

# Information structures in an incomplete information system: A granular computing viewpoint

Guangji Yu\*

*School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, Guangxi  
530003, P.R.China*

*E-mail: guangjiyu100@126.com*

Received 10 February 2018

Accepted 18 May 2018

## Abstract

Granular computing is an essential mathematical tool in artificial intelligence. An incomplete information system is an important model and its basic structures are information structures. This paper investigates information structures in an incomplete information system from granular computing viewpoint, i.e., information structures are viewed as granular structures. Information structures in an incomplete information system are first described through set vectors. Then, dependence between information structures is depicted, and information distance for calculating the difference between information structures is proposed. Next, properties of information structures in an incomplete information system are given. Finally, group and lattice characterizations of information structures in an incomplete information system are obtained. These results will be helpful for establishing a framework of granular computing in information systems.

*Keywords:* Granular computing; Incomplete information system; Information granule; Information structure; Inclusion degree; Characterization.

## 1. Introduction

Granular computing, proposed by Zadeh<sup>17</sup>, is an important tool in artificial intelligence and information processing<sup>17,18,19,20</sup>. The study of granular computing mainly concentrates in three aspects, including information granulation, organization and causation. Zadeh thought that the information granulation involves decomposition of whole into parts, the organization involves integration of parts into whole and the causation involves association of causes with effects.

The aim of granular computing is to seek for an approximation scheme, which allows us to view a

phenomenon with different levels of granularity and then can effectively solve complex problems. Information granules and granular structures are two important notions in granular computing. Information granule is a collection of objects that are drawn together by some constraints, such as indistinguishability, similarity or functionality. The process of constructing information granules is said to be information granulation. It granulates a universe of discourse into a collection of disjoint or overlapping information granules. Granular structure is the collection of information granules, in which the internal structure of each information granule is visible as a sub-structure. To manage the construction, interpre-

\* Corresponding author: Guangji Yu

tation and representation of information granules is an important problem in granular computing. Until now, the research on granular computing mainly has four methods, i.e., rough set <sup>8,9,10,11</sup>, fuzzy set <sup>16</sup>, concept lattice <sup>15</sup> and quotient space <sup>22</sup>.

Rough set is a mathematical tool for disposing uncertainty of knowledge and has been successfully applied to intelligent systems, machine learning, knowledge discovery, expert systems, decision analysis, inductive reasoning, pattern recognition, signal analysis.

An information system based on rough sets was introduced by Pawlak. Most applications of rough sets, such as rule extraction, reasoning with uncertainty, uncertainty modeling, classification and feature selection are related to information systems.

In granular computing in information systems, information granular and information structures are two important concepts. An equivalence relation is a special kind of similarities among objects from a data set. In an information system, an attribute subset determines an equivalence relation. The object set of this information system is divided into some disjoint classes by this equivalence relation, and these classes are said to be equivalence classes. If two objects of the universe belong to the same equivalence class, then we may say that they cannot be distinguished under this equivalence relation. Thus, every equivalence class can be seen as an information granule consisting of indistinguishable objects <sup>3,4,13</sup>. All these equivalence classes or information granules constitutes a vector, this vector is called an information structures in the information system induced by this attribute subset. Obviously, information structures in this information system are granular structures in the meaning of granular computing. Liang et al. <sup>7</sup> introduced information structures in an information system. Chen et al. <sup>1</sup> studied information structures in a lattice-valued information system and obtained their invariant characterizations under homomorphisms based on data compression.

So far, however, information structures in an incomplete information system have not been reported. A set vector is better than a set family in displaying the image of information structures. Based on granular computing, this paper investigates infor-

mation structures in an incomplete information system by using set vectors.

The remaining part of this paper is organized as follows. In Section 2, we recall some basic concepts about set vectors, binary relations and incomplete information systems. In Section 3, we describe information structures in an incomplete information system through set vectors. In Section 4, we divide dependence between information structures into two classes and introduce information distance between information structures. In Section 5, we give properties of information structures in an incomplete information system. In Sections 6, we obtain group and lattice characterizations of information structures in an incomplete information system. Sections 7 summarizes this paper.

## 2. Preliminaries

In this section, we recall some basic concepts about set vectors, binary relations and incomplete information systems.

Throughout this paper,  $U$  and  $V$  denotes two non-empty finite sets called the universes,  $E$  denotes the set of parameters,  $2^U$  denotes the family of all subsets of  $U$ ,  $|X|$  denotes the cardinality of  $X \in 2^U$ . All mappings are assumed to be surjective.

Put

$$\begin{aligned} U &= \{u_1, u_2, \dots, u_n\}, \\ V &= \{v_1, v_2, \dots, v_m\}, \\ \delta &= U \times U, \Delta = \{(u, u) : u \in U\}, \\ p(X) &= \frac{|X|}{|U|} (X \in 2^U). \end{aligned}$$

### 2.1. Set vectors

Recall that  $(A_1, A_2, \dots, A_n)$  is called a set vector on  $U$  if for any  $i \leq n$ ,  $A_i \in 2^U$ . We write  $\mathbf{A} = (A_1, A_2, \dots, A_n)$ .

The family of all set vectors on  $U$  is denoted by  $2^U$ .

If  $A_1 = A_2 = \dots = A_n = B$ , then  $\mathbf{A}$  is called constant. We write  $\mathbf{A} = \tilde{B}$ . Especially,

$$\tilde{U} = (U, U, \dots, U), \quad \tilde{\emptyset} = (\emptyset, \emptyset, \dots, \emptyset).$$

Let  $\mathbf{A} = (A_1, A_2, \dots, A_n)$ ,  $\mathbf{B} = (B_1, B_2, \dots, B_n) \in 2^U$ . Define

$$\mathbf{A} = \mathbf{B} \Leftrightarrow \forall i, A_i = B_i;$$

$$\mathbf{A} \preceq \mathbf{B} \Leftrightarrow \forall i, A_i \subseteq B_i;$$

$$\mathbf{A} \sqsubseteq \mathbf{B} \Leftrightarrow \exists i, A_i \subseteq B_i;$$

$$\mathbf{A} \bowtie \mathbf{B} \Leftrightarrow \forall i, A_i \not\subseteq B_i;$$

$$\mathbf{A} \odot \mathbf{B} = (A_1 \cap B_1, A_2 \cap B_2, \dots, A_n \cap B_n);$$

Obviously,

$$\mathbf{A} = \mathbf{B} \iff \mathbf{A} \preceq \mathbf{B} \text{ and } \mathbf{B} \preceq \mathbf{A};$$

$$\mathbf{A} \preceq \mathbf{B} \implies \mathbf{A} \sqsubseteq \mathbf{B}.$$

## 2.2. Binary relations

Recall that  $R$  is a binary relation on  $U$  whenever  $R \subseteq U \times U$ . If  $(x, y) \in R$ , then we denote it by  $xRy$ .

Let  $R$  be a binary relation on  $U$ . Then  $R$  is called

- (1) reflexive, if  $xRx$  for any  $x \in U$ .
- (2) symmetric, if  $xRy$  implies  $yRx$  for any  $x, y \in U$ .
- (3) transitive, if  $xRy$  and  $yRz$  imply  $xRz$  for any  $x, y, z \in U$ .

Let  $R$  be a binary relation on  $U$ . Then  $R$  is called an equivalence relation on  $U$ , if  $R$  is reflexive, symmetric and transitive.  $R$  is called a similarity relation on  $U$ , if  $R$  is reflexive and transitive.  $R$  is called a tolerance relation on  $U$ , if  $R$  is reflexive and symmetric.

