J Syst Sci Syst Eng (Dec 2007) 16(4): 434-449
ISSN: 1004-3756 (Paper) 1861-9576 (Online)
DOI: 10.1007/s11518-007-5059-1
CN11-2983/N

KNOWLEDGE DISTANCE IN INFORMATION SYSTEMS*

Yuhua QIAN[1,2] Jiye LIANG[1] Chuangyin DANG[2] Feng WANG[1] Wei XU[3]

[1] *School of Computer & Information Technology, Shanxi University, Taiyuan, 030006, China*
ljy@sxu.edu.cn (✉)
[2] *Department of Manufacturing Engineering and Engineering Management,*
City University of Hong Kong, Hong Kong
mecdang@cityu.edu.hk
[3] *School of Management, Graduate University of Chinese Academy of Sciences, Chinese Academy of Sciences*
Beijing, 100080, China

Abstract

In this paper, we first introduce the concepts of knowledge closeness and knowledge distance for measuring the sameness and the difference among knowledge in an information system, respectively. The relationship between these two concepts is a strictly mutual complement relation. We then investigate some important properties of knowledge distance and perform experimental analyses on two public data sets, which show the presented measure appears to be well suited to characterize the nature of knowledge in an information system. Finally, we establish the relationship between the knowledge distance and knowledge granulation, which shows that two variants of the knowledge distance can also be used to construct the knowledge granulation. These results will be helpful for studying uncertainty in information systems.

Keywords: Information systems, knowledge, knowledge distance, knowledge granulation

1. Introduction

As a recently renewed research topic, granular computing (GrC) is an umbrella term to cover any theories, methodologies, techniques, and tools that make use of granules in problem solving（Zadeh 1996, Zadeh 1997, Zadeh 1998）. Basic ideas of GrC have appeared in related fields, such as interval analysis, rough set theory, cluster analysis, machine learning, databases, and many others (Zadeh 1979). Zadeh (1997) identified three basic concepts that underlie the process of human cognition, namely, granulation, organization, and causation. A granule is a clump of objects (points), in the universe of discourse, drawn together by indistinguishability, similarity, proximity, or functionality. In

* This work was supported by the National Natural Science Foundation of China under Grant Numbers 60773133, 70471003, and 60573074, the High Technology Research and Development Program of China under Grant No. 2007AA01Z165, the Foundation of Doctoral Program Research of the Ministry of Education of China under Grant No. 20050108004, and Key Project of Science and Technology Research of the Ministry of Education of China.

situations involving incomplete, uncertain, or vague information, it may be difficult to differentiate different elements and instead it is convenient to consider granules, i.e., clump or group of indiscernible elements, for performing operations. Although detailed information may be available, it may be sufficient to use granules in order to have an efficient and practical solution. Very precise solutions may not be required for many practical problems. The acquisition of precise information may be too costly and coarse-grained information reduced cost. There is clearly a need for the systematic studies of granular computing.

A general framework of granular computing was presented by Zadeh (1997) in the context of fuzzy set theory. Granules are defined by generalized constraints. Examples of constraints are equality, possibilistic, probabilistic, fuzzy, and veristic constraints. Many specific models of granular computing have also been proposed. Pawlak (1991), Polkowski and Skowron (1998), and Yao (2006) examined granular computing in connection with the theory of rough sets. Yao (1996, 2000) suggested the use of hierarchical granulations for the study of stratified rough set approximations. Lin (1998) and Yao (1999) studied granular computing using neighborhood systems. Klir (1998) investigated some basic issues of computing with granular probabilities. In the literature (Zhang and Zhang 2003), Zhang extended the theory of quotient space into the theory of fuzzy quotient space based on fuzzy equivalence relation, in which they studied topology relation among objects, and provided theory basis for fuzzy granular computing. Liang et al. (Liang and Shi 2004, Liang and Li 2005) gave a measure called knowledge granulation for measuring the uncertainty of knowledge in rough set theory from the view of granular computing. Liang and Qian (2005) studied rough sets approximation based on dynamic granulation and its application for rule extracting. Qian and Liang (2006a) extended the Pawlak's rough set model to rough set model based on multi-granulations (MGRS), where the set approximations are defined by using multi-equivalences on the universe.

Recently, the rough set theory proposed by Pawlak (1991) has become a popular mathematical framework for granular computing. 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”. The primary use of rough set theory has so far mainly been in generating logical rules for classification and prediction (Skowron and Rauszer 1992) using information granules; thereby making it a prospective tool for pattern recognition, image processing, feature selection, data mining and knowledge discovery process from large data sets. Use of rough set rules based on reducts has a significant role for dimensionality reduction/feature selection by discarding redundant features; thereby having potential application for mining large data sets (Komorouski et al. 1999).

