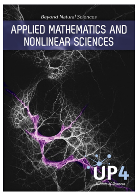
CrossMark

Applied Mathematics and Nonlinear Sciences 2(2) (2017) 329–340



Applied Mathematics and Nonlinear Sciences

http://journals.up4sciences.org

Three-way weighted combination-entropies based on three-layer granular structures

Wang Jun[1,†], Tang Lingyu[2], Zhang Xianyong[2], Luo Yuyan[3]

[1]Business School, Sichuan Normal University, China.

[2]School of Mathematical Science, Sichuan Normal University, China.

[3]Management Science School, Chengdu University of Technology, China.

Submission Info

Communicated by Juan L.G. Guirao
Received 7th January 2017
Accepted 28th July 2017
Available online 28th July 2017

Abstract

Rough set theory is an important theory for the uncertain information processing. The information theoretic measures have been introduced into rough set theory and provided a new effective method in uncertainty measurement and attribute reduction. However, most of them did not consider the hierarchical structure of a decision table (D-Table). Thus, this paper concretely constructs three-way weighted combination-entropies based on the D-Table's three-layer granular structures and Bayes' theorem from a new perspective, and reveals the granulation monotonicity and systematic relationships of three-way weighted combination-entropies. The relevant conclusion provides a more complete and updated interpretation of granular computing for the uncertainty measurement, and it also establishes a more effective basis for the quantitative application in attribute reduction.

Keywords: Rough set; Granular computing; Three-layer granular structures; Three-way weighted combination-entropies; Three-way decisions; Bayes' theorem
AMS 2010 codes: 68T30, 68T37.

1 Introduction

Rough set theory, introduced by Pawlak [1], is a kind of important theory about uncertainty information processing. So far, it has been successfully applied in data analysis, pattern recognition, machine learning and knowledge discovery, artificial intelligence, and so on [2–7]. Using the tool of entropy to deal with the uncertainty problem in rough set theory has been already studied [8–11], and the combination entropy was proposed in the literature [12].

†Corresponding author.
Email address: wangjun.sicnu@qq.com

ISSN 2444-8656 doi:10.21042/AMNS.2017.2.00027

As it is known that the rough set theory is mainly utilized to address the problem of information granules approximation problem based on a D-Table, and the granularity reflects the different levels of a given problem, then the granular computing (GrC) concerns the processing of complex information entities, information granules, which arise in the process of data abstraction and derivation of knowledge from D-Table. Hence, GrC is a kind of important structure technology, and can be able to resolve the hierarchies problems in rough set theory [13–17]. Note that three-way decisions serve as a fundamental methodology with extensive applications. Hu [18] discussed three-way decisions based on semi-three-way decision spaces, and Li et al. [17] adopted the multi-granularity to study three-way cognitive concept learning. In particular, Yao [19] pointed out that the three-level analysis falls into the category of three-way decisions, so the three-layer attribute reduction and relevant three-level measure construction become a typical case and a good example of three-way decisions.But the definition of combination entropy proposed in literature [12] did not consider the hierarchical structure of the a decision table (D-Table). Hence, on the basis of Ref. [12, 16], this paper concretely constructs three-way weighted combination-entropies based on the new perspective of a D-Table's three-layer granular structures and Bayes' theorem, and reveals the granulation monotonicity and systematic relationships of the weighted combination-entropies.

The relevant conclusions of the study has been deepened information theory of rough set theory, provides a more complete and updated interpretation of granular computing for the uncertainty measurement, and establish more effective basis of the quantitative application for attribute reduction.

2 Preliminaries

2.1 The three-layer granular structures and the three-way probabilities

This section reviews the three-layer granular structures of a given D-Table and the three-way probabilities in Ref. [16].

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

1. U is a non-empty finite set of objects;
2. AT is a non-empty finite set of attributes;
3. for every $a \in AT$, there is a mapping f_a, $f_a : U \to V_{f_a}$, where V_{f_a} is called the value set of U.

The D-Table is a special type of information table with $AT = C \cup D$ and $C \cap D = \emptyset$, where C and D denote the sets of condition attribute and decision attribute, respectively.

Each subset of attributes $A \subseteq C$ determines a binary indistinguishable relation $IND(A)$ as follows $IND(A) = \{(u,v) \in U \times U | \forall a \in A, f_a(u) = f_a(v)\}$.

$IND(A)$ serves as an equivalence relation to cause C-Class $[x]_A$, which implies a type of basic granule. The classified structure $U/IND(A) = \{[x]_A : x \in U\}$ means knowledge or C-Classification. Suppose that $U/IND(A) = \{[x]_A^i : i = 1,2,\ldots,n\}$, thus $|U/IND(A)| = n$. Similarly, D can induce the equivalence relation $IND(D)$ and further D-Classification $U/IND(D) = \{X_j : j = 1,2,\ldots,m\}$, thus $|U/IND(D)| = m$.