## 2.3. Incomplete information systems

**Definition 2.1** <sup>(8)</sup> Let  $U$  be a set of objects and  $A$  a set of attributes. Suppose that  $U$  and  $A$  are finite sets. Then the pair  $(U, A)$  is called an information system, if each attribute  $a \in A$  determines an information function  $a : U \rightarrow V_a$ , where  $V_a$  is the information function values set of the attribute  $a$ .

If  $P \subseteq A$ , then  $(U, P)$  is called a subsystem of  $(U, A)$ .

**Definition 2.2** <sup>(8)</sup> Let  $(U, A)$  be an information system. Then the pair  $(U, A)$  is called an incomplete information system, if there are the object  $x \in U$  and the attribute  $a \in A$  such that  $a(x)$  is missing or unknown.

If  $P \subseteq A$ , then  $(U, P)$  is called a subsystem of  $(U, A)$ .

If  $(U, A)$  is an incomplete information system and  $P \subseteq A$ , then a binary relation  $sim(P)$  on  $U$  can be defined as

$$(u, v) \in sim(P) \Leftrightarrow \forall a \in A, a(u) = a(v) \text{ or } a(u) = * \text{ or } a(v) = *,$$

where  $*$  is a missing value.

Clearly,  $sim(P)$  is a tolerance relation on  $U$  and  $sim(P) = \bigcap_{a \in P} sim(\{a\})$ .

In particular, if  $sim(P) = \delta$ , then  $sim(P)$  is called a universal relation; if  $sim(P) = \Delta$ , then  $sim(P)$  is called an identity relation.

In this paper, we stipulate that  $sim(\emptyset) = \delta$ .

## 3. Concepts and operators of information structures in an incomplete information system

In this section, we give concepts and operators of information structures and study their properties.

### 3.1. Information granules and information structures

Let  $(U, A)$  be an incomplete information system, for each  $u \in U$ , denote

$$S_{sim(P)}(u) = \{v \in U : (u, v) \in sim(P)\}.$$

Then,  $S_{sim(P)}(u)$  is called the tolerance class of  $x$  under the tolerance relation  $sim(P)$ . We briefly denote  $S_{sim(P)}(u)$  by  $S_P(u)$ . If two objects  $v_1, v_2$  of  $U$  belong to the same tolerance class, then we may say that  $v_1, v_2$  cannot be distinguished under the tolerance relation  $sim(P)$  <sup>3,4,13</sup>.

The universe  $U$  may be divided into some classes by the tolerance relation  $sim(P)$ , which is said to be a quotient set. This quotient set is denoted by  $U/sim(P)$ , i.e.,

$$U/sim(P) = \{S_P(u) : u \in U\}.$$

We briefly denote  $U/sim(P)$  by  $U/P$ . An tolerance relation is a special kind of similarities among objects from a data set.

It is known that an unknown target concept can be characterized approximately by existing knowledge structures in a knowledge base, which is one of the strengths of rough set theory. Qian et al. <sup>13</sup> have studied knowledge structures in a knowledge base by means of set families. Li et al. <sup>5,6</sup> studied knowledge structures in a knowledge base and discussed relationships between knowledge bases. Their results have been shown to be very helpful for knowledge discovery from knowledge bases and significant for establishing a framework of granular computing in a knowledge base (see <sup>14</sup>).

An information structure is the collection of information granules. How do we construct information structures? It should be noted that every tolerance class consists of indistinguishable objects under some tolerance relation. Then every tolerance class can be viewed as an information granule. In this way, information structures are constructed.

Considering that a set vector is better than a set family in displaying the image of structures, information structures in an incomplete information system may be described by means of set vectors.

**Definition 3.1** Let  $(U, A)$  be an incomplete information system. Then for  $P \subseteq A$ ,

$$K(P) = (S_P(u_1), S_P(u_2), \dots, S_P(u_n))$$

is called the information structure of the subsystem  $(U, P)$ .

If  $sim(P)$  is an identity relation, then  $K(P) = (\{u_1\}, \{u_2\}, \dots, \{u_n\})$ .

**Example 3.2**  $K(\emptyset) = \tilde{U}$

**Definition 3.3** Let  $(U, A)$  be an incomplete information system. Then

$$\mathbf{K}(U, A) = \{K(P) : P \in 2^A\}$$

is called the information structure base of  $(U, A)$ .

## 4. Relationships between information structures in an incomplete information system

### 4.1. The dependency between information structures

**Definition 4.1** Let  $(U, A)$  be an incomplete information system. Given that  $K(P), K(Q)$  are the information structures of  $P, Q \subseteq A$ , respectively. Then  $K(P)$  and  $K(Q)$  are called to be the same, if for each  $i, S_P(u_i) = S_Q(u_i)$ . We still write  $K(P) = K(Q)$ .

The following definition depicts the dependency between information structures in an incomplete information system from three aspects.

**Definition 4.2** Let  $(U, A)$  be an incomplete information system. Given that  $K(P), K(Q)$  are the information structures of  $P, Q \subseteq A$ , respectively.

(1)  $K(Q)$  is called to depend on  $K(P)$ , if  $K(P) \preceq K(Q)$ ; moreover,  $K(P) \prec K(Q)$ , if  $K(P) \preceq K(Q)$  and  $K(P) \neq K(Q)$ .

(2)  $K(Q)$  is called to depend partially on  $K(P)$ , if  $K(P) \sqsubseteq K(Q)$ ; moreover,  $K(P) \sqsubset K(Q)$ , if  $K(P) \sqsubseteq K(Q)$  and  $K(P) \neq K(Q)$ .

(3)  $K(Q)$  is called to be independent of  $K(P)$ , if  $K(P) \bowtie K(Q)$ .

Obviously,

$$K(P) = K(Q) \iff K(P) \preceq K(Q) \text{ and } K(Q) \preceq K(P);$$

$$K(P) \prec K(Q) \implies K(P) \sqsubset K(Q).$$

**Example 4.3** Table 1 depicts an incomplete information system about several cars where  $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$  is the set of cars and  $A = \{Price, Mileage, Size, Max-speed\} = \{a_1, a_2, a_3, a_4\}$  is the set of attributes.

Table 1. An incomplete information system about cars

	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	*

Pick

$P_1 = \{a_1\}, P_2 = \{a_2\}, P_3 = \{a_3\}, P_4 = \{a_4\},$   
 $P_5 = \{a_1, a_2\}, P_6 = \{a_1, a_3\}, P_7 = \{a_1, a_4\}, P_8 =$   
 $\{a_2, a_3\}, P_9 = \{a_2, a_4\}, P_{10} = \{a_3, a_4\}, P_{11} =$   
 $\{a_1, a_2, a_3\}, P_{12} = \{a_1, a_2, a_4\}, P_{13} = \{a_1, a_3, a_4\},$   
 $P_{14} = \{a_2, a_3, a_4\}.$

It should be noted that

$sim\{a_1\} = \Delta \cup \{(u_1, u_4), (u_4, u_1), (u_2, u_6), (u_6, u_2)\},$   
 $sim\{a_2\} = \Delta,$   
 $sim\{a_3\} = \Delta \cup \{(u_1, u_2), (u_2, u_1), (u_1, u_4), (u_4, u_1), (u_1,$   
 $u_5), (u_5, u_1), (u_1, u_6), (u_6, u_1), (u_2, u_4), (u_4, u_2), (u_2, u_5)$   
 $, (u_5, u_2), (u_2, u_6), (u_6, u_2), (u_4, u_5), (u_5, u_4), (u_4, u_6), (u_6,$   
 $u_4), (u_5, u_6), (u_6, u_5)\},$   
 $sim\{a_4\} = \Delta \cup \{(u_1, u_2), (u_2, u_1), (u_1, u_3), (u_3, u_1), (u_2,$   
 $u_3), (u_3, u_2), (u_4, u_5), (u_5, u_4)\}.$  Then  
 $sim(\emptyset) = \delta,$   
 $sim(P_1) = sim(P_6) = sim\{a_1\},$   
 $sim(P_2) = sim(P_5) = sim(P_7) = sim(P_8) = sim(P_9) =$   
 $sim(P_{11}) = sim(P_{12}) = sim(P_{13}) = sim(P_{14}) =$   
 $sim(A) = sim\{a_2\},$   
 $sim(P_3) = sim\{a_3\},$   
 $sim(P_4) = sim\{a_4\},$   
 $sim(P_{10}) = \Delta \cup \{(u_1, u_2), (u_2, u_1), (u_4, u_5), (u_5, u_4)\}.$

