Combination Entropy and Combination Granulation in Rough Set Theory.

2 authors, including:

Yuhua Qian
Shanxi University
213 PUBLICATIONS   **9,750** CITATIONS

Some of the authors of this publication are also working on these related projects:

Granular Computing View project

Machine learning and mining View project

COMBINATION ENTROPY AND COMBINATION GRANULATION IN ROUGH SET THEORY

YUHUA QIAN, JIYE LIANG

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

Based on the intuitionistic knowledge content nature of information gain, the concepts of combination entropy and combination granulation are introduced in rough set theory. The conditional combination entropy and the mutual information are defined and their several useful properties are derived. Furthermore, the relationship between the combination entropy and the combination granulation is established, which can be expressed as $CE(R) + CG(R) = 1$. All properties of the above concepts are all special instances of those of the concepts in incomplete information systems. These results have a wide variety of applications, such as measuring knowledge content, measuring the significance of an attribute, constructing decision trees and building a heuristic function in a heuristic reduct algorithm in rough set theory.

Keywords: Rough set theory; combination entropy; combination granulation

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 information granulation. The rough set 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 the 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[1-6]. The use of the indiscernibility relation results in knowledge granulation. The focus of rough set theory is on the ambiguity caused by limited discernibility of objects in the domain of discourse. Its key concepts are those of object indiscernibility and set approximation, and its main

perspectives are information view and algebra view[7].

The entropy of a system as defined by Shannon (1948) gives a measure of uncertainty about its actual structure[8]. It has been a useful mechanism for characterizing the information content in various modes and applications in many diverse fields. Several authors (see, e.g.[9−13]) have used Shannon's concept and its variants to measure uncertainty in rough set theory. But Shannon's entropy is not a fuzzy entropy, and cannot measure the fuzziness in rough set theory. A new information entropy was proposed by Liang in[14−16], and some important properties of this entropy are also derived. Unlike the logarithmic behavior of Shannon's entropy, the gain function of this entropy possesses the complement nature. This entropy can be used to measure the fuzziness of rough set and rough classification. In[17], Mi el al. gave a new fuzzy entropy and applied it for measuring the fuzziness of a fuzzy-rough set based partition. In[18], a new information entropy (combination entropy) and a new information granulation (combination granulation) were introduced to measure the uncertainty of an incomplete information system, and the relationship between these two concepts was established. The gain function of combination entropy possesses a intuitionistic knowledge content nature. It is mentioning that the equation $CE(P;Q) = CE(Q) - CE(Q \mid P)$ holds, where $CE(P;Q)$ denotes the mutual information between the attribute set P and Q, $CE(Q)$ is the combination entropy of Q, and $CE(Q \mid P)$ represents the conditional combination entropy Q with respect to P in incomplete information systems, respectively.

This paper aims to establish combination entropy and combination granulation in rough set theory. Some preliminary concepts such as knowledge, incomplete information systems, approximation space and partial relation are reviewed in Section 2. In Section 3, the combination entropy, the conditional combination entropy and the mutual information in rough set theory are introduced, and their several important properties are induced. In Section 3, the combination granulation in rough set theory is given. The relationship between the combination entropy and the combination granulation is established as well. Section 4 concludes the paper.

2. Preliminaries

In this section, we review some basic concepts such as knowledge, incomplete information systems, approximation space and partial relation.

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, $a : U \rightarrow V_a$, 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

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

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 $U/IND(P)$.

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[19,20]. 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 and the concept of a tolerance relation has a wide variety of applications in classification[19]. 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 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 $\bigcup_{u \in U} S_P(u) = U$.

Let $K = (U, R)$ be an approximation space, where U: a non-empty, finite set called the universe; R: an equivalence relation (i.e., indiscernibility relation) on U. $K = (U, \mathbf{R})$ can be regarded as a knowledge base about U. $\forall R \in \mathbf{R}$, the partition $U/R = \{X_1, X_2, \cdots, X_m\}$ is called the knowledge induced by equivalence relation R on U. An equivalence relation $IND(A)$ can be induced by the attribute set A in a complete information system.