Aiming at the D-Table $(U, C \cup D)$ according to the four basic notions of the D-Table and four granular notions presented in Table 1, the relevant classification and class lead to three-layer granular structures, as shown in Table 2.

The three-layer granular structures (Macro-Top, Meso-Middle, and Micro-Bottom) are mainly considered from a systematic viewpoint, the numeric result and hierarchical/granular relationships are described in Fig. 1.

At the Micro-Bottom, C-Class $[x]_A^i$ and D-Class X_j are of concern. They exist in approximate space (U, AT) and can produce some fundamental measures, including probabilities. By connecting the Meso-Middle and its reasoning mechanism, three-way probabilities become bottomed measures that underlie informational construction at higher levels.

Table 1 Conditional and decisional classifications and classes

Item	C-Classification	C-Class	D-Classification	D-Class
Mathematical symbol	$U/IND(A)$	$[x]_A^i, i = 1,2,\ldots,n$	$U/IND(D)$	X_j,j=1,2,...,m
Granular essence	Conditional granule set	Conditional granule	Decisional granule set	Decisional granule

Table 2 Basic descriptions of a D-Table's three-layer granular structures

Structure	Composition	Granular scale	Granular level	Simple name
(1)	$U/IND(A), U/IND(D)$	Macro	Top	Macro-Top
(2)	$U/IND(A), X_j$	Meso	Middle	Meso-Middle
(3)	$[x]_A^i, X_j$	Micro	Bottom	Micro-Bottom

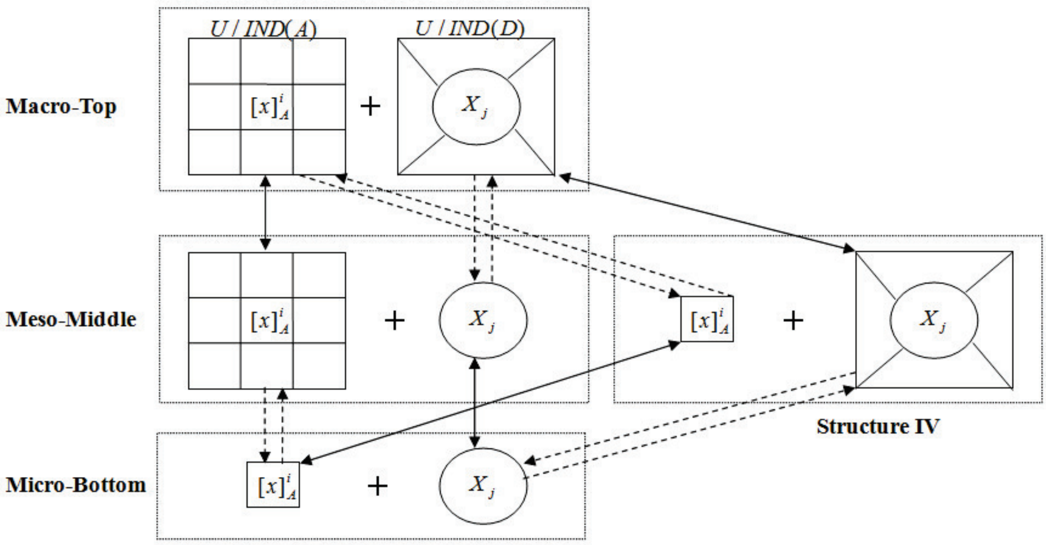


Fig. 1 Hierarchical/Granular relationships of the D-Table's three-layer granular structures.

Definition 1. At the Micro-Bottom, C-Class $[x]_A^i$ and D-Class X_j are of concern. The three-way probabilities are defined by:

$$P([x]_A^i) = \frac{|[x]_A^i|}{|U|}, P([x]_A^i/X_j) = \frac{|[x]_A^i \cap X_j|}{|X_j|}, P(X_j/([x]_A^i)) = \frac{|[x]_A^i \cap X_j|}{|[x]_A^i|}. \tag{1}$$

Theorem 1. *Three-way probabilities hold systematicness with regard to Bayes' theorem:*

$$P([x]_A^i/X_j) = \frac{P([x]_A^i) \times P(X_j/[x]_A^i)}{P(X_j)}. \tag{2}$$

2.2 Information theory of combination entropy

Regarding the D-Table, this subsection reviews the information theory of combination entropy on classifications by Ref. [12].

Definition 2. Let $K = (U,R)$ be an approximation space, $U/IND(R) = \{X_1, X_2, \ldots, X_m\}$ a partition of U. Combination entropy of R is defined as:

$$CE(R) = \sum_{i=1}^{m} \frac{\left|[x]_R^i\right|}{|U|} \frac{C_{|U|}^2 - C_{|[x]_R^i|}^2}{C_{|U|}^2} = \sum_{i=1}^{m} \frac{\left|[x]_R^i\right|}{|U|} \left(1 - \frac{C_{|[x]_R^i|}^2}{C_{|U|}^2}\right), \tag{3}$$

where $C_{|[x]_R^i|}^2 = \frac{|[x]_R^i| \times (|[x]_R^i| - 1)}{2}$, $\frac{|[x]_R^i|}{|U|}$ represents the probability of an equivalence X_i within the universe U, and $\frac{C_{|U|}^2 - C_{|[x]_R^i|}^2}{C_{|U|}^2}$ 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.