Knowledge base and indiscernibility relation are two basic concepts in Pawlak's rough set theory. Research on the uncertainty of knowledge in knowledge base becomes an important issue in resent years and information entropy and knowledge granulation are two main approaches. For our further development, we briefly review some relative researches. The entropy of a system, as defined by Shannon gives a measure of uncertainty about its actual structure (Shannon 1948). It has been a useful mechanism for characterizing the uncertainty in various modes and applications in many diverse fields. Several authors have used Shannon's entropy and its variants to measure uncertainty of knowledge in rough set theory (Beaubouef et al. 1998, Duntsch and Gediga 1998, Chakik et al. 2004). A new definition for information entropy in rough set theory was presented by Liang in the literature (Liang et al. 2002). Unlike the logarithmic behavior of Shannon entropy, the gain function considered there possesses the complement nature. Combination entropy and combination granulation in incomplete information system were proposed by Qian and Liang for measuring uncertainty of knowledge (Qian and Liang 2006b), their gain function possesses intuitionistic knowledge content nature. Especially, Wierman (1999) presented a well justified measure of uncertainty, the measure of granularity, along with an axiomatic derivation. Its strong connections to the Shannon entropy and the Hartley measure of uncertainty (Hartley 1928) also lend strong support to its correctness and applicability. Furthermore, the relationships among information entropy, rough entropy and knowledge granulation in information systems were established (Liang and Shi 2004). In essence, knowledge granulation characterizes and defines average measure of information granules in a given partition or cover on the universe. Although the information entropy and knowledge granulation can effectively characterizes the uncertainty of knowledge, the difference in between all knowledge in a knowledge base. In fact, if knowledge granulations or information entropy of two knowledge have the same value, then these two knowledge have the same discernibility ability in information systems. Therefore, this kind of measures cannot be used to characterize the difference between two knowledge on the universe. In many practical issues, however, we need often to distinguish any two knowledge for uncertain data processing. Thus, a more comprehensive and effective measure for depicting the difference between knowledge is desirable.

This paper aims to present an approach to measure the difference among knowledge in an information system. The rest of this paper is organized as follows. In Section 2, some preliminary concepts such as complete information systems, incomplete information systems and partial relation are brief recalled. In Section 3, the concept of a knowledge closeness is introduced to measure the similarity between two knowledge in information systems. In Section 4, the concept of a knowledge distance is presented for describing the difference among knowledge on the universe and its some important mathematical properties are derived. In Section 5, we establish the relationship between the knowledge distance and the knowledge granulation. Section 6 concludes the paper.

2. Preliminaries

In this section, some basic concepts are reviewed, which are complete information systems, incomplete information systems and partial relation among knowledge.

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 $a: U \rightarrow V_a$, where V_a is called the value set of a.

For an information system $S=(U,A)$, if $\forall a \in A$, every element in V_a is a definite value, then S is called a complete information system.

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

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

It is easily shown that

$$IND(P)=\bigcap_{a\in P} IND(\{a\}).$$

$U/IND(P)$ constitutes a partition of U. $U/IND(P)$ is called a knowledge in U and every equivalence class is called a knowledge granule or information granule (Liang et al. 2006). Information granulation, in some sense, denotes the average measure of information granules (equivalence classes) in P. In general, we denote the knowledge induced by $P \subseteq A$ by U/P.

Example 2.1 Consider descriptions of several cars in Table 1.

This is a complete information system, where $U=\{u_1,u_2,u_3,u_4,u_5,u_6\}$ and $A=\{a_1,a_2,a_3,a_4\}$, with a_1-Price, a_2-Mileage, a_3-Size, a_4-Max-Speed. By computing, it follows that

$$U/IND(A)=\{\{u_1\},\{u_2,u_6\},\{u_3\},\{u_4,u_5\}\}.$$

It may happen that some of the attribute values for an object are missing. For example, in medical information systems there may exist a group of patients for which it is impossible to perform all the required tests. These missing values can be represented by the set of all possible values for the attribute or equivalence by the domain of the attribute. To indicate such a situation, a distinguished value, a so-called null value is usually assigned to those attributes.

Table 1 The complete information system about car (Kryszkiewicz 1998, Kryszkiewicz 1999)

Car	Price	Mileage	Size	Max-Speed
u_1	High	Low	Full	Low
u_2	Low	High	Full	Low
u_3	Low	Low	Compact	Low
u_4	High	High	Full	High
u_5	High	High	Full	High
u_6	Low	High	Full	Low

If V_a contains a null value for at least one attribute $a \in A$, then S is called an incomplete information system (Liang et al. 2006, Kryszkiewicz 1998, Kryszkiewicz 1999), otherwise it is called a complete information system. From now on, we will denote the null value by $*$.

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

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

In fact, $SIM(P)$ is a tolerance relation on U. The concept of a tolerance relation has a wide variety of applications in classification (Liang et al. 2006, Kryszkiewicz 1998, Kryszkiewicz 1999). It can be easily shown that

$$SIM(P)=\bigcap_{a\in P} SIM(\{a\}).$$

Let $U/SIM(P)$ denote the family sets

$\{S_P(u) \mid u \in U\}$, 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 cover of U, i.e., $S_P(u) \neq \varnothing$ for every $u \in U$ and $U_{u\in U} S_P(u) = U$.

Of course, $SIM(P)$) degenerates into an equivalence relation in a complete information system.

Example 2.2 Consider descriptions of several cars in Table 2.