Of particular interest is the discrete partition,

$$U/R = \omega = \{\{x\}, x \in U\},$$

and the indiscrete partition,

$$U/R = \delta = \{U\},$$

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

Now we define a partial order on all partition sets of U. Let P and Q be two equivalence relations of U, $U/P = \{P_1, P_2, \cdots, P_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ be partitions of the finite set U, and we define that the partition U/Q is coarser than the partition U/P (or the partition U/P is finer than the partition U/Q), i.e., $P \preceq Q$, between partitions by

$$P \preceq Q \Leftrightarrow \forall P_i \in U/P, \ \exists Q_j \in U/Q \rightarrow P_i \subseteq Q_j.$$

If $P \preceq Q$ and $P \neq Q$, then we say that U/Q is strictly coarser than U/P (or U/P is strictly finer than U/Q) and write as $P \prec Q$.

4   *YUHUA QIAN, JIYE LIANG*

3. Combination Entropy in Rough Set Theory

In general, the elements in an equivalence class cannot be distinguished each other, but the elements in different equivalence classes can be distinguished each other in rough set theory. Therefore, in a sense, the knowledge content of an approximation space $K = (U, R)$ is the whole number of pairs of the elements which can be distinguished each other on the universe U. Based on the consideration, in this section, the combination entropy, the conditional combination entropy and the mutual information in rough set theory are presented, and their some important properties are discussed.

In the first part of this section, we first given the definition of combination entropy in rough set theory.

Definition 1. Let $K = (U, R)$ be an approximation space, $U/R = \{X_1, X_2, \cdots, X_m\}$ a partition of U. Combination entropy of R is defined as

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

where $C_{|X_i|}^2 = \frac{|X_i| \times (|X_i| - 1)}{2}$, $\frac{|X_i|}{|U|}$ represents the probability of an equivalence X_i within the universe U, and $\frac{C_{|U|}^2 - C_{|X_i|}^2}{C_{|U|}^2}$ denotes the probability of pairs of the elements which are distinguishable each other within the whole number of pairs of the elements on the universe U.

If $U/R = \omega$, then the combination entropy of R achieves the maximum value 1. If $U/R = \delta$, then the combination entropy of R achieves the minimum value 0. Obviously, when U/R is a partition of U, we have that $0 \leq CE(R) \leq 1$.

Definition 2. [18] Let $S = (U, A)$ be an incomplete information system. Combination entropy of A is defined as

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

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 each other within the whole number of pairs of the elements on the universe U.

From Definition 1 and Definition 2, one can obtain the following proposition.

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

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

i.e.,

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

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

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

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

Hence, we have that

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

This completes the proof. □

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.

Definition 3. Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces, $U/P = \{P_1, P_2, \cdots, P_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ be two partitions on U. The combination entropy induced by the equivalence relation $P \cap Q$ can be defined as

$$CE(P \cap Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \frac{C_{|U|}^2 - C_{|P_i \cap Q_j|}^2}{C_{|U|}^2} \tag{4}$$

For our further development, we introduce the following lemma.

Lemma 1. Let a be a natural number and N is the set of natural numbers. If $a = \sum_{i=1}^{s} a_i$ $(s > 1)$, $a_i \in N$, then $C_a^2 > \sum_{i=1}^{s} C_{a_i}^2$.

From Definition 1 and Lemma 1, one can get the following proposition.

Proposition 2. *Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces, then $CE(P) > CE(Q)$ if $P \prec Q$.*

6   *YUHUA QIAN, JIYE LIANG*

Proof. Let $U/P = \{P_1, P_2, \cdots, P_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$. Since $P \prec Q$, we have that $m > n$ and there exists a partition $C = \{C_1, C_2, \cdots, C_n\}$ of $\{1, 2, \cdots, m\}$ such that $Q_j = \bigcup\limits_{i \in C_j} P_i$ $(j = 1, 2, \cdots, n)$. And, it follows from the definition of $P \prec Q$ that there exists $C_{j_0} \in C$ such that $|C_{j_0}| > 1$.