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

3 Three-way weighted combination-entropies at the Meso-Middle

Based on three-way probabilities at the Micro-Bottom, this subsection constructs three-way weighted combination-entropies at the Meso-Middle using the Bayes' theorem and discusses their granulation monotonicity and systematicness. Relevant results take a link function to underlie the latter informational construction at the Macro-Top.

A promotional measure at the Meso-Middle requires probability fusion when integrating C-Classes into C-Classification. And because Bayes' theorem provides systematicness of three-way probabilities. So it becomes the starting point. Herein, we first make a key transformation for Bayes' theorem. According to Theorem 1. with stable X_j,

Because, $P([x]_A^i/X_j) = \frac{P([x]_A^i) \times P(X_j/([x]_A^i))}{P(X_j)}, \forall i = 1, 2, \ldots, n$. So,

$$P(X_j)P(([x]_A^i)/X_j) = P([x]_A^i) \times P(X_j/([x]_A^i)). \tag{4}$$

Then, the Eq.(4) on both sides is multiplied by $\left(1 - \frac{C_{|[x]_A^i \cap X_j|}^2}{C_{|X_j|}^2}\right)$, we have:

$$P(X_j) \times P([x]_A^i)/X_j) \left(1 - \frac{C_{|[x]_A^i \cap X_j|}^2}{C_{|X_j|}^2}\right) = P([x]_A^i)P(X_j/[x]_A^i) \left(1 - \frac{C_{|[x]_A^i \cap X_j|}^2}{C_{|X_j|}^2}\right) \tag{5}$$

then, the right side of Eq.(5) can be calculated as follows:

$$
\begin{aligned}
&P([x]_A^i)P(X_j/[x]_A^i)\left(1-\frac{C^2_{|[x]_A^i\cap X_j|}}{C^2_{|X_j|}}\right)\\
&=P([x]_A^i)P\left(X_j/([x]_A^i)\right)-P([x]_A^i)P(X_j/([x]_A^i))\,\frac{C^2_{|[x]_A^i\cap X_j|}}{C^2_{|X_j|}}\\
&=P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)P(X_j/([x]_A^i))\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{|X_j|}}\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{\left|[x]_A^i\right|}}\\
&=P([x]_A^i)P\left(X_j/([x]_A^i)\right)-P\left(X_j/([x]_A^i)\right)\left(\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{\left|[x]_A^i\right|}}\right)P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)\\
&=P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)P(X_j/([x]_A^i))+P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)\left(P(X_j/([x]_A^i))-P(X_j/([x]_A^i))\left(\frac{C_{|[x]_A^i\cap X_j|^2}}{C^2_{\left|[x]_A^i\right|}}\right)\right)\\
&=P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)P(X_j/([x]_A^i))\left(\frac{C^2_{|U|}}{C^2_{|U|}}\right)+P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)\times\\
&\quad\left[P(X_j/([x]_A^i))-P(X_j/([x]_A^i))\left(\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{|[x]_A^i|}}\right)\right]\\
&=P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|U|}}\right)P(X_j/([x]_A^i)\left(\frac{C^2_{|U|}}{C^2_{|X_j|}}\right)+P([x]_A^i)P(X_j/([x]_A^i))\left(\frac{C^2_{|U|}}{C^2_{|X_j|}}\right)-\\
&\quad P([x]_A^i)P(X_j/([x]_A^i))\frac{C^2_{|U|}}{C^2_{|X_j|}}+P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)\left[P(X_j/[x]_A^i)-P(X_j/[x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{\left|[x]_A^i\right|}}\right)\right]\\
&=P(X_j/[x]_A^i)\left(\frac{C^2_{|U|}}{C^2_{|X_j|}}\right)\left[P([x]_A^i)-P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|U|}}\right)\right]+P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)P(X_j/([x]_A^i))\left(\frac{C^2_{|U|}}{C^2_{|X_j|}}\right)+\\
&\quad P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)\left[P(X_j/[x]_A^i)-P(X_j/([x]_A^i))\left(\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{|[x]_A^i|}}\right)\right],
\end{aligned}
\tag{6}
$$

thus,

$$
\begin{aligned}
&P(X_j)\times P([x]_A^i)/X_j)\left(1-\frac{C^2_{|[x]_A^i\cap X_j|}}{C^2_{|X_j|}}\right)\\
&=P(X_j/[x]_A^i)\left(\frac{C^2_{|U|}}{C^2_{|X_j|}}\right)\left[P([x]_A^i)-P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|U|}}\right)\right]+P([x]_A^i)\left(\frac{C^2_{\left|[x]_A^i\right|}}{C^2_{|X_j|}}\right)[P(X_j/([x]_A^i))-\\
&\quad P(X_j/[x]_A^i)\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{\left|[x]_A^i\right|})}]+P([x]_A^i)P(X_j/([x]_A^i))-P([x]_A^i)P(X_j/[x]_A^i)\frac{C^2_{|U|}}{C^2_{|X_j|}}.
\end{aligned}
\tag{7}
$$