Obviously,

$$K(\emptyset) = (U, U, U, U, U, U) = \tilde{U}.$$

$$K(P_1) = K(P_6) = (\{u_1, u_4\}, \{u_2, u_6\}, \{u_3\}, \{u_4,$$
  
 $u_1\}, \{u_5\}, \{u_6, u_2\}),$

$$K(P_2) = K(P_5) = K(P_7) = K(P_8) = K(P_9) =$$
  
 $K(P_{11}) = K(P_{12}) = K(P_{13}) = K(P_{14}) = K(A) =$   
 $(\{u_1\}, \{u_2\}, \{u_3\}, \{u_4\}, \{u_5\}, \{u_6\}),$

$$K(P_3) = (\{u_1, u_2, u_4, u_5, u_6\}, \{u_1, u_2, u_4, u_5, u_6\},$$
  
 $\{u_3\}, \{u_1, u_2, u_4, u_5, u_6\}, \{u_1, u_2, u_4, u_5, u_6\}, \{u_1, u_2,$   
 $u_4, u_5, u_6\}),$

$$K(P_4) = (\{u_1, u_2, u_3\}, \{u_1, u_2, u_3\}, \{u_1, u_2, u_3\},$$
  
 $\{u_4, u_5\}, \{u_4, u_5\}, \{u_6\}),$

$$K(P_{10}) = (\{u_1, u_2\}, \{u_1, u_2\}, \{u_3\}, \{u_4, u_5\},$$
  
 $\{u_4, u_5\}, \{u_6\}).$

Then

(1)

$$\mathbf{K}(U, A) = \{K(\emptyset), K(P_1), K(P_3), K(P_4), K(P_{10}), K(A)\}.$$

(2)  $K(A) \prec K(P_1) \prec K(P_3) \prec K(\emptyset), K(A) \prec$   
 $K(P_{10}) \prec K(P_4) \prec K(\emptyset); K(P_{10}) \prec K(P_3); K(P_3) \sqsubset$   
 $K(P_1) \sqsubset K(A), K(P_3) \sqsubset K(A); K(P_4) \sqsubset K(P_{10}) \sqsubset$

$K(A), K(P_4) \sqsubset K(A); K(P_1) \sqsubset K(P_{10}), K(P_{10}) \sqsubset$   
 $K(P_1); K(P_3) \sqsubset K(P_4), K(P_4) \sqsubset K(P_3); K(P_1) \sqsubset$   
 $K(P_4), K(P_4) \sqsubset K(P_1).$

#### 4.2. Information distance between information structures

For  $A, B \subseteq U$ , denote

$$A \oplus B = A \cup B - A \cap B.$$

Then  $A \oplus B$  is called the symmetric difference  $A$  and  $B$ .

Obviously,  $|A \oplus B| = |A \cup B| - |A \cap B|.$

In rough set theory, information entropy and information granulation are two main approaches to measuring the uncertainty of information structures in an information system. If the information granulation (or information entropy) of an information structure is equal to that of the other information structure in the same information system, we say that these two information structures have the same uncertainty. However, it does not mean that these two information structures are equivalent each other. Thus, information entropy and information granulation cannot characterize the difference between any two information structures in an information system.

To differentiate two information structures in the same incomplete information system, the concept of information distance is defined as follows.

**Definition 4.4** Let  $(U, A)$  be an incomplete information system. Given that  $K(P), K(Q)$  are the information structures of  $P, Q \subseteq A$ , respectively. Information distance between  $K(P)$  and  $K(Q)$  is defined as

$$\rho(K(P), K(Q)) = \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus S_Q(u_i)|.$$

**Lemma 4.5** Let  $A, B \subseteq U$ . Then

$$A = B \iff |A \oplus B| = 0.$$

**Proof.** “ $\implies$ ”. This is obvious.

“ $\impliedby$ ”. Suppose  $|A \oplus B| = 0$ . Then  $|A \cup B - A \cap B| = 0.$

It follows that  $A \cup B - A \cap B = \emptyset$ . So  $A \cup B \subseteq A \cap B$ .

Since  $A \cap B \subseteq A \cup B$ , we have  $A \cap B = A \cup B$ .  
Thus  $A = B$ . □

**Lemma 4.6** Let  $A, B, C \subseteq U$ . Then

$$|A \oplus B| + |B \oplus C| \geq |A \oplus C|.$$

**Proof.** Denote

$$\begin{aligned} A_1 &= A - (A \cap B + A \cap C - A \cap B \cap C), \\ A_2 &= A \cap B - A \cap B \cap C, \\ A_3 &= B - (A \cap B + B \cap C - A \cap B \cap C), \\ A_4 &= A \cap C - A \cap B \cap C, \\ A_5 &= A \cap B \cap C, \\ A_6 &= B \cap C - A \cap B \cap C, \\ A_7 &= C - (A \cap C + B \cap C - A \cap B \cap C). \end{aligned}$$

Then

$$\begin{aligned} |A \oplus B| &= |A_1| + |A_4| + |A_3| + |A_6|, \\ |B \oplus C| &= |A_2| + |A_3| + |A_4| + |A_7|, \\ |A \oplus C| &= |A_1| + |A_2| + |A_6| + |A_7|. \end{aligned}$$

So  $|A \oplus B| + |B \oplus C| - |A \oplus C| = 2(|A_3| + |A_4|) > 0$ .

Thus  $|A \oplus B| + |B \oplus C| \geq |A \oplus C|$ . □

**Lemma 4.7** Let  $A, B, C \in 2^U$ . If  $A \subseteq B \subseteq C$  or  $C \subseteq B \subseteq A$ , then

$$|A \oplus B| + |B \oplus C| = |A \oplus C|.$$

**Proof.** Suppose  $A \subseteq B \subseteq C$ . Then

$$\begin{aligned} |A \oplus B| &= |A \cup B| - |A \cap B| = |B| - |A|, \\ |B \oplus C| &= |B \cup C| - |B \cap C| = |C| - |B|, \\ |A \oplus C| &= |A \cup C| - |A \cap C| = |C| - |A|. \end{aligned}$$

Thus

$$|A \oplus B| + |B \oplus C| = |A \oplus C|.$$

Suppose  $C \subseteq B \subseteq A$ . Similarly, we can prove that

$$|A \oplus B| + |B \oplus C| = |A \oplus C|.$$

□

**Theorem 4.8** Let  $(U, A)$  be an incomplete information system. Then  $(\mathbf{K}(U, A), \rho)$  is a distance space.