Thus, one has that

$$
\begin{aligned}
CE(Q) &= \sum_{j=1}^{n} \frac{|Q_j|}{|U|} \frac{C_{|U|}^2 - C_{|Q_j|}^2}{C_{|U|}^2} \\
&= 1 - \sum_{j=1}^{n} \frac{|Q_j|}{|U|} \frac{C_{|Q_j|}^2}{C_{|U|}^2} \\
&= 1 - \sum_{j=1}^{n} \frac{|\bigcup\limits_{i \in C_j} P_i|}{|U|} \frac{C_{|\bigcup\limits_{i \in C_j} P_i|}^2}{C_{|U|}^2} \\
&= 1 - \sum_{j=1}^{n} \frac{\sum\limits_{i \in C_j} |P_i|}{|U|} \frac{C_{\sum\limits_{i \in C_j} |P_i|}^2}{C_{|U|}^2} \\
&= 1 - \left(\sum_{j=1, j \neq j_0}^{n} \frac{\sum\limits_{i \in C_j} |P_i|}{|U|} \frac{C_{\sum\limits_{i \in C_j} |P_i|}^2}{C_{|U|}^2} + \frac{\sum\limits_{i \in C_{j_0}} |P_i|}{|U|} \frac{C_{\sum\limits_{i \in C_{j_0}} |P_i|}^2}{C_{|U|}^2} \right) \\
&< 1 - \sum_{i=1}^{m} \frac{|P_i|}{|U|} \frac{C_{|P_i|}^2}{C_{|U|}^2} \\
&= \sum_{i=1}^{m} \frac{|P_i|}{|U|} \frac{C_{|U|}^2 - C_{|P_i|}^2}{C_{|U|}^2} \\
&= CE(P).
\end{aligned}
$$

That is, $CE(P) > CE(Q)$. This completes the proof. □

Proposition 2 states that the combination entropy of a knowledge increases as the equivalence classes become smaller through finer partitioning in rough set theory.

In the following, in the view of the above combination entropy, we discuss the conditional combination entropy and the mutual information in rough set theory.

Definition 4. Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces. Conditional combination entropy of Q with respect to P is defined as

$$
CE(Q \mid P) = \sum_{i=1}^{m} \left(\frac{|P_i|}{|U|} \frac{C_{|P_i|}^2}{C_{|U|}^2} - \sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \frac{C_{|P_i \cap Q_j|}^2}{C_{|U|}^2} \right). \tag{5}
$$

Proposition 3. *Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces. Then,*

$$
CE(Q \mid P) = CE(P \cap Q) - CE(P). \tag{6}
$$

Proof. It easily follows from the definition of the conditional combination entropy that

$$
\begin{aligned}
CE(Q \mid P) &= \sum_{i=1}^{m} \left(\frac{|P_i|}{|U|} \frac{C_{|P_i|}^2}{C_{|U|}^2} - \sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \frac{C_{|P_i \cap Q_j|}^2}{C_{|U|}^2} \right) \\
&= \sum_{i=1}^{m} \left[\sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \left(1 - \frac{C_{|P_i \cap Q_j|}^2}{C_{|U|}^2} \right) - \frac{|P_i|}{|U|} \left(1 - \frac{C_{|P_i|}^2}{C_{|U|}^2} \right) \right]
\end{aligned}
$$

$$= CE(P \cap Q) - CE(P).$$

This completes the proof. □

Definition 5. Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces. Mutual information between P and Q is defined as

$$CE(P; Q) = CE(P) + CE(Q) - CE(P \cap Q). \tag{7}$$

The relationship among the combination entropy, the conditional combination entropy and the mutual information can be established by the following proposition.

Proposition 4. *Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces. Then,*

$$CE(P; Q) = CE(Q) - CE(Q \mid P). \tag{8}$$

Proof. It follows from Proposition 3 that
$$
\begin{aligned}
CE(P; Q) &= CE(P) + CE(Q) - CE(P \cap Q) \\
&= CE(Q) - (CE(P \cap Q) - CE(P)) \\
&= CE(Q) - CE(Q \mid P).
\end{aligned}
$$
This completes the proof. □

As follows, we investigate three important properties of the conditional combination entropy and the mutual information.