According to the i-based summation,

$$
\begin{aligned}
&P(X_j)\times\sum_{i=1}^{m}P([x]_A^i)/X_j)\left(1-\frac{C^2_{|[x]_A^i\cap X_j|}}{C^2_{|X_j|}}\right)\\
&=\sum_{i=1}^{m}P(X_j/[x]_A^i)\frac{C^2_{|U|}}{C^2_{|X_j|}}\left[P([x]_A^i)\left(1-\frac{C^2_{|[x]_A^i|}}{C^2_{|U|}}\right)\right]+\sum_{i=1}^{m}P([x]_A^i)\frac{C^2_{|[x]_A^i|}}{C^2_{|X_j|}}\left[P(X_j/([x]_A^i)\left(1-\frac{C^2_{\left|[x]_A^i\cap X_j\right|}}{C^2_{|[x]_A^i|}}\right)\right]+\\
&\quad\sum_{i=1}^{m}P([x]_A^i)P(X_j/[x]_A^i))\left(1-\frac{C^2_{|U|}}{C^2_{|X_j|}}\right).
\end{aligned}
\tag{8}
$$

The final item in Eq. (8) becomes

$$\sum_{i=1}^{m} P([x]_A^i)P(X_j/([x]_A^i))\left(1-\frac{C_{|U|}^2}{C_{|X_j|}^2}\right) = P(X_j)\left(1-\frac{C_{|U|}^2}{C_{|X_j|}^2}\right). \tag{9}$$

The above step-by-step deduction implies the hierarchical evolution of Bayes' theorem. Bayes' theorem and its three-way probabilities at the Micro-Bottom evolve in the combination-entropy direction, and thus, weight-based combination-entropies and their relationships emerge at the Meco-Middle. Concretely, Eq. (9) provides a constant that is based on X_j, and thus, systematic Eq. (8) concerns three weighted and informational items. In Eq. (8) , except the final item, the others three terms derived from the combination entropy proposed in [12] are multiplied by the corresponding weight coefficients of specific probabilities. Next, we introduce the weighted combination-entropy. Suppose that (ξ, p_i) denotes a probability distribution and $\omega_i \geq 0$ means the weight, then, the weighted combination-entropy is defined as:

$$CE_\omega(\xi) = \sum_{i=1}^{n} \omega_i P_i(1-CP_i). \tag{10}$$

Definition 3. At the Meso-Middle, three-way weighted combination-entropies are defined by:

$$\begin{aligned} CE_\omega^{X_j}(A) &= \sum_{i=1}^{n} P(X_j/[x]_A^i)\frac{C_{|U|}^2}{C_{|X_j|}^2}\left[P([x]_A^i)\left(1-\frac{C_{\left|[x]_A^i\right|}^2}{C_{|U|}^2}\right)\right], \\ CE_\omega(A/X_j) &= P(X_j) \times \sum_{i=1}^{n} P([x]_A^i/X_j)\left(1-\frac{C_{\left|[x]_A^i \cap X_j\right|}^2}{C_{|X_j|}^2}\right), \\ C_\omega(X_j/A) &= \sum_{i=1}^{n} P([x]_A^i)\frac{C_{\left|[x]_A^i\right|}^2}{C_{|X_j|}^2}\left[P(X_j/[x]_A^i)\left(1-\frac{C_{\left|[x]_A^i \cap X_j\right|}^2}{C_{\left|[x]_A^i\right|}^2}\right)\right]. \end{aligned} \tag{11}$$

The weighted combination-entropy introduces weights into the combination entropy, where the weights refect the importance degrees for information receivers or attention degrees of information receivers. Concretely, $CE_\omega^{X_j}(A)$ improves absolute $\sum_{i=1}^{n} P([x]_A^i)\left(1-\frac{C_{|[x]_A^i|}^2}{C_{|U|}^2}\right)$ introducing relative $P(X_j/[x]_A^i)\frac{C_{|U|}^2}{C_{|X_j|}^2}$ to the importance weights, while $C_\omega(X_j/A)$ and $CE_\omega(A/X_j)$ respectively improve the relative $\sum_{i=1}^{n}\left[P(X_j/[x]_A^i)\left(1-\frac{C_{\left|[x]_A^i \cap X_j\right|}^2}{C_{\left|[x]_A^i\right|}^2}\right)\right]$ and $\sum_{i=1}^{n} P([x]_A^i/X_j)\left(1-\frac{C_{\left|[x]_A^i \cap X_j\right|}^2}{C_{|X_j|}^2}\right)$ by introducing absolute $P([x]_A^i)\frac{C_{\left|[x]_A^i\right|}^2}{C_{|X_j|}^2}$ and $P(X_j)$. In other words, three-way weighted combination-entropies inherit the essential uncertainty semantics of three-way properties by using different probability weights, and thus, can better describe the system regarding cause A and result X_j, hence, they become robust for uncertainty measurement. Next, we discuss their properties.

