}$. So $\frac{1}{s}\sum_{k=1}^{s}C^2_{|P^{i_k}|}<C^2_{|Q^{j_0}|}$.

Therefore, we have that

$$\begin{aligned} CE(C_Q) &= \frac{1}{n}\sum_{j=1}^{n}\frac{C^2_{|Q^j|}}{C^2_{|U|}} \\ &= \frac{1}{n}\left(\sum_{j=1,j\neq j_0}^{n}\frac{C^2_{|Q^j|}}{C^2_{|U|}}+\frac{C^2_{|Q^{j_0}|}}{C^2_{|U|}}\right) \end{aligned}$$

$$
\begin{aligned}
&> \frac{1}{n}\left(\sum_{j=1, j\neq j_0}^{n} \frac{C^2_{|Q^j|}}{C^2_{|U|}} + \frac{1}{s}\sum_{k=1}^{s} \frac{C^2_{|P^{i_k}|}}{C^2_{|U|}}\right) \\
&> \frac{1}{m}\sum_{i=1}^{m} \frac{C^2_{|P^i|}}{C^2_{|U|}} \\
&= CE(C_P)\,.
\end{aligned}
$$

It is clear that $CE(C_P) < CE(C_Q)$.
This completes the proof. □

Proposition 13 states that the combination granulation of knowledge decreases as maximal consistent blocks become smaller through finer classification.

Proposition 14. *Let $S = (U, A)$ be an incomplete information system and $P, Q \subseteq A$ two subsets on A. If $P \preceq' Q$, then $CG(P) \leq CG(Q)$.*

Proof. Analogous to the proof of Proposition 8, it is easy to be proved. □

Proposition 14 shows that the combination granulation decreases with maximal consistent blocks becoming smaller through finer classification.

Analogous to Proposition 12, one can establish the relationship between the combination entropy $CE(C_A)$ and the combination granulation $CG(C_A)$ in incomplete information systems.

Proposition 15. *Let $S = (U, A)$ be an incomplete information system, $C_A = \{A^1, A^2, \ldots, A^m\}$, then the relationship between the combination entropy $CE(C_A)$ and combination granulation $CG(C_A)$ is as follows*

$$CE(C_A) + CG(C_A) = 1\,. \tag{11}$$

Proof. It is straightforward. □

Remark. Proposition 15 shows the relationship between the combination entropy $CE(C_A)$ and the combination granulation $CG(C_A)$ is the strict complement relationship, i.e., they possess the same capability on depicting the uncertainty of an incomplete information system in the view of maximal consistent block technique.

Example 7. For Table 1, $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$, $A = \{$ Price, Mileage, Size, Max-Speed $\}$, $C_A = \{\{u_1\}, \{u_2, u_6\}, \{u_3\}, \{u_4, u_5\}, \{u_5, u_6\}\}$. By computing, it follows that

$$CE(C_A) = \frac{1}{m}\sum_{i=1}^{m}\frac{C^2_{|U|} - C^2_{|A^i|}}{C^2_{|U|}}$$
$$= \frac{1}{5}\left(1 + \frac{14}{15} + 1 + \frac{14}{15} + 1\right)$$
$$= \frac{24}{25}$$

and

$$CG(C_A) = \frac{1}{m}\sum_{i=1}^{m}\frac{C^2_{|A^i|}}{C^2_{|U|}}$$
$$= \frac{1}{5}\left(0 + \frac{1}{15} + 0 + \frac{1}{15} + \frac{1}{15}\right)$$
$$= \frac{1}{25}\,.$$

It is obvious that $CE(C_A) + CG(C_A) = 1$, i.e., the sum of the combination entropy and the combination granulation of A is constant 1.

6. Comparison Analysis of Combination Entropy with Several Representative Entropies

Information entropy is one more kind of approaches to measuring the uncertainty of an information system. In this section, we review several existing forms of information entropy in information systems, and perform comparison analysis of combination entropy with each of them in the context of information systems.

The entropy of a system as defined by Shannon[11] also can be used to measure the uncertainty of an information system. In Shannon's entropy, an equivalence partition is regarded as a finite probability distribution, and the proportion of each equivalence class from a given partition within the universe is seen as its probability on the universe. It can be formally defined as follows.