Proposition 5. *Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces. Then, $P \preceq Q$ if and only if $CE(Q \mid P) = 0$.*

Proof. (1) Suppose $P \preceq Q$. Hence, for arbitrary $P_i \in U/P$ and arbitrary $Q_j \in U/Q$, we have that $P_i \cap Q_j = \varnothing$ or $P_i \subseteq Q_j$, i.e., $|P_i \cap Q_j| = 0$ or $|P_i \cap Q_j| = |P_i|$. Therefore, one can obtain that
$$
\begin{aligned}
CE(Q \mid P) &= \sum_{i=1}^{m} \left(\frac{|P_i|}{|U|} \frac{C^2_{|P_i|}}{C^2_{|U|}} - \sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \frac{C^2_{|P_i \cap Q_j|}}{C^2_{|U|}} \right) \\
&= \frac{1}{|U|C^2_{|U|}} \sum_{i=1}^{m} \left(|P_i|C^2_{|P_i|} - \sum_{j=1}^{n} |P_i \cap Q_j|C^2_{|P_i \cap Q_j|} \right) \\
&= \frac{1}{|U|C^2_{|U|}} \sum_{i=1}^{m} \left(|P_i|C^2_{|P_i|} - |P_i|C^2_{|P_i|} \right) \\
&= 0.
\end{aligned}
$$
(2) Suppose $CE(Q \mid P) = 0$. We need to prove $P \preceq Q$. If $P \preceq Q$ does not hold, then there exists $P_k \in U/P$ such that $P_k \subseteq Q_j$ does not hold $(\forall Q_j \in U/Q)$. Let $\{Q_j \in Q \mid Q_j \cap P_k \neq \varnothing\} = \{Q_{j_1}, Q_{j_2}, \cdots, Q_{j_{k'}}\}$, where $k' > 1$, then $|P_k \cap Q_{j_l}| > 0$, $l = 1, 2, \cdots, k'$.

Therefore, we have that
$$
CE(Q \mid P) = \sum_{i=1}^{m} \left(\frac{|P_i|}{|U|} \frac{C^2_{|P_i|}}{C^2_{|U|}} - \sum_{j=1}^{n} \frac{|P_i \cap Q_j|}{|U|} \frac{C^2_{|P_i \cap Q_j|}}{C^2_{|U|}} \right)
$$

8   *YUHUA QIAN, JIYE LIANG*

$$= \frac{1}{|U|C^2_{|U|}} \sum_{i=1}^{m}(|P_i|C^2_{|P_i|} - \sum_{j=1}^{n}|P_i \cap Q_j|C^2_{|P_i \cap Q_j|})$$

$$= \frac{1}{|U|C^2_{|U|}}[\sum_{i=1,i\neq k}^{m}(|P_i|C^2_{|P_i|} - \sum_{j=1}^{n}|P_i \cap Q_j|C^2_{|P_i \cap Q_j|}) + (|P_k|C^2_{|P_k|}$$

$$- \sum_{j=1}^{n}|P_k \cap Q_j|C^2_{|P_k \cap Q_j|})]$$

$$\geq \frac{1}{|U|C^2_{|U|}}(|P_k|C^2_{|P_k|} - \sum_{j=1}^{n}|P_k \cap Q_j|C^2_{|P_k \cap Q_j|})$$

$$= \frac{1}{|U|C^2_{|U|}}(|P_k|C^2_{|P_k|} - \sum_{l=1}^{k'}|P_k \cap Q_{j_l}|C^2_{|P_k \cap Q_{j_l}|})$$

$$> 0.$$

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

Proposition 6. *Let $K_1 = (U, P)$ and $K_2 = (U, Q)$ be two approximation spaces, U/D be a decision (i.e., a known partition) on U, then $CE(P \mid D) > CE(Q \mid D)$ if $P \prec Q$.*

Proof. Let $U/P = \{P_1, P_2, \cdots, P_m\}$, $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ and $U/D = \{D_1, D_2, \cdots, D_r\}$. It follows from $P \prec Q$ that there exists a partition $C = \{C_1, C_2, \cdots, C_n\}$ of $\{1, 2, \cdots, m\}$ such that $Q_j = \bigcup_{s \in C_j} P_s$, $j = 1, 2, \cdots, n$. And, there exists $C_{j_0} \in C$ such that $|C_{j_0}| > 1$.