Theorem 3. *At the Meso-Middle, three-way weighted combination-entropies have granulation monotonicity. Concretely, if $P \preceq Q$, then, (1)$CE_\omega^{(X_j)}(P) \geq CE_\omega^{(X_j)}(Q)$, (2) $CE_\omega(P/X_j) \geq CE_\omega(Q/X_j)$, (3)$CE_\omega(X_j/P) \leq CE_\omega(X_j/Q)$.*

Proof. Since $P \preceq Q$, let $\cup_{t=1}^{k}[x]_P^t = [x]_Q$, then we have

(1)

$$
\begin{aligned}
CE_{\omega}^{X_j}(P) &= \sum_{t=1}^{k} P(X_j/([x]_P^t)\left(\frac{C_{|U|}^2}{C_{|X_j|}^2}\right)\left[P([x]_P^t)\left(1-\frac{C_{|[x]_P^t|}^2}{C_{|U|}^2}\right)\right]\\
&= \sum_{t=1}^{k} P(X_j/([x]_P^t)P([x]_P^t)\frac{C_{|U|}^2}{C_{|X_j|}^2} - \sum_{t=1}^{k} P(X_j/[x]_P^t)P([x]_P^t)\frac{C_{|U|}^2}{C_{|X_j|}^2}\frac{C_{|[x]_P^t|}^2}{C_{|U|}^2}\\
&= \sum_{t=1}^{k} P(X_j\cap[x]_P^t)\frac{C_{|U|}^2}{C_{|X_j|}^2} - \sum_{t=1}^{k} P(X_j\cap[x]_P^t)\frac{C_{|[x]_P^t|}^2}{C_{|X_j|}^2}\\
&= \sum_{t=1}^{k} P(X_j\cap[x]_P^t)\frac{C_{|U|}^2-C_{|[x]_P^t|}^2}{C_{|X_j|}^2}\\
&= P(X_j\cap[x]_P^1)\left(\frac{C_{|U|}^2-C_{|[x]_P^1|}^2}{C_{|X_j|}^2}\right)+P(X_j\cap[x]_P^2)\left(\frac{C_{|U|}^2-C_{|[x]_P^2|}^2}{C_{|X_j|}^2}\right)+\cdots+P(X_j\cap[x]_P^k)\left(\frac{C_{|U|}^2-C_{|[x]_P^k|}^2}{C_{|X_j|}^2}\right)\\
&\geq P(X_j\cap[x]_P^1)\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)+P(X_j\cap[x]_P^2)\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)+\cdots+P(X_j\cap[x]_P^k)\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)\\
&= \left[P(X_j\cap[x]_P^1)+P(X_j\cap[x]_P^2)+\cdots+P(X_j\cap[x]_P^k)\right]\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)\\
&= \sum_{t=1}^{k} P(X_j\cap[x]_P^t)\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)\\
&= P(X_j\cap[x]_Q)\left(\frac{C_{|U|}^2-C_{|[x]_Q|}^2}{C_{|X_j|}^2}\right)\\
&= CE_{\omega}^{X_j}(Q).
\end{aligned}
\tag{12}
$$

(2)

$$
\begin{aligned}
&CE_{\omega}(P/X_j) = P(X_j)\times\sum_{t=1}^{k} P([x]_P^t/X_j)\left(1-\frac{C_{\left|[x]_P^t\cap X_j\right|}^2}{C_{|X_j|}^2}\right)\\
&= P(X_j)\times\left[P([x]_Q/X_j)-\sum_{t=1}^{k} P([x]_P^t/X_j)\frac{C_{\left|[x]_P^t\cap X_j\right|}^2}{C_{|X_j|}^2}\right]\\
&= P(X_j)\times P([x]_Q/X_j)-P(X_j)\times\left[P([x]_P^1/X_j)\frac{C_{\left|[x]_P^1\cap X_j\right|}^2}{C_{|X_j|}^2}+P([x]_P^2/X_j)\frac{C_{\left|[x]_P^2\cap X_j\right|}^2}{C_{|X_j|}^2}+\cdots+P([x]_P^k/X_j)\frac{C_{\left|[x]_P^k\cap X_j\right|}^2}{C_{|X_j|}^2}\right]\\
&\geq P(X_j)\times P([x]_Q/X_j)-P(X_j)\times\left[P([x]_P^1/X_j)\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}+P([x]_P^2/X_j)\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}+\cdots+P([x]_P^k/X_j)\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}\right]\\
&= P(X_j)\times P([x]_Q/X_j)-P(X_j)\times\left[P([x]_P^1/X_j)+P([x]_P^2/X_j)+\cdots+P([x]_P^k/X_j)\right]\left(\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}\right)\\
&= P(X_j)\times P([x]_Q/X_j)-P(X_j)\times\sum_{i=1}^{k} P([x]_P^i/X_j)\left(\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}\right)\\
&= P(X_j)\times P([x]_Q/X_j)-P(X_j)\times P([x]_Q/X_j)\left(\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}\right)\\
&= P(X_j)\times P([x]_Q/X_j)\left(1-\frac{C_{\left|[x]_Q\cap X_j\right|}^2}{C_{|X_j|}^2}\right) = CE_{\omega}(Q/X_j).
\end{aligned}
\tag{13}
$$