**Proof.** Suppose  $P, Q, L \subseteq A$ .

Obviously,

$$\begin{aligned} \rho(K(P), K(Q)) &\geq 0, \\ \rho(K(P), K(Q)) &= \rho(K(Q), K(P)). \end{aligned}$$

By Lemma 4.5,

$$\begin{aligned} \rho(K(P), K(Q)) = 0 &\Leftrightarrow \forall i, |S_P(u_i) \oplus S_Q(u_i)| = 0 \\ &\Leftrightarrow \forall i, S_P(u_i) = S_Q(u_i) \Leftrightarrow K(P) = K(Q). \end{aligned}$$

By Lemma 4.6,

$$|S_P(u_i) \oplus S_Q(u_i)| + |S_Q(u_i) \oplus S_L(u_i)| \geq |S_P(u_i) \oplus S_L(u_i)|.$$

Then  $\rho(K(P), K(Q)) + \rho(K(Q), K(L))$

$$\begin{aligned} &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus S_Q(u_i)| + \frac{1}{n^2} \sum_{i=1}^n |S_Q(u_i) \oplus S_L(u_i)| \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i) \oplus S_Q(u_i)| + |S_Q(u_i) \oplus S_L(u_i)|) \\ &\geq \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus S_L(u_i)| \\ &= \rho(K(P), K(L)) \end{aligned}$$

Thus  $(\mathbf{K}(U, A), \rho)$  is a distance space. □

The following theorem illustrates the fact that information entropy in information system can be expressed by information distance between two information structures.

**Theorem 4.9** Let  $(U, A)$  be an incomplete information system. Then for  $P \subseteq A$ ,  $\rho(K(P), K(\emptyset))$  is the information entropy of  $P$ .

**Proof.** By Definition 11 in <sup>7</sup>, the information entropy of  $P$  is defined as:

$$E(P) = \sum_{i=1}^n \frac{1}{n} \left(1 - \frac{S_P(u_i)}{n}\right).$$

But  $\rho(K(P), K(\emptyset)) = \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus U|$

$$\begin{aligned} &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i) \cup U| - |S_P(u_i) \cap U|) \\ &= \frac{1}{n^2} \sum_{i=1}^n (n - |S_P(u_i)|). \end{aligned}$$

Thus

$$\rho(K(P), K(\emptyset)) = E(P).$$

□

**Proposition 4.10** Let  $(U, A)$  be an incomplete information system. Then for  $P, Q \subseteq A$ ,

(1)

$$0 \leq \rho(K(P), K(Q)) \leq 1 - \frac{1}{n};$$

(2) If  $K(P) \preceq K(Q)$  and  $\text{sim}(P^*)(P^* \subseteq A)$  is an identity relation on  $U$ , then

$$\rho(K(P), K(P^*)) \leq \rho(K(Q), K(P^*));$$

(3) If  $K(P) \preceq K(Q)$ , then

$$\rho(K(P), K(\emptyset)) \geq \rho(K(Q), K(\emptyset)).$$

**Proof.** (1) Obviously,

$1 \leq |S_P(u_i) \cup S_Q(u_i)| \leq n$  and  $1 \leq |S_P(u_i) \cap S_Q(u_i)| \leq n$  ( $i = 1, 2, \dots, n$ ). Then

$$0 \leq |S_P(u_i) \oplus S_Q(u_i)| \leq n - 1 \quad (i = 1, 2, \dots, n).$$

Thus

$$0 \leq \rho(K(P), K(Q)) \leq \frac{1}{n^2} \sum_{i=1}^n (n-1) = \frac{n^2 - n}{n^2} = 1 - \frac{1}{n}.$$

(2) Since  $K(P) \preceq K(Q)$ ,  $\forall i$ , we have  $S_P(u_i) \subseteq S_Q(u_i)$ . Then,  $\forall i$ ,  $|S_P(u_i)| \leq |S_Q(u_i)|$ . Thus

$$\begin{aligned} \rho(K(P), K(P^*)) &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus \{u_i\}| \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i) \cup \{u_i\}| - |S_P(u_i) \cap \{u_i\}|) \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i)| - 1) \leq \frac{1}{n^2} \sum_{i=1}^n (|S_Q(u_i)| - 1) \\ &= \rho(K(Q), K(P^*)). \end{aligned}$$

(2) It should be noted that  $K(P) \preceq K(Q)$ . Then,  $\forall i$ ,  $|S_P(u_i)| \leq |S_Q(u_i)|$ . Thus

$$\begin{aligned} \rho(K(P), K(\emptyset)) &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus U| \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i) \cup U| - |S_P(u_i) \cap U|) \\ &= \frac{1}{n^2} \sum_{i=1}^n (n - |S_P(u_i)|) \geq \frac{1}{n^2} \sum_{i=1}^n (n - |S_Q(u_i)|) \\ &= \rho(K(Q), K(\emptyset)). \quad \square \end{aligned}$$

**Proposition 4.11** Let  $(U, A)$  be an incomplete information system. If  $\text{sim}(P^*)(P^* \subseteq A)$  is an identity relation on  $U$ , then for  $P \subseteq A$ ,

$$\rho(K(P), K(P^*)) + \rho(K(P), K(\emptyset)) = 1 - \frac{1}{n}.$$

**Proof.** We can obtain that

$$\begin{aligned} &\rho(K(P), K(P^*)) + \rho(K(P), K(\emptyset)) \\ &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus \{u_i\}| + \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus U| \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i)| - 1) + \frac{1}{n^2} \sum_{i=1}^n (n - |S_P(u_i)|) \\ &= \frac{1}{n^2} \sum_{i=1}^n (n - 1) = 1 - \frac{1}{n}. \quad \square \end{aligned}$$

**Proposition 4.12** Let  $(U, A)$  be an incomplete information system. Given  $P, Q, L \subseteq A$ . If  $K(P) \preceq K(Q) \preceq K(L)$  or  $K(L) \preceq K(Q) \preceq K(P)$ , then

$$\rho(K(P), K(Q)) + \rho(K(Q), K(L)) = \rho(K(P), K(L)).$$

**Proof.** Since  $K(P) \preceq K(Q) \preceq K(L)$  or  $K(L) \preceq K(Q) \preceq K(P)$ , we have

$$S_P(u_i) \subseteq S_Q(u_i) \subseteq S_L(u_i) \text{ or } S_L(u_i) \subseteq S_Q(u_i) \subseteq S_P(u_i) \quad (i = 1, 2, \dots, n).$$

By Lemma 4.7,

$$|S_P(u_i) \oplus S_Q(u_i)| + |S_Q(u_i) \oplus S_L(u_i)| = |S_P(u_i) \oplus S_L(u_i)| \quad (i = 1, 2, \dots, n).$$

$$\begin{aligned} &\rho(K(P), K(Q)) + \rho(K(Q), K(L)) \\ &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus S_Q(u_i)| + \frac{1}{n^2} \sum_{i=1}^n |S_Q(u_i) \oplus S_L(u_i)| \\ &= \frac{1}{n^2} \sum_{i=1}^n (|S_P(u_i) \oplus S_Q(u_i)| + |S_Q(u_i) \oplus S_L(u_i)|) \\ &= \frac{1}{n^2} \sum_{i=1}^n |S_P(u_i) \oplus S_L(u_i)| \\ &= \rho(K(P), K(L)). \quad \square \end{aligned}$$

**Example 4.13** (Continued from Example 4.3) By Definition 4.4, one can obtain that

$$\rho(K(P_1), K(\emptyset)) = \frac{10}{16}, \quad \rho(K(P_2), K(\emptyset)) = \frac{8}{16},$$

$$\rho(K(P_3), K(\emptyset)) = \frac{8}{16}, \quad \rho(K(P_5), K(\emptyset)) = \frac{12}{16},$$

$$\rho(K(P_1), K(P_2)) = \frac{2}{16}, \quad \rho(K(P_1), K(P_3)) = \frac{6}{16},$$

$$\rho(K(P_1), K(P_5)) = \frac{2}{16}, \quad \rho(K(P_2), K(P_3)) = \frac{8}{16},$$

$$\rho(K(P_2), K(P_5)) = \frac{4}{16}, \quad \rho(K(P_3), K(P_5)) = \frac{4}{16}.$$