Therefore, we have that

$$CE(Q \mid D) = \sum_{k=1}^{r}(\frac{|D_k|}{|U|}\frac{C^2_{|D_k|}}{C^2_{|U|}} - \sum_{j=1}^{n}\frac{|Q_j \cap D_k|}{|U|}\frac{C^2_{|Q_j \cap D_k|}}{C^2_{|U|}})$$

$$= \frac{1}{|U|C^2_{|U|}}\sum_{k=1}^{r}(|D_k|C^2_{|D_k|} - \sum_{j=1}^{n}|\bigcup_{s \in C_j}(P_s \cap D_k)|C^2_{|\bigcup_{s \in C_j}(P_s \cap D_k)|})$$

$$= \frac{1}{|U|C^2_{|U|}}\sum_{k=1}^{r}(|D_k|C^2_{|D_k|} - \sum_{j=1}^{n}(\sum_{s \in C_j}|P_s \cap D_k|)C^2_{\sum_{s \in C_j}|P_s \cap D_k|})$$

$$= \frac{1}{|U|C^2_{|U|}}\sum_{k=1}^{r}(|D_k|C^2_{|D_k|} - \sum_{j=1,j\neq j_0}^{n}(\sum_{s \in C_j}|P_s \cap D_k|)C^2_{\sum_{s \in C_j}|P_s \cap D_k|}$$

$$- \sum_{s \in C_{j_0}}|P_s \cap D_k|C^2_{\sum_{s \in C_{j_0}}|P_s \cap D_k|})$$

$$< \frac{1}{|U|C^2_{|U|}}\sum_{k=1}^{r}(|D_k|C^2_{|D_k|} - \sum_{i=1}^{m}|P_i \cap D_k|C^2_{|P_i \cap D_k|})$$

$$= CE(P \mid D),$$

i.e., $CE(P \mid D) > CE(Q \mid D)$. This completes the proof. □

However, the reverse relation of this proposition cannot be established in general. This is illustrated by the following example.

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

$$U/P = \{\{1, 3, 4\}, \{2, 5, 6\}, \{7, 8, 9, 10\}\},$$

$$U/Q = \{\{1,5\},\{2,3,4,6,7\},\{8,9,10\}\},$$

and

$$U/D = \{\{1,3,5,8,9\},\{2,4,6,7,10\}\}.$$

It is easily computed that

$$CE(P \mid D) = CE(P \cap D) - CE(D) = \frac{221}{225} - \frac{7}{9} = \frac{46}{225},$$
$$CE(Q \mid D) = CE(Q \cap D) - CE(D) = \frac{211}{225} - \frac{7}{9} = \frac{36}{225},$$

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

However, $P \prec Q$ can not hold in fact.

Proposition 7. *Let $K_1 = (U,P)$ and $K_2 = (U,Q)$ be two approximation spaces, and U/D a decision (i.e., a known partition) on U. Then, $CE(P;D) \geq CE(Q;D)$ if $P \prec Q$.*

Proof. Let $U/P = \{P_1, P_2, \cdots, P_m\}$, $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$ and $U/D = \{D_1, D_2, \cdots, D_r\}$. It follows from $P \prec Q$ that there exists a partition $C = \{C_1, C_2, \cdots, C_n\}$ of $\{1, 2, \cdots, m\}$ such that $Q_j = \bigcup\limits_{s \in C_j} P_s$, $j = 1, 2, \cdots, n$.