(3)

$$\begin{aligned}
C_\omega(X_j/P) &= \sum_{t=1}^{k} P([x]_P^t)\frac{C^2_{|[x]_P^t|}}{C^2_{|X_j|}}\left[P(X_j/[x]_P^t)\left(1-\frac{C^2_{|[x]_P^t\cap X_j|}}{C^2_{|[x]_P^t|}}\right)\right]\\
&= \sum_{t=1}^{k} p([x]_p^t)p(X_j/[x]_p^t)\left[\frac{C^2_{|[x]_p^t|}}{C^2_{|X_j|}}\left(1-\frac{C^2_{|X_j\cap[x]_p^t|}}{C^2_{|[x]_p^t|}}\right)\right]\\
&= \sum_{t=1}^{k} p(X_j\cap[x]_p^t)\left(\frac{C^2_{|[x]_p^t|}-C^2_{|X_j\cap[x]_p^t|}}{C^2_{|X_j|}}\right)\\
&= p(X_j\cap[x]_p^1)\left(\frac{C^2_{|[x]_p^1|}-C^2_{|X_j\cap[x]_p^1|}}{C^2_{|X_j|}}\right)+p(X_j\cap[x]_p^2)\left(\frac{C^2_{|[x]_p^2|}-C^2_{|X_j\cap[x]_p^2|}}{C^2_{|X_j|}}\right)+\cdots+p(X_j\cap[x]_p^k)\left(\frac{C^2_{|[x]_p^k|}-C^2_{|X_j\cap[x]_p^k|}}{C^2_{|X_j|}}\right)\\
&\le p(X_j\cap[x]_p^1)\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right)+p(X_j\cap[x]_p^2)\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right)+\cdots+p(X_j\cap[x]_p^k)\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right)\\
&= \sum_{t=1}^{k} p(X_j\cap[x]_p^t)\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right)\\
&= \left[\sum_{t=1}^{k} p(X_j\cap[x]_p^t)\right]\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right)\\
&= p(X_j\cap[x]_Q)\left(\frac{C^2_{|[x]_Q|}-C^2_{|X_j\cap[x]_Q|}}{C^2_{|X_j|}}\right) = CW_\omega(X_j/Q).\square
\end{aligned} \tag{14}$$

Theorem 4. *Three-way weighted combination-entropies have systematicness:*

$$CE_\omega(A/X_j) = CE_\omega^{X_j}(A) + CE_\omega(X_j/A) + P(X_j)\left(1-\frac{C^2_{|U|}}{C^2_{|X_j|}}\right). \tag{15}$$

Theorem 4 provides an important relationship for the three-way weighted combination-entropies. In other words, $CE_\omega(A/X_j)$ is a linear translation of the sum of $CE_\omega^{X_j}(A)$ and $CE_\omega(X_j/A)$, where $P(X_j)[1-(C^2_{|U|})/(C^2_{|X_j|})]$ is a constant at the Meso-Middle. And it develops Bayes' theorem at the Micro-Bottom to establish a systematic equation of three-way weighted combination-entropies. Furthermore, eliminating the conversion distance can produce a new measure to simplify the systematic equation.

Definition 4. At the Meso-Middle, the linear weighted combination-entropy with regard to the weighted combination-entropy $CE_\omega(X_j/A)$ is defined as:

$$CE_\omega^{lin}(X_j/A) = CE_\omega(X_j/A) + P(X_j)\left(1-\frac{C^2_{|U|}}{C^2_{|X_j|}}\right). \tag{16}$$

Corollary 5. *At the Meso-Middle, the linear weighted combination-entropy has granulation monotonicity. Concretely, if $P \preceq Q$, then, $CE_\omega^{lin}(X_j/P) \le CE_\omega^{lin}(X_j/Q)$.*

Corollary 6. *Three-way weighted combination-entropies have the equivalent systematicness:*

$$CE_\omega(A/X_j) = CE_\omega^{X_j}(A) + CE_W^{lin}(X_j/A). \tag{17}$$

The linear weighted combination-entropy $CE_\omega^{lin}(X_j/A)$ corresponds to $CE_\omega(X_j/A)$ by virtue of a specific linear transformation. The former uses the superscript *lin* (which means linear) to different from the latter, but both are viewed as only one item for three-way weighted combination-entropies. In contrast to $CE_\omega(X_j/A)$,

$CE_{\omega}^{lin}(X_j/A)$ exhibits same granulation monotonicity, and it simplfies the systematicness of three-way weighted combination-entropies.