Definition 8.[11] Let $S = (U, A)$ be a complete information system and $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$. Shannon's entropy of A is defined as (see Fig. 1)

$$H(A) = -\sum_{i=1}^{m} p_i \log_2 p_i = -\sum_{i=1}^{m}\frac{|X_i|}{|U|}\log_2\frac{|X_i|}{|U|}, \tag{12}$$

where $p_i = \frac{|X_i|}{|U|}$ represents the probability of equivalence class X_i within the universe U.

Although several authors have used Shannon's entropy and its variants to measure uncertainty in information systems, it has some limitations. In fact, Shannon's entropy is not a fuzzy entropy, and can not measure the fuzziness in information

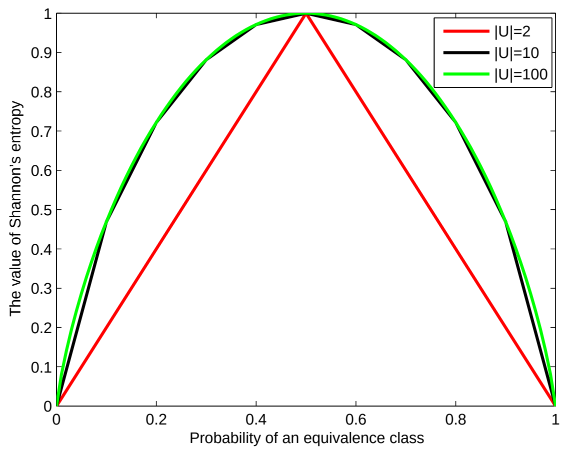

Fig. 1. Sketch map of Shannon's entropy with two equivalence classes.

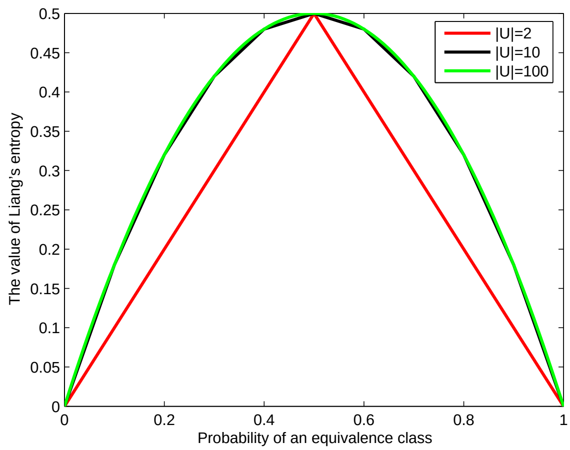

Fig. 2. Sketch map of Liang's entropy with two equivalence classes.

systems. To overcome the limitation, in the literature,[18] Liang *et al.* proposed a new information entropy. Unlike the logarithmic behavior of Shannon's entropy, the gain function of this entropy possesses the complement nature. The new entropy can be used to measure both the uncertainty of an information system and the fuzziness of a rough set and a rough classification in rough set theory. In complete information systems, Liang's information entropy is defined by the following.

Definition 9.[18] Let $S = (U, A)$ be a complete information system and $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$. The information entropy of A is defined as (see Fig. 2)

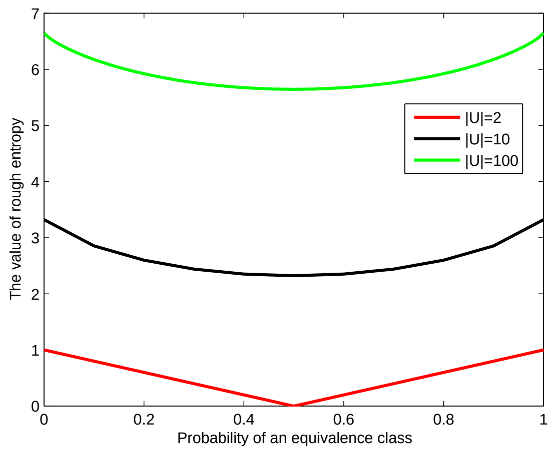

Fig. 3. Sketch map of rough entropy with two equivalence classes.

$$IE(A) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{|X_i^c|}{|U|}, \tag{13}$$

where X_i^c is the complement set of X_i, i.e., $X_i^c = U - X_i$.

In rough set theory, there is a kind of especial uncertainty, i.e., roughness.[1,13,17,23] For a given information system, we need to assess its roughness for a target concept or a target decision. An uncertainty measure, called rough entropy, is always employed to calculate roughness degree of an information system. Liang *et al.*[13] introduced the concept of rough entropy to Palawk's rough set theory for measuring the roughness degree of a complete information system. The following definition gives the depiction of the rough entropy.