This example illustrates the following facts:

(1) It should be noted that  $K(P_1) = K(P_4)$ ,  $K(P_5) = K(P_6) = K(A)$ . Then

$$E(P_1) = E(P_4) = \rho(K(P_1), K(\emptyset)) = \frac{10}{16},$$

$$E(P_2) = \rho(K(P_2), K(\emptyset)) = \frac{8}{16},$$

$$E(P_3) = \rho(K(P_3), K(\emptyset)) = \frac{8}{16},$$

$$E(P_5) = E(P_6) = E(A) = \rho(K(P_5), K(\emptyset)) = \frac{12}{16}.$$

(2) We have  $K(P_1) \prec K(P_2)$  and  $\text{ind}(P_5) = \Delta$ . It is clear that

$$\rho(K(P_1), K(P_5)) = \frac{2}{16} \leq \frac{4}{16} = \rho(K(P_2), K(P_5)),$$

$$\rho(K(P_1), K(\emptyset)) = \frac{10}{16} \geq \frac{8}{16} = \rho(K(P_2), K(\emptyset)).$$

(3) We have  $\text{ind}(P_5) = \Delta$ . It is clear that for  $P \subseteq A$ ,

$$\rho(K(P), K(P_5)) + \rho(K(P), K(\emptyset)) = 1 - \frac{1}{4} = 1 - \frac{1}{n}.$$

(4) We have  $K(P_5) \prec K(P_1) \prec K(P_2) \prec K(\emptyset)$ . It is clear that

$$\rho(K(P_5), K(P_1)) + \rho(K(P_1), K(P_2)) = \frac{4}{16} = \rho(K(P_5), K(P_2)),$$

$$\rho(K(P_1), K(P_2)) + \rho(K(P_2), K(\emptyset)) = \frac{10}{16} = \rho(K(P_1), K(\emptyset)).$$

## 5. Properties of information structures in incomplete information systems

**Theorem 5.1** Let  $(U, A)$  be an incomplete information system. Then for  $P, Q \subseteq A$ , the following are equivalent:

- (1)  $K(P) = K(Q)$ ;
- (2)  $U/P = U/Q$ ;
- (3)  $\text{sim}(P) = \text{sim}(Q)$ ;
- (4)  $\rho(K(P), K(Q)) = 0$

**Proof.** (1)  $\Leftrightarrow$  (2)  $\Leftrightarrow$  (3) are obvious.

(1)  $\Leftrightarrow$  (4) holds by Theorem 3.9.  $\square$

**Definition 5.2** <sup>(21)</sup> Let  $\mathcal{A} = \{X_1, X_2, \dots, X_k\}$  and  $\mathcal{B} = \{Y_1, Y_2, \dots, Y_l\}$  be two partitions on  $U$ .

(1)

$$H(\mathcal{A}) = \sum_{i=1}^k p(X_i)p(U - X_i)$$

is called the information amount of  $\mathcal{A}$ , where  $p(X_i) = \frac{|X_i|}{|U|}$  means the probability that the element of  $U$  belongs to  $X_i$ .

(2)

$$H(\mathcal{B}/\mathcal{A}) = \sum_{i=1}^k \sum_{j=1}^l p(X_i \cap Y_j)p(X_i - Y_j)$$

is called the condition information amount of  $\mathcal{B}$  with respect to  $\mathcal{A}$ .

The following theorem quantitatively depict the dependence of information structures by the condition information amount.

**Theorem 5.3** Let  $(U, A)$  be an incomplete information system. Then for  $P, Q \subseteq A$ , the following are equivalent:

- (1)  $K(P) \preceq K(Q)$ ;
- (2)  $U/P$  refines  $U/Q$ , i.e., for each  $X \in U/P$ , there exists  $Y \in U/Q$  such that  $X \subseteq Y$ ;
- (3)  $\text{sim}(P) \subseteq \text{sim}(Q)$ ;
- (4)  $H((U/Q)/(U/P)) = 0$ .

**Proof.** (1)  $\Leftrightarrow$  (2)  $\Leftrightarrow$  (3) are obvious.

(3)  $\Rightarrow$  (4). Denote

$$U/P = \{X_1, X_2, \dots, X_k\}, \quad U/Q = \{Y_1, Y_2, \dots, Y_l\}.$$

$\forall i$ , since  $U/P$  refines  $U/Q$ , we have  $X_i \subseteq Y_{j_i}$  for some  $j_i \leq l$ . Then  $X_i - Y_{j_i} = \emptyset$ .

$\forall j \neq j_i, Y_j \cap Y_{j_i} = \emptyset$ . Then  $X_i \cap Y_j = \emptyset$ .

Thus  $\forall i, j, X_i - Y_j = \emptyset$  or  $X_i \cap Y_j = \emptyset$ . This implies

$$p(X_i \cap Y_j)p(X_i - Y_j) = 0.$$

Hence  $H((U/Q)/(U/P)) = 0$ .

(4)  $\Rightarrow$  (2). Since  $H((U/Q)/(U/P)) = 0$ , we have  $\forall i, j, X_i - Y_j = \emptyset$  or  $X_i \cap Y_j = \emptyset$ . So  $\forall i, j, X_i \subseteq Y_j$  or  $X_i \cap Y_j = \emptyset$ .

$\forall i, X_i = \bigcup_{j=1}^l (X_i \cap Y_j)$ . Since  $X_i \neq \emptyset$ , we have

$X_i \cap Y_{j_i} \neq \emptyset$  for some  $j_i \leq l$ . Then  $X_i \subseteq Y_{j_i}$ .

Thus  $U/P$  refines  $U/Q$ .  $\square$

**Proposition 5.4** Let  $(U, A)$  be an incomplete information system. Given  $P, Q \subseteq A$ . If  $P \subseteq Q$ , then  $K(Q) \preceq K(P)$ .

**Proof.** Since  $P \subseteq Q$ , we have  $sim(Q) \subseteq sim(P)$ . By Theorem 5.3,  $K(Q) \preceq K(P)$ . □

**Proposition 5.5** Let  $(U, A)$  be an incomplete information system. Given  $P \subseteq A$ . Then  $K(A) \preceq K(P) \preceq K(\emptyset)$ .

**Proof.** This holds by Proposition 5.4. □

Let  $(U, A)$  be an incomplete information system. For  $P \subseteq A$ , denote

$$\sigma(U/P) = \left\{ \bigcup_{u \in X} S_P(u) : X \in 2^U \right\}.$$

**Theorem 5.6** Let  $(U, A)$  be an incomplete information system and  $P, Q \subseteq A$ . Denote

$$U/Q = \{D_1, D_2, \dots, D_r\}.$$

Then the following are equivalent:

- (1)  $K(P) \preceq K(Q)$ ;
- (2) for each  $j$ ,  $D_j \in \sigma(U/P)$ ;
- (3) for each  $j$ ,  $\underline{P}(D_j) = D_j$ ;
- (4)  $\bigcup_{j=1}^r \underline{P}(D_j) = U$ ;

**Proof.** (1)  $\Rightarrow$  (2).  $\forall u \in D_j$ , we have  $D_j = S_Q(u)$ . Since  $K(P) \preceq K(Q)$ ,  $S_P(u) \subseteq S_Q(u)$ . Then  $\{u\} \subseteq S_P(u) \subseteq D_j$ . So

$$D_j = \bigcup_{u \in D_j} \{u\} \subseteq \bigcup_{u \in D_j} S_P(u) \subseteq D_j.$$

Thus  $D_j = \bigcup_{u \in D_j} S_P(u) \in \sigma(U/P)$ .