In summary, this section at the Meso-Middle becomes important to link the Micro-Bottom and Macro-Top. Bayes' theorem provides three-way probabilities systematicness, and it further plays a fundamental role in the informational evolution of weighted combination-entropies. It induces essential measures and systematic equations of three-way weighted combination-entropies. Next, three-way weighted combination-entropies are promoted from the Meso-Middle to the Macro-Top. combination

4 Three-way weighted combination-entropies at the macro-top

For three-way weighted combination-entropies at the Meso-Middle, their monotonicity and systematicness are established. They can hierarchically evolve to Macro-Top by using the natural sum integration with regard to multiple D-Classes. This subsection constructs three-way weighted combination-entropies at the Macro-Top and offers their monotonicity and systematicness.

Definition 5. At Macro-Top, three-way weighted combination-entropies are defined by:

$$\begin{aligned} CE_{\omega}^{D}(A) &= \sum_{j=1}^{m} CE_{\omega}^{X_j}(A), \\ CE_{\omega}(A/D) &= \sum_{j=1}^{m} CE_{\omega}(A/X_j), \\ CE_{\omega}(D/A) &= \sum_{j=1}^{m} CE_{\omega}(X_j/A). \end{aligned} \tag{18}$$

Corollary 7. *$CE_{\omega}^{lin}(D/A)$ is a linear transformation of $CE_{\omega}(D/A)$. Thus,*

$$CE_{\omega}^{lin}(D/A) = \sum_{j=1}^{m} CE_{\omega}(X_j/A) + \sum_{j=1}^{m} P(X_j)\left(1 - \frac{C_{|U|}^{2}}{C_{|X_j|}^{2}}\right) = CE_{\omega}(D/A) + CE(D), \tag{19}$$

where combination entropy $CE(D) = \sum_{j=1}^{m} P(X_j)[1-(C_{|U|}^{2})/(C_{|X_j|}^{2})]$ is a constant.

$CE_{\omega}^{lin}(D/A)$ and $CE_{\omega}(D/A)$ exhibit a linear transformation to be viewed as only one item. Three-way weighted combination-entropies at Macro-Top depend on the sum integration to naturally inherit monotonicity and systematicness at the Meso-Middle, and the relevant features are presented as follows.

Theorem 8. *At Macro-Top, three-way weighted combination-entropies have granulation monotonicity. Concretely, if $P \preceq Q$, then, $CE_{\omega}^{D}(P) \geq CE_{\omega}^{D}(Q)$; $CE_{\omega}(P/D) \geq CE_{\omega}(Q/D)$, $CE_{\omega}(D/P) \leq CE_{\omega}(D/Q)$, $CE_{\omega}^{lin}(D/P) \leq CE_{\omega}^{lin}(D/Q)$.*

Theorem 9. *Three-way weighted combination-entropies have systematicness:*

$$CE_{\omega}(A/D) = CE_{\omega}^{D}(A) + CE_{\omega}(D/A) + CE(D) = CE_{\omega}^{D}(A) + CE_{\omega}^{lin}(D/A) \tag{20}$$

At Macro-Top, Theorem 9 describes an important relationship of the three-way weighted combination-entropies by introducing $CE(D)$. Thus, $CE_{\omega}(A/D)$ is a linear translation of the summation of $CE_{\omega}^{D}(A)$ and $CE_{\omega}^{lin}(D/A)$ or the difference between $CE_{\omega}(D/A)$ and $CE_{\omega}^{lin}(D/A)$.

With regard to the Meso-Middle, the Macro-Top exhibits the hierarchical promotion and systematic integration from D-Classes to D-Classification. Accordingly, three-way weighted combination-entropies at Macro-Top are interestedly fused by three-way weighted combination-entropies at the Meso-Middle, and they exhibit a type of informational summation. The relevant results are well clarified in a relationship as shown Fig. 2.

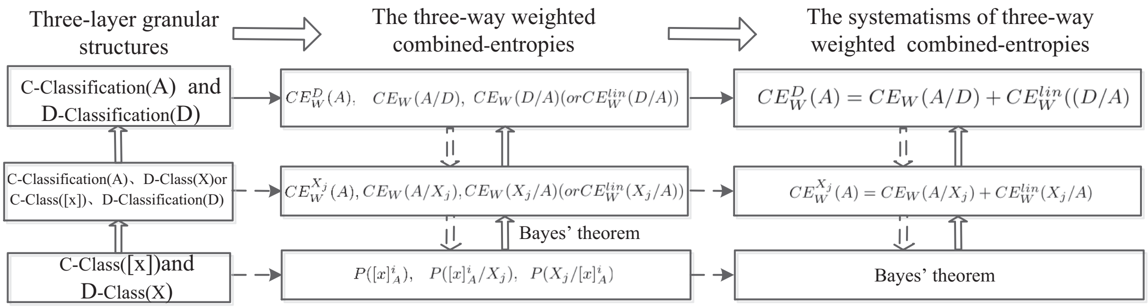


Fig. 2 Three-way weighted combination-entropies based on three-layer granular structures.