Therefore, we have that

$$CE(Q;D) = CE(Q) + CE(D) - CE(Q \cap D)$$

$$= 1 + \sum_{j=1}^{n} \sum_{k=1}^{r} \frac{|Q_j \cap D_k|}{|U|} \frac{C^2_{|Q_j \cap D_k|}}{C^2_{|U|}} - \sum_{j=1}^{n} \frac{|Q_j|}{|U|} \frac{C^2_{|Q_j|}}{C^2_{|U|}} - \sum_{k=1}^{r} \frac{|D_k|}{|U|} \frac{C^2_{|D_k|}}{C^2_{|U|}}$$

$$= 1 + \sum_{j=1}^{n} \sum_{k=1}^{r} \frac{|\bigcup\limits_{s \in C_j}(P_s \cap D_k)|}{|U|} \frac{C^2_{|\bigcup\limits_{s \in C_j}(P_s \cap D_k)|}}{C^2_{|U|}} - \sum_{j=1}^{n} \frac{|\bigcup\limits_{s \in C_j} P_s|}{|U|} \frac{C^2_{|\bigcup\limits_{s \in C_j} P_s|}}{C^2_{|U|}}$$

$$- \sum_{k=1}^{r} \frac{|D_k|}{|U|} \frac{C^2_{|D_k|}}{C^2_{|U|}}$$

$$= 1 + \sum_{j=1}^{n} \sum_{k=1}^{r} \frac{\sum\limits_{s \in C_j}|P_s \cap D_k|}{|U|} \frac{C^2_{\sum\limits_{s \in C_j}|P_s \cap D_k|}}{C^2_{|U|}} - \sum_{j=1}^{n} \frac{\sum\limits_{s \in C_j}|P_s|}{|U|} \frac{C^2_{\sum\limits_{s \in C_j}|P_s|}}{C^2_{|U|}}$$

$$- \sum_{k=1}^{r} \frac{|D_k|}{|U|} \frac{C^2_{|D_k|}}{C^2_{|U|}}$$

$$\leq 1 + \sum_{i=1}^{m} \sum_{k=1}^{r} \frac{|P_i \cap D_k|}{|U|} \frac{C^2_{|P_i \cap D_k|}}{C^2_{|U|}} - \sum_{i=1}^{m} \frac{|P_i|}{|U|} \frac{C^2_{|P_i|}}{C^2_{|U|}} - \sum_{k=1}^{r} \frac{|D_k|}{|U|} \frac{C^2_{|D_k|}}{C^2_{|U|}}$$

$$= CE(P;D).$$

That is $CE(P;D) \geq CE(Q;D)$. This completes the proof. $\qquad\square$

Similar to Proposition 6, the reverse relation of this proposition can not hold in general.

4. Combination Granulation

In this section, the combination granulation and its very useful properties are investigated in rough set theory. The relationship between the combination entropy and the combination granulation in rough set theory is established as well.

10   *YUHUA QIAN, JIYE LIANG*

Definition 6. Let $K = (U, R)$ be an approximation space and $U/R = \{X_1, X_2, \cdots, X_m\}$ a partition of U. Combination granulation of R is defined as

$$CG(R) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C^2_{|X_i|}}{C^2_{|U|}}, \tag{9}$$

where $\frac{|X_i|}{|U|}$ represents the probability of an equivalence class X_i within the universe U and $\frac{C^2_{|X_i|}}{C^2_{|U|}}$ denotes the probability of pairs of the elements on equivalence class X_i within the whole number of pairs of the elements on the universe U.

If $U/R = \delta$, then the combination granulation of R achieves the maximum value 1.

If $U/R = \omega$, then the combination granulation of R achieves the minimum value 0.

Obviously, when U/R is a partition of U, we have that $0 \leq CG(R) \leq 1$.

Definition 7. [18] Let $S = (U, A)$ be an incomplete information system. Combination granulation of A is defined as

$$CG(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}}, \tag{10}$$

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.

The following proposition shows the relationship between these two knowledge granulations.

Proposition 8. *Let $S = (U, A)$ be an incomplete information system and $U/IND(A) = \{X_1, X_2, \cdots, X_m\}$. Then, the knowledge granulation of knowledge A degenerates into*

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

i.e.,

$$CG(A) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{C^2_{|S_A(u_i)|}}{C^2_{|U|}} = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C^2_{|X_i|}}{C^2_{|U|}}. \tag{11}$$

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

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

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

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

Therefore, one has 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. $\square$

Proposition 8 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 of complete information systems is a consistent extension to that of incomplete information systems.