Definition 10.[13] Let $S = (U, A)$ be a complete information system and $U/IND(A) = \{X_1, X_2, \ldots, X_m\}$. Rough entropy of A is defined as (see Fig. 3)

$$E_r(A) = -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \log_2 \frac{1}{|X_i|}\,. \tag{14}$$

For a complete information system, the mechanism of combination entropy can be illustrated by the following Fig. 4.

In an given partition, the elements in an equivalence class cannot be distinguished each other, but the elements in different equivalence classes can be distinguished each other. In the view of combination entropy, the knowledge content of an information system is the whole number of pairs of the elements which can be distinguished each other on the universe. Based on this consideration and analyses in this paper, we can obtain two advantages of combination entropy, which are listed as follows.

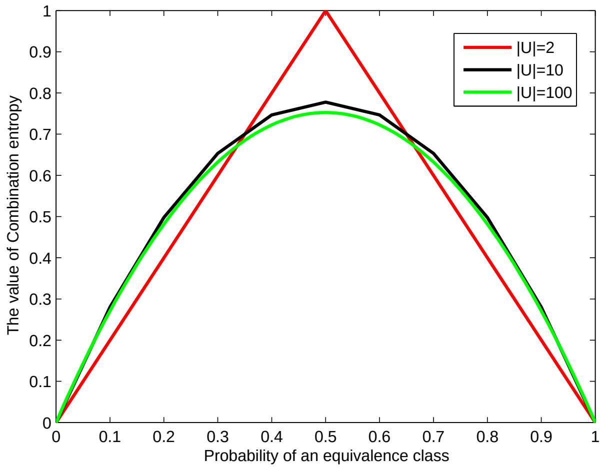

Fig. 4. Sketch map of Combination entropy with two equivalence classes.

• It can measure more clearly the knowledge content of an information system.

In fact, from Figs. 1, 2 and 4, it can be seen that the maximum value of each of Shannon's entropy and Liang's entropy is one, and the maximum value of the combination entropy varies from big to small with the size of the universe becoming larger. It is because that, in Shannon's entropy and Liang's entropy, probability distribution usually is considered. In other words, the uncertainty of an information system is based on its probability distribution. Rough entropy, in some sense, is an information granulation, which is used to measure the roughness of an information system. However, the combination entropy is based on the intuitionistic knowledge content nature of information gain, in which we need to compute the number of pairs of the elements that can be distinguished each other on the universe. Hence, the combination entropy can measure more clearly the knowledge content of an information system.

• Unlike all existing measures for the uncertainty in incomplete information systems, the relationships among the three concepts (combination entropy, conditional combination entropy and mutual information) can be established, which are formally expressed as $CE(Q \mid P) = CE(P \cup Q) - CE(P)$ and $CE(P;Q) = CE(P) - CE(P \mid Q)$. This relationship is very significant for reasonably applying an uncertainty measure to incomplete information systems. However, all existing entropies and their extensions in incomplete information systems can not establish the above relationship, which is inconsistent with classical entropy theory in statistics.

Therefore, the combination entropy may be a much better uncertainty measure for measuring the knowledge content of an information system.

7. Conclusions

In the present research, the concepts of the combination entropy $CE(A)$, the conditional combination entropy $CE(Q \mid P)$ and the mutual information $CE(P;Q)$ are introduced to incomplete information systems. Their gain functions possess the intuitionistic knowledge content nature. Unlike all existing measures for the uncertainty in incomplete information systems, the relationships among these three concepts can satisfy the equation $CE(Q \mid P) = CE(P \cup Q) - CE(P)$ and the equation $CE(P;Q) = CE(P) - CE(P \mid Q)$. Furthermore, based on the maximal consistent block technique, a variant $CE(C_A)$ of the combination entropy with maximal consistent block nature is introduced for measuring the uncertainty of an incomplete information system. Its monotonicity is the same as that of the combination entropy. Finally, the combination granulation $CG(A)$ and its variant $CG(C_A)$ with maximal consistent block nature are defined to measure knowledge granulation in incomplete information systems. The relationship between the combination entropy and the combination granulation is established as well, which is the strict complement relationship. These results have a wide variety of applications, such as measuring the knowledge content, measuring the significance of an attribute set, constructing a decision tree and building a heuristic function in a heuristic reduct algorithm in incomplete information systems. Note that the combination entropy also can be further extended to measure the uncertainty in non-equivalence-based information systems.