(2)  $\Rightarrow$  (3). Since  $D_j \in \sigma(U/R)$ ,  $D_j = \bigcup_{u \in X} S_P(u)$

for some  $X \in 2^U$ .

Thus  $\underline{P}(D_j) = \underline{P}(\bigcup_{u \in X} S_P(u)) = \bigcup_{u \in X} S_P(u) = D_j$ .

(3)  $\Rightarrow$  (4). This is obvious.

(4)  $\Rightarrow$  (1).  $\forall i \leq n$ ,  $\forall v \in S_P(u_i)$ . Since

$\bigcup_{j=1}^r \underline{P}(D_j) = U$ ,  $u_i \in \underline{P}(D_j)$  for some  $j \leq r$ . Then  $S_P(u_i) \subseteq D_j$ . Denote  $D_j = S_Q(u^*)$ .  $u_i \in D_j$  implies

$S_Q(u_i) = S_Q(u^*)$ . Then  $v \in S_Q(u_i)$ . So  $S_P(u_i) \subseteq S_Q(u_i)$ . This shows  $K(P) \preceq K(Q)$ . □

**Definition 5.7** <sup>(21)</sup> Let  $(U, A)$  be an incomplete information system. Given that  $\mathbf{K}(U, A)$  is the information structure base of  $(U, A)$ . Then a mapping  $D : \mathbf{K}(U, A) \times \mathbf{K}(U, A) \rightarrow [0, 1]$  is called the inclusion degree on  $\mathbf{K}(U, A)$ , if

- (1)  $0 \leq D(K(Q)/K(P)) \leq 1$ ;
- (2)  $K(P) \preceq K(Q)$  implies  $D(K(Q)/K(P)) = 1$ ;
- (3)  $K(P) \sqsubseteq K(Q) \sqsubseteq K(L)$  implies

$D(K(P)/K(L)) \leq D(K(P)/K(Q))$ .

**Definition 5.8** Let  $(U, A)$  be an incomplete information system. For any  $P, Q \subseteq A$ , define

$$D(K(Q)/K(P)) = \sum_{i=1}^n \frac{|S_Q(u_i)|}{\sum_{i=1}^n |S_Q(u_i)|} \chi_{S_Q(u_i)}(S_P(u_i)),$$

where

$$\chi_{S_Q(u_i)}(S_P(u_i)) = \begin{cases} 1, & \text{if } S_P(u_i) \subseteq S_Q(u_i), \\ 0, & \text{if } S_P(u_i) \not\subseteq S_Q(u_i). \end{cases}$$

**Proposition 5.9**  $D$  in Definition 5.8 is the inclusion degree under Definition 5.7.

**Proof.** This is obvious. □

**Example 5.10** (Continued from Example 4.3) We have

$$\mathbf{K}(U, A) = \{K(\emptyset), K(P_1), K(P_3), K(P_4), K(P_{10}), K(A)\}.$$

Pick  $P = P_1$ ,  $Q = P_4$ .

$$K(P) = (S_P(u_1), S_P(u_2), S_P(u_3), S_P(u_4), S_P(u_5), S_P(u_6))$$

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

$$K(Q) = (S_Q(u_1), S_Q(u_2), S_Q(u_3), S_Q(u_4), S_Q(u_5), S_Q(u_6))$$

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

It is easy to verify that

$$D(K(P)/K(P)) = D(K(Q)/K(Q)) = 1.$$

We have

$$D(K(Q)/K(P)) = \frac{5}{14}, \quad D(K(P)/K(Q)) = \frac{2}{10}.$$

This example illustrates that

$$D(K(Q)/K(P)) + D(K(P)/K(Q)) \neq 1.$$

The following theorem show that the dependency between information structures in an incomplete information system can be quantitatively described by the inclusion degree.

**Theorem 5.11** *Let  $(U, A)$  be an incomplete information system. Given  $P, Q \subseteq A$ . Then*

- (1)  $K(P) \preceq K(Q) \iff D(K(Q)/K(P)) = 1$ .
- (2)  $K(P) \bowtie K(Q) \iff D(K(Q)/K(P)) = 0$ .
- (3)  $K(P) \sqsubseteq K(Q) \iff 0 < D(K(Q)/K(P)) \leq 1$ .

**Proof.** (1) “ $\implies$ ” is obvious.

“ $\impliedby$ ”. Put

$$|S_Q(u_l)| = q_l, \quad \sum_{l=1}^n |S_Q(u_l)| = q.$$

Then  $q = \sum_{l=1}^n q_l$ . Since  $D(K(Q)/K(P)) = 1$ , we have

$$\sum_{l=1}^n q_l \chi_{S_Q(u_l)}(S_P(u_l)) = q = \sum_{l=1}^n q_l.$$

Then

$$\sum_{l=1}^n q_l (1 - \chi_{S_Q(u_l)}(S_P(u_l))) = 0.$$

Thus  $\forall l$ ,

$$1 - \chi_{S_Q(u_l)}(S_P(u_l)) = 0.$$

It follows that  $\forall l, S_P(u_l) \subseteq S_Q(u_l)$ .

Hence  $K(P) \preceq K(Q)$ .

(2) “ $\implies$ ”. Since  $K(P) \bowtie K(Q)$ , we have  $\forall l, S_P(u_l) \not\subseteq S_Q(u_l)$ . Then  $\forall l$ ,

$$\chi_{S_Q(u_l)}(S_P(u_l)) = 0.$$

Thus  $D(K(Q)/K(P)) = 0$ .

“ $\impliedby$ ”. Since  $D(K(Q)/K(P)) = 0$ , we obtain that  $\forall l$ ,

$$\chi_{S_Q(u_l)}(S_P(u_l)) = 0.$$

Then  $\forall l, S_P(u_l) \not\subseteq S_Q(u_l)$ . Thus  $K(P) \bowtie K(Q)$ .

(3) This holds by (2).  $\square$

## 6. Characterizations of information structures in incomplete information systems

In this section, group, and lattice characterizations of information structures in incomplete information systems are given.

### 6.1. Group characterizations of information structures

**Definition 6.1** <sup>(2)</sup> *Let  $S$  be a non-empty set and “ $\cdot$ ” a binary operation on  $S$ .*

(1)  $(S, \cdot)$  is called a semigroup, if for any  $a, b \in S, a \cdot b \in S$  and for any  $a, b, c \in S, (a \cdot b) \cdot c = a \cdot (b \cdot c)$ .

(2)  $(S, \cdot)$  is called an exchangeable semigroup, if it is a semigroup and for any  $a, b \in S, a \cdot b = b \cdot a$ .

(3)  $e \in S$  is called the identity element of  $S$ , if for any  $a \in S, e \cdot a = a \cdot e = a$ .

(4)  $(S, \cdot)$  is called a group, if it is a semigroup and every element has an inverse element.

**Theorem 6.2**  $(\mathbf{K}(U, A), \odot)$  is a exchangeable semigroup with the identity element  $K(\emptyset)$ .

**Proof.** Suppose  $P, Q, L \subseteq A$ . Then

$$\begin{aligned} K(P) \odot K(Q) &= (S_{sim(P)}(u_1) \cap S_{sim(Q)}(u_1), S_{sim(P)}(u_2) \cap S_{sim(Q)}(u_2), \dots, S_{sim(P)}(u_n) \cap S_{sim(Q)}(u_n)) \\ &= (S_{sim(P) \cap sim(Q)}(u_1), S_{sim(P) \cap sim(Q)}(u_2), \dots, S_{sim(P) \cap sim(Q)}(u_n)) \\ &= (S_{sim(P \cup Q)}(u_1), S_{sim(P \cup Q)}(u_2), \dots, S_{sim(P \cup Q)}(u_n)) \\ &= K(P \cup Q). \end{aligned}$$

Similarly,  $K(Q) \odot K(P) = K(Q \cup P)$ .