Table 2 The complete information system about car (Kryszkiewicz 1998, Kryszkiewicz 1999)

Car	*Price*	*Mileage*	*Size*	*Max-Speed*
u_1	High	Low	Full	Low
u_2	Low	*	Full	Low
u_3	*	*	Compact	Low
u_4	High	*	Full	High
u_5	*	*	Full	High
u_6	Low	High	Full	*

This is an incomplete information system, where $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$ and $A = \{a_1, a_2, a_3, a_4\}$, with a_1 -Price, a_2 -Mileage, a_3 -Size, a_4-Max-Speed. By computing, it follows that

$$U / SIM(A) = \{S_A(u_1), S_A(u_2), S_A(u_3), S_A(u_4), S_A(u_5), S_A(u_6)\},$$

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

Of particular interest is the discrete classification

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

and the indiscrete classification

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

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

Now we define a partial order on the set of all classifications of U. Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$, $U / SIM(P) = \{S_P(u_1), S_P(u_2), \cdots, S_P(u_{|U|})\}$ and $U / SIM(Q) = \{S_Q(u_1), S_Q(u_2), \cdots, S_Q(u_{|U|})\}$. We define a partial relation $\preceq$ as follows

$$P \preceq Q \Leftrightarrow S_P(u) \subseteq S_Q(u), \forall u \in U.$$

When S be a complete information system, there are two partitions $U/IDN(P) = \{P_1, P_2, \cdots, P_m\}$ and $U/IDN(P) = \{Q_1, Q_2, \cdots, Q_n\}$. Then the partial relation has the following property (Liang and Li 2005)

$P \preceq Q \Leftrightarrow$ for any $P_i \in U/IND(P)$, there exists $Q_j \in U/IND(Q)$ such that $P_i \subseteq Q_j$.

3. Knowledge Closeness

In this section, we extend the concept of set closeness to the concept of knowledge closeness for measuring the closeness degree between two knowledge in an information system.

Tolerance classes induced by attribute set A are described by a family of sets $\{S_A(u) \mid u \in U\}$ in an incomplete information system. In fact, a complete information system is a special form of incomplete information systems. Let $S = (U, A)$ be a complete information system, $U/SIM(A) = \{S_A(u_1), S_A(u_2), \cdots, S_A(u_{|U|})\}$, $U/IND(A) = \{X_1, X_2, \cdots, X_m\}$ and $X_i = \{u_{i1}, u_{i2}, \cdots, u_{im_i}\}$, where $|X_i| = s_i$ and $\sum_{i=1}^{m} |X_i| = |U|$. Then, the relationship between the elements in $U/SIM(A)$ and the elements in $U/IND(A)$ is as follows (Liang et al. 2006)

$X_i = S_A(u_{i1}) = S_A(u_{i2}) = \cdots = S_A(u_{im_i})$ and

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

Definition 3.1 (Yao 2001) Let A , B be two finite sets. Set closeness between A and B is defined as

$$H(A,B) = \frac{|A \cap B|}{|A \cup B|} \ (A \cup B) \neq \varnothing , \qquad (1)$$

where $0 \le H(A,B) \le 1$ and we assume that $H(A,B) = 1$ if $(A \cup B) \neq \varnothing$.

If $A = B$, then the set closeness between A and B achieves maximum value 1.

If $A \cap B = \varnothing$, then the set closeness between A and B achieves minimum value 0.

The set closeness denotes the measure of the similarity between two sets. The more the overlap between these two sets is, the large the value of H is, and vice versa.

In order to investigate the measure of the similarity between two knowledge and its some properties, we here introduce the concept of complement of knowledge. Let $U/SIM(A) = \{S_A(u_1),\ S_A(u_2), \cdots, S_A(u_{|U|})$ be the knowledge induced by attribute set A on the universe U, then the complement of this knowledge is defined as

$\sim (U/SIM(A)) = \{ \{u_1\} \cup (U - S_A(u_1)),$
$\{u_2\} \cup (U - S_A(u_2)), \cdots, \{u_{|U|}\} \cup (U - S_A(u_{|U|}))\}$.

That is

$\sim (U/SIM(A)) = \{ \{u_i\} \cup (U - S_A(u_i)) \mid i \in U\}$.

Proposition 3.1 *Let $S = (U, A)$ be an information system, $P,Q \subseteq A$ and $U/SIM(P)$, $U/SIM(Q)$ two knowledge on the universe U. If $U/SIM(P) = \sim (U/SIM(Q))$, then U/SIM $(P \cup Q) = \omega$.*

Proof. From the definition of tolerance relation, we have that for arbitrary $i \in U$, the tolerance classes induced by u_i in $U/SIM(P)$, $U/SIM(Q)$ and $U/SIM(P \cup Q)$ are $S_P(u_i)$, $S_P(u_i)$ and $S_{P\cup Q}(u_i)$, respectively. Since $U/SIM(P) = \sim (U/SIM(Q))$, we have that

$S_Q(u_i) = \{u_i\} \cup (U - S_P(u_i)) \ (i \le |U|)$.

Hence, for