Proposition 9. *Let $K = (U, R)$ be an approximation space, and U/P and U/Q be partitions of the finite set U. If $P \prec Q$, then $CG(P) < CE(Q)$.*

Proof. Let $U/P = \{P_1, P_2, \cdots, P_m\}$ and $U/Q = \{Q_1, Q_2, \cdots, Q_n\}$. Since $P \prec Q$, we have that $m > n$ and there exists a partition $C = \{C_1, C_2, \cdots, C_n\}$ of $\{1, 2, \cdots, m\}$ such that $Q_j = \bigcup_{i \in C_j} P_i$ $(j = 1, 2, \cdots, n)$. And, it follows from the definition of $P \prec Q$ that there exists $C_{j_0} \in C$ such that $|C_{j_0}| > 1$.

Thus, one can have that

$$
\begin{aligned}
CG(Q) &= \sum_{j=1}^{n} \frac{|Q_j|}{|U|} \frac{C^2_{|Q_j|}}{C^2_{|U|}} \\
&= \sum_{j=1}^{n} \frac{|\bigcup_{i \in C_j} P_i|}{|U|} \frac{C^2_{|\bigcup_{i \in C_j} P_i|}}{C^2_{|U|}} \\
&= \sum_{j=1}^{n} \frac{\sum_{i \in C_j} |P_i|}{|U|} \frac{C^2_{\sum_{i \in C_j} |P_i|}}{C^2_{|U|}} \\
&= \sum_{j=1, j \neq j_0}^{n} \frac{\sum_{i \in C_j} |P_i|}{|U|} \frac{C^2_{\sum_{i \in C_j} |P_i|}}{C^2_{|U|}} + \frac{\sum_{i \in C_{j_0}} |P_i|}{|U|} \frac{C^2_{\sum_{i \in C_{j_0}} |P_i|}}{C^2_{|U|}} \\
&> \sum_{i=1}^{m} \frac{|P_i|}{|U|} \frac{C^2_{|P_i|}}{C^2_{|U|}} \\
&= CG(Q).
\end{aligned}
$$

That is, $CG(P) < CG(Q)$. This completes the proof. $\square$

Proposition 9 states that the knowledge granulation decreases as the equivalence classes become smaller through finer partitioning.

Then, we will establish the relationship between the combination entropy and the combination granulation in rough set theory.

Proposition 10. *Let $K = (U, R)$ be an approximation space and $U/R = \{X_1, X_2, \cdots, X_m\}$ a partition of U. Then, the relationship between the combination entropy $CE(R)$ and the combination granulation $CG(R)$ is as follows*

$$CE(R) + CG(R) = 1. \tag{12}$$

Proof. It follows from the definition of $CE(R)$ and $CG(R)$ that

$$\begin{aligned}
CE(R) &= \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C_{|U|}^2 - C_{|X_i|}^2}{C_{|U|}^2} \\
&= \sum_{i=1}^{m} \frac{|X_i|}{|U|} (1 - \frac{C_{|X_i|}^2}{C_{|U|}^2}) \\
&= 1 - \sum_{i=1}^{m} \frac{|X_i|}{|U|} \frac{C_{|X_i|}^2}{C_{|U|}^2} \\
&= 1 - CG(R).
\end{aligned}$$

Obviously,

$$CG(R) + CG(R) = 1.$$

This completes the proof. $\square$

Remark. Proposition 10 states the relationship between the combination entropy and the combination granulation is strictly complement relationship. In other words, they posses the same capability for depicting the uncertainty of an approximation space. This proposition is illustrated by the following Example 2.

Example 2. Given by $U = \{$medium, small, little, tiny, big, large, huge, enormous$\}$. Let R be an equivalence relation, U/R a partition of U and $U/R = \{\{$medium$\},\{$small, little, tiny$\},\{$big, large$\},\{$huge, enormous$\}\}$.

By computing, it follows that

$CE(R) = \frac{1}{8}(1 - \frac{0}{28}) + \frac{3}{8}(1 - \frac{3}{28}) + \frac{2}{8}(1 - \frac{1}{28}) + \frac{2}{8}(1 - \frac{1}{28}) = \frac{211}{224}$,

$CG(R) = \frac{1}{8} \times \frac{0}{28} + \frac{3}{8} \times \frac{3}{28} + \frac{2}{8} \times \frac{1}{28} + \frac{2}{8} \times \frac{1}{28} = \frac{13}{224}$.