Thus

$$K(P) \odot K(Q) = K(Q) \odot K(P).$$

Since  $(K(P) \odot K(Q)) \odot K(L) = K(P \cup Q) \odot K(L) = K(P \cup Q \cup L)$ ,

$K(P) \odot (K(Q) \odot K(L)) = K(P) \odot K(Q \cup L) = K(P \cup Q \cup L)$ ,

we have

$$(K(P) \odot K(Q)) \odot K(L) = K(P) \odot (K(Q) \odot K(L)).$$

Hence  $(\mathbf{K}(U), \odot)$  is a exchangeable semigroup. Obviously,  $K(\emptyset)$  is the identity element.  $\square$

**Example 6.3**  $(\mathbf{K}(U, A), \odot)$  is not a group.

Consider Example 4.3. By Theorem 6.2,  $(\mathbf{K}(U, A), \odot)$  is a exchangeable semigroup with the identity element  $K(\emptyset)$ .

It should be noted that  $\forall P \subseteq A$ ,

$$K(P) \odot K(A) = K(P \cup A) = K(A) \neq K(\emptyset).$$

Then  $K(A)$  has no inverse elements.

Thus  $(\mathbf{K}(U, A), \odot)$  is not a group.

## 6.2. Lattice characterizations of information structures

**Theorem 6.4** Let  $(U, A)$  be an incomplete information system. Then

(1)  $L = (\mathbf{K}(U, A), \preceq)$  is a lattice with  $1_L = K(\emptyset)$  and  $0_L = K(A)$ ;

(2) If  $P_1, P_2 \subseteq A$  and  $\mathcal{A}(P_1, P_2) = \{P \mid P \subseteq A, \text{sim}(P_1) \cup \text{sim}(P_2) \subseteq \text{sim}(P)\}$ , then

$$K(P_1) \wedge K(P_2) = K(P_1) \odot K(P_2) = K(P_1 \cup P_2),$$

$$K(P_1) \vee K(P_2) = \bigodot_{P \in \mathcal{A}(P_1, P_2)} K(P) = K(\cup \mathcal{A}(P_1, P_2)).$$

**Proof.** Obviously,  $\forall P \subseteq A, K(P) \preceq K(P)$ .

Suppose that  $K(P) \preceq K(Q), K(Q) \preceq K(P)$ . By Theorem 5.3,  $\text{sim}(P) \subseteq \text{sim}(Q), \text{sim}(Q) \subseteq \text{sim}(P)$ . Then  $\text{sim}(P) = \text{sim}(Q)$ . Thus  $K(P) = K(Q)$ .

Suppose that  $K(P) \preceq K(Q), K(Q) \preceq K(L)$ . By Theorem 5.3,  $\text{sim}(P) \subseteq \text{sim}(Q), \text{sim}(Q) \subseteq \text{sim}(L)$ . Then  $\text{sim}(P) \subseteq \text{sim}(L)$ .

By Theorem 5.3,  $K(P) \preceq K(L)$ .

Hence  $(\mathbf{K}(U, A), \preceq)$  is a partial order set.

By the proof of Theorem 5.1, we have

$$K(P_1) \odot K(P_2) = K(P_1 \cup P_2),$$

$$\bigodot_{P \in \mathcal{A}(P_1, P_2)} K(P) = K(\cup \mathcal{A}(P_1, P_2)).$$

Since  $P_1, P_2 \subseteq P_1 \cup P_2$ , by Proposition 5.4, we have  $K(P_1 \cup P_2) \preceq K(P_1)$  and  $K(P_1 \cup P_2) \preceq K(P_2)$ . Then  $K(P_1 \cup P_2)$  is the lower bound of  $\{K(P_1), K(P_2)\}$ .

Suppose  $K(P)$  is the lower bound of  $\{K(P_1), K(P_2)\}$  with  $P \subseteq A$ . Then  $K(P) \preceq K(P_1), K(P) \preceq K(P_2)$ . By Theorem 5.3, we have  $\text{sim}(P) \subseteq \text{sim}(P_1)$  and  $\text{sim}(P) \subseteq \text{sim}(P_2)$ . Then  $\text{sim}(P) \subseteq \text{sim}(P_1) \cap \text{sim}(P_2) = \text{sim}(P_1 \cup P_2)$ . By Theorem 5.3,  $K(P) \preceq K(P_1 \cup P_2)$ .

Thus

$$K(P_1) \wedge K(P_2) = K(P_1 \cup P_2).$$

$\forall P \in \mathcal{A}(P_1, P_2)$ , since  $\text{sim}(P_1) \cup \text{sim}(P_2) \subseteq \text{sim}(P)$ , we have

$$\text{sim}(P_1), \text{sim}(P_2) \subseteq \text{sim}(P).$$

Then

$$\text{sim}(P_1), \text{sim}(P_2) \subseteq \bigcap_{P \in \mathcal{A}(P_1, P_2)} \text{sim}(P) = \text{sim}\left(\bigcup_{P \in \mathcal{A}(P_1, P_2)} P\right) = \text{sim}(\cup \mathcal{A}(P_1, P_2)).$$

By Theorem 5.3,  $K(P_1), K(P_2) \preceq K(\cup \mathcal{A}(P_1, P_2))$ . Then  $K(\cup \mathcal{A}(P_1, P_2))$  is the upper bound of  $\{K(P_1), K(P_2)\}$ .

Suppose  $K(S)$  is the upper bound of  $\{K(P_1), K(P_2)\}$  with  $S \subseteq A$ . Then  $K(P_1) \preceq K(S), K(P_2) \preceq K(S)$ .

By Theorem 5.3,  $\text{sim}(P_1) \subseteq \text{sim}(S)$  and  $\text{sim}(P_2) \subseteq \text{sim}(S)$ .

Then  $S \in \mathcal{A}(P_1, P_2)$ . So  $S \subseteq \cup \mathcal{A}(P_1, P_2)$ .

By Proposition 5.4,  $K(\cup \mathcal{A}(P_1, P_2)) \preceq K(S)$ .

Thus  $K(P_1) \vee K(P_2) = K(\cup \mathcal{A}(P_1, P_2))$ .

Hence  $L$  is a lattice.

Obviously,  $1_L = K(\emptyset), 0_L = K(A)$ . □

**Corollary 6.5**  $(\mathbf{K}(U, A), \wedge)$  is a exchangeable semigroup with the identity element  $K(\emptyset)$ .

**Example 6.6**  $L = (\mathbf{K}(U, A), \preceq)$  is not a distributive lattice.

Consider Example 4.3.

By Theorem 6.4,  $L = (\mathbf{K}(U, A), \preceq)$  is a lattice with  $1_L = K(\emptyset)$  and  $0_L = K(A)$ .