Acknowledgements

This work was supported by the national natural science foundation of China (No. 60773133, No. 60573074), the high technology research and development program (No. 2007AA01Z165), the foundation of doctoral program research of the ministry of education of China (No. 20050108004) and key project of science and technology research of the ministry of education of China.

References

1. Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data* (Kluwer Academic Publishers, Dordrecht, 1991).
2. Z. Pawlak, J.W. Grzymala-Busse, R. Slowiski, W. Ziarko, Rough sets, *Commun. ACM* **38(11)** (1995) 89–95.
3. J. S. Mi, W. Z. Wu, and W. X Zhang, Approaches to knowledge reduction based on variable precision rough set model, *Information Sciences* **159** (2004) 255–272.
4. W. X. Zhang, W. Z. Wu, J. Y. Liang, and D. Y. Li, *Theory and Method of Rough Sets* (Science Press, Beijing, China, 2001).
5. Q. Hu, D. Yu, Z. Xie, and J. Liu, Fuzzy probabilistic approximation spaces and their information measures, *IEEE Trans. Fuzzy Systems* **14(2)** (2006) 191–201.
6. Y. H. Qian, J. Y. Liang, D. Y. Li, H. Y. Zhang, and C. Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, *Information Sciences* **178** (2008) 181–202.

7. W. Zhu and F. Y. Wang, On three types of covering-based rough sets, *IEEE Trans. Knowledge and Data Engineering* **19(8)** (2007) 1131–1144.
8. W. Zhu and F. Y. Wang, Topological approaches to covering rough sets, *Information Sciences* **177** (2007) 1499–1508.
9. Q. Hu, D. Yu, and Z. X. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, *Pattern Recognition Letters* **27(5)** (2006) 414–423.
10. Y. H. Qian, J. Y. Liang, and C. Y. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, *Fuzzy Sets and Systems* **159** (2008) 2353–2377.
11. C. E. Shannon, The mathematical theory of communication, *The Bell System Technical Journal* **27(3,4)** (1948) 373–423; 623–656.
12. J. Y. Liang, Z. Z. Shi, D. Y. Li, and M. J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information system, *Int. J. General Systems* **35(6)** (2006) 641–654.
13. J. Y. Liang and Z. Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems* **12(1)** (2004) 37–46.
14. I. Düntsch and G. Gediga, Uncertainty measures of rough set prediction, *Artificial Intelligence* **106** (1998) 109–137.
15. T. Beaubouef, F. E. Perty, and G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, *Information Sciences* **109** (1998) 185–195.
16. G. J. Klir and M. J. Wierman, *Uncertainty Based Information* (Physica-Verlag, New York, 1998).
17. J. Y. Liang and Z. B. Xu, The algorithm on knowledge reduction in incomplete information systems, *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems* **24(1)** (2002) 95–103.
18. J. Y. Liang, K. S. Chin, C. Y. Dang, and C. M. Yam Richard, A new method for measuring uncertainty and fuzziness in rough set theory, *Int. J. General Systems* **31(4)** (2002) 331–342.
19. J. Y. Liang and D. Y. Li, *Uncertainty and Knowledge Acquisition in Information Systems* (Science Press, Beijing, China, 2005).
20. Y. Leung and D. Y. Li, Maximal consistent block technique for rule acquisition in incomplete information system, *Information Sciences* **153** (2003) 85–106.
21. Y. Y. Guan and H. K. Wang, Set-valued information systems, *Information Science* **176** (2006) 2507–2525.
22. J. S. Mi, Y. Leung, and W. Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, *Int. J. General Systems* **34(1)** (2005) 77–90.
23. Y. H. Qian and J. Y. Liang, Combination entropy and combination granulation in rough set theory, *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems*, **16(2)** (2008) 179–193.
24. Y. H. Qian and J. Y. Liang, Combination entropy and combination granulation in incomplete information systems, *Lecture Notes in Artificial Intelligence* **4062** (2006) 184–190.
25. M. Kryszkiewicz, Rough set approach to incomplete information systems, *Information Sciences* **112** (1998) 39–49.
26. M. Kryszkiewicz, Rule in incomplete information systems, *Information Sciences* **113** (1999) 271–292.
27. J. S. Mi, Y. Leung, and W. Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, *Int. J. General Systems* **34(1)** (2005) 77–90.
28. D. R. Yu, Q. H. Hu, and C. X. Wu, Uncertainty measures on fuzzy relations and their applications, *Applied Soft Computing* **7** (2007) 1135–1143.