5 Conclusion

In summary, based on the new perspective of three-layer granular structures and Bayes' theorem, this paper concretely constructed three-way weighted combination-entropies, and revealed the granulation monotonicity and systematic relationships of the weighted combination-entropies. The relevant conclusion provided a more complete and updated the interpretation of granular computing for the uncertainty measurement, and established more effective basis of the quantitative application with attribute reduction.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (No.71501138, 71371130, 71501019, 61673285); Sichuan Province Social Science Planning Project (No.SC15C030, SC15C005); Scientific research project in Sichuan province department of education (No.16SB0049, 16SB0071); Specialists Development Project of Sichuan Normal University (251 Project); Open experimental project of Sichuan Normal University (No.KFSY2017044); Development and Funding Program for Middle-aged Core Teachers of Chengdu University of Technology (KYGG201519), University innovative research project of Sichuan province (15TD0004).

References

[1] Pawlak, Z. (1991). Rough Sets. Kluwer Academic Publishers.
[2] Bishop, C. M. (2006). Pattern recognition and machine learning. Company New York Ny, 16(4), 140-155.
[3] Guyon, I., Elisseeff, A. (2003). An introduction to variable feature selection. Journal of Machine Learning Research, 3, 1157-1182. doi 10.1162/153244303322753616
[4] Singh, K., Thakur, S. S., Lal, M. (2008). Vague rough set techniques for uncertainty processing in relational database model. Informatica, 19(1), 113-134. http://dl.acm.org/citation.cfm?id=1414621.1414629
[5] Chen, Y. S., Cheng, C. H. (2012). A soft-computing based rough sets classifier for classifying IPO returns in the financial markets. Applied Soft Computing, 12(1), 462-475. doi 10.1016/j.asoc.2011.08.023
[6] Fan, B., Tsang, E. C. C., Xu, W., Yu, J. (2017). Double-quantitative rough fuzzy set based decisions: a logical operations method. Information Sciences. 378, 264-281. doi 10.1016/j.ins.2016.05.035
[7] Fang, B. W., Hu, B. Q. (2016). Probabilistic graded rough set and double relative quantitative decision-theoretic rough set. International Journal of Approximate Reasoning, 74, 1-12. doi 10.1016/j.ijar.2016.03.004
[8] Weaver, Warren. (1949). The mathematical theory of communication. University of Illinois Press.
[9] Jianhua D, Qing Xu, Wentao W, Haowei Tian. (2012). Conditional entropy for incomplete decision systems and its application in data mining. International Journal of General Systems, 41(7), 713-728. doi 10.1080/03081079.2012.685471
[10] Liang, J., Shi, Z., D., L. I., Wierman, M. J. (2006). Information entropy, rough entropy and knowledge

granulation in incomplete information systems. International Journal of General Systems, 35(6), 641-654. doi 10.1080/03081070600687668

[11] Jiang, F., Sui, Y., Zhou, L. (2015). A relative decision entropy-based feature selection approach. Pattern Recognition, 48(7), 2151-2163. doi 10.1016/j.patcog.2015.01.023

[12] Qiu Y. H., Liang J. Y.. (2008). Combination entropy and combination granulation in rough set theory. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 16(02), 179-193. doi 10.1142/S0218488508005121

[13] Skowronabcd, A. (2012). Modeling rough granular computing based on approximation spaces. Information Sciences, 184(1), 20-43. doi 10.1016/j.ins.2011.08.001

[14] Zhang, X., Miao, D. (2014). Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade. Information Sciences, 268(2), 147-168. doi 10.1016/j.ins.2013.09.020

[15] Yao Y. Y. (2016). A triarchic theory of granular computing. Granular Computing 1(2), 145-157. doi 10.1007/s41066-015-0011-0

[16] Zhang, X. Y., Miao, D. (2017). Three-layer granular structures and three-way informational measures of a decision table. Information Sciences. 412-413, 67-86. doi 10.1016/j.ins.2017.05.032

[17] Li, J., Huang, C., Qi, J., Qian, Y., Liu, W. (2017). Three-way cognitive concept learning via multi-granularity. Information Sciences, 378(1), 244-263. doi 10.1016/j.ins.2016.04.051

[18] Hu, B. Q. (2017). Three-way decisions based on semi-three-way decision spaces. Information Sciences. 382, 415-440. doi 10.1016/j.ins.2016.12.012

[19] Yao, Y. Y. (2016). Three-way decisions and cognitive computing. Cognitive Computation, 8(4), 543-554. doi 10.1007/s12559-016-9397-5