We have

$$\mathcal{A}(P_2, P_1 \cup P_4) = \mathcal{A}(P_2, P_2) = \{P \subseteq A : \text{sim}(P_2) \cup \text{sim}(P_2) \subseteq \text{sim}(P)\} = \{P \subseteq A : \text{sim}(P_2) \subseteq \text{sim}(P)\} = 2^A,$$

$$\mathcal{A}(P_2, P_1) = \{P \subseteq A : \text{sim}(P_2) \cup \text{sim}(P_1) \subseteq \text{sim}(P)\} = \{P \subseteq A : \text{sim}(P_1) \subseteq \text{sim}(P)\} = \{\emptyset, P_1, P_3, P_6\},$$

$$\mathcal{A}(P_2, P_4) = \{P \subseteq A : \text{sim}(P_2) \cup \text{sim}(P_4) \subseteq \text{sim}(P)\} = \{P \subseteq A : \text{sim}(P_4) \subseteq \text{sim}(P)\} = \{\emptyset, P_4\}.$$

Then

$$\cup \mathcal{A}(P_2, P_1 \cup P_4) = \{a_1, a_2, a_3, a_4\},$$

$$(\cup \mathcal{A}(P_2, P_1)) \cup (\cup \mathcal{A}(P_2, P_4)) = \{a_1, a_3, a_4\}.$$

By Theorem 5.1,

$$K(\cup \mathcal{A}(P_2, P_1 \cup P_4)) \neq K((\cup \mathcal{A}(P_2, P_1)) \cup (\cup \mathcal{A}(P_2, P_4))).$$

By Theorem 6.4,

$$\begin{aligned} K(P_2) \vee (K(P_1) \wedge K(P_4)) &= K(P_2) \vee K(P_1 \cup P_4) \\ &= K((\cup \mathcal{A}(P_2, P_1 \cup P_4))), \end{aligned}$$

$$\begin{aligned} (K(P_2) \vee K(P_1)) \wedge (K(P_2) \vee K(P_4)) \\ &= K((\cup \mathcal{A}(P_2, P_1)) \cup (\cup \mathcal{A}(P_2, P_4))) \\ &= K((\cup \mathcal{A}(P_2, P_1)) \cup (\cup \mathcal{A}(P_2, P_4))). \end{aligned}$$

Thus

$$\begin{aligned} K(P_2) \vee (K(P_1) \wedge K(P_4)) \\ \neq (K(P_2) \vee K(P_1)) \wedge (K(P_2) \vee K(P_4)). \end{aligned}$$

Therefore,  $(\mathbf{K}(U, A), \preceq)$  is not a distributive lattice.

**Example 6.7**  $(\mathbf{K}(U, A), \sqsubseteq)$  is not a partial order set.

Consider Example 4.3. We have  $K(P_1) \sqsubseteq K(P_2)$ ,  $K(P_2) \sqsubseteq K(P_1)$ . But  $K(P_1) \neq K(P_2)$ . This implies that  $\sqsubseteq$  is not anti-symmetric.

Therefore,  $L = (\mathbf{K}(U, A), \preceq)$  is not a distributive lattice.

**Example 6.8**  $(\mathbf{K}(U, A), \sqsubseteq)$  is not a partial order set.

Consider Example 4.3. We have  $K(P) \sqsubseteq K(Q)$ ,  $K(Q) \sqsubseteq K(P)$ . But  $K(P) \neq K(Q)$ . This implies that  $\sqsubseteq$  is not anti-symmetric.

Hence  $(\mathbf{K}(U, A), \sqsubseteq)$  is not a partial order set.

## 7. Conclusions

In this paper, information structures in an incomplete information system have been considered as a special case of set vectors. Based on this consideration, relationships between information structures have been discussed from two aspects of dependence and separability. Properties of information structures are given. Group and lattice characterizations of information structures have been obtained. These results will be significant for establishing a

framework of granular computing in a information system. This vector-based framework can be used to represent different dimension of information in an incomplete information system and may have potential applications to knowledge discovery in incomplete information systems. In the future, we will consider the fuzzification of the proposed results and give some applications such as dealing with knowledge discovery in incomplete information systems.

## Acknowledgements

The authors would like to thank the editors and the anonymous reviewers for their valuable comments and suggestions which have helped immensely in improving the quality of the paper. This work is supported by National Natural Science Foundation of China (11461005), Natural Science Foundation of Guangxi (2016GXNSFAA380045, 2016GXNSFAA380282), Key Laboratory of Optimization Control and Engineering Calculation in Department of Guangxi Education, and Special Funds of Guangxi Distinguished Experts Construction Engineering.

## References

1. N.Chen, B.Qin, "Invariant characterizations of information structures in a lattice-valued information system under homomorphisms based on data compression," *Journal of Intelligent & Fuzzy Systems*, **33**, 3987-3998 (2017).
2. T.W.Hungerford, "Algebra", Springer-Verlag Press, New York, 1974.
3. T.Y.Lin, "Granular computing on binary relations I: data mining and neighborhood systems," In: *Rough Sets In Knowledge Discovery*, A.Skowron and L.Polkowski (eds), Physica-Verlag, pp. 107-121, 1998.
4. T.Y.Lin, "Granular computing on binary relations I-I: rough set representations and belief functions," In: *Rough Sets In Knowledge Discovery*, A.Skowron and L.Polkowski (eds), Physica-Verlag, pp. 121-140, 1998.
5. Z.Li, Y.Liu, Q.Li, B.Qin, "Relationships between knowledge bases and related results," *Knowledge and Information Systems*, **49**, 171-195 (2016).
6. Z.Li, Q.Li, R.Zhang, N.Xie, "Knowledge structures in a knowledge base," *Expert Systems*, **33**, 581-591 (2016).
7. J.Liang, Y.Qian, "Information granules and entropy theory," *Science in China Series F: Information Sciences*, **51**, 1427-1444 (2008).
8. Z.Pawlak, "Rough sets: Theoretical aspects of reason-

- ing about data”, Kluwer Academic Publishers, Dordrecht, 1991.
9. Z.Pawlak, A.Skowron, “Rudiments of rough sets,” *Information Sciences*, **177**, 3-27 (2007).
  10. Z.Pawlak, A.Skowron, “Rough sets: some extensions,” *Information Sciences*, **177**, 28-40 (2007).
  11. Z.Pawlak, A.Skowron, “Rough sets and Boolean reasoning,” *Information Sciences*, **177**, 41-73 (2007).
  12. W.Pedrycz, G.Vukovich, “Granular worlds: representation and communication problems,” *International Journal of Intelligent Systems*, **15**, 1015-1026 (2000).
  13. Y.Qian, J.Liang, C.Dang, “Knowledge structure, knowledge granulation and knowledge distance in a knowledge base,” *International Journal of Approximate Reasoning*, **50**, 174-188 (2009).
  14. Y.Qian, H.Zhang, F.Li, Q.Hu, J.Liang, “Set-based granular computing: A lattice model,” *International Journal of Approximate Reasoning*, **55**, 834-852 (2014).
  15. W.Z. Wu, Y. Leung, J.S. Mi, “Granular computing and knowledge reduction in formal contexts,” *IEEE Transactions on Knowledge and Data Engineering*, **21**, 1461-1474 (2009).
  16. L.A.Zadeh, “Fuzzy sets,” *Information and Control*, **8**, 338-353 (1965).
  17. L.A.Zadeh, “Fuzzy logic equals computing with words,” *IEEE Transactions on Fuzzy Systems*, **4**, 103-111 (1996).
  18. L.A.Zadeh, “Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic,” *Fuzzy Sets and Systems*, **90**, 111-127 (1997).
  19. L.A.Zadeh, “Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information / intelligent systems,” *Soft Computing*, **2**, 23-25 (1998).
  20. L.A.Zadeh, “A new direction in AI-Toward a computational theory of perceptions,” *Ai Magazine*, **22**, 73-84 (2001).
  21. W.Zhang, G.Qiu, “Uncertain decision making based on rough sets”, Tsinghua University Publishers, Beijing, 2005.
  22. L.Zhang, B.Zhang, “Theory and application of problem solving-theory and application of granular computing in quotient spaces, Tsinghua University Publishers, Beijing, 2007.