It is clear that $CE(R) + CG(R) = 1$.

5. Conclusions

In the present research, the combination entropy, the conditional combination entropy, the mutual information and the combination granulation with the intuitionistic knowledge content nature are introduced in rough set theory, respectively. Some important properties of these concepts are derived. All properties of the above concepts are all special instances of those of the concepts in incomplete information systems. Finally, the relationship between the combination entropy and the combination granulation is established, which can be formally expressed as $CE(R) + CG(R) = 1$. These results have a wide variety of applications, such as measuring knowledge content, measuring the significance of an attribute, constructing decision trees and building a heuristic function in a heuristic reduct algorithm in rough set theory.

Acknowledgements

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 and W. Ziarko, "Rough sets", *Communication of ACM*, **38(11)** (1995) 89–95.
3. J.F. Peters, A. Skowron (Eds.), *Transactions on Rough Sets III* (Springer-Verlag Berlin Heidelberg, 2005).
4. J. Bazan, H.S. Nguyen, A. Skowron, "Tough set methods in approximation of hiererchical concepts", in *Rough Sets and Current Trends in Computing TSCTC 2004*, eds. S. Tsumoto el al., Lecture Notes in Computer Science, **3006** (2004) 342-351.
5. 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.
6. W.X. Zhang, W.Z. Wu, J.Y. Liang, D.Y. Li, *Theory and Method of Rough Sets* (Science Press, Beijing, China, 2001).
7. G.Y. Wang, H. Yu, D.C. Yang, "Decision table reduction based on conditional information entropy", *Journal of Computers* (in Chinese), **25(11)** (2002) 1-8.
8. C.E. Shannon, "The mathematical theory of communication", *The Bell System Technical Journal*, **27(3, 4)** (1948) 373-423.
9. I. Düntsch, G. Gediga, "Uncertainty measures of rough set prediction", *Aritificial Intelligence*, **106** (1998) 109-137.
10. T. Beaubouef, F.E. Perty, G. Arora, "Information-theoretic measures of uncertainty for rough sets and rough relational databases", *Information Sciences*, **109** (1998) 185-195.
11. G.J. Klir, M.J. Wierman, *Uncertainty Based Information* (Physica-Verlag, New York, 1998).
12. J.Y. Liang, K.S. Qu, "Information measures of roughness of knowledge and rough sets in incomplete information systems", *Journal of System Science and System Engineering*, **24(5)** (2001) 544-547.
13. J.Y. Liang, Z.B. Xu, "The algorithm on knowledge reduction in incomplete information systems", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **24(1)** (2002) 95-103.
14. J.Y. Liang, K.S. Chin, C.Y. Dang, C.M. Yam Richard, "A new method for measuring uncertainty and fuzziness in rough set theory", *International Journal of General Systems*, **31(4)** (2002) 331-342.
15. J.Y. Liang, Z.Z. Shi, "The information entropy, rough entropy and knowledge granulation in rough set theory", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **12(1)** (2004) 37-46.
16. J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wierman, "The information entropy, rough entropy and knowledge granulation in incomplete information system", *International Journal of General Systems*, **35(6)** (2006) 641-654.

14   *YUHUA QIAN, JIYE LIANG*

17. J.S. Mi, Y. Leung, W.Z. Wu, "An uncertainty measure in partition-based fuzzy rough
     sets", *International Journal of General Systems*, **34(1)** (2005) 77-90.
18. Y.H. Qian, J.Y. Liang, "Combination entropy and combination granulation in in-
     compelte information system", *Lecture Notes in Artificial Intelligence*, **4062** (2006)
     184-190.
19. M. Kryszkiewicz, "Rough set approach to incomplete information systems", *Informa-
     tion Sciences*, **112** (1998) 39-49.
20. M. Kryszkiewicz, "Rule in incomplete information systems", *Information Sciences*,
     113 (1999) 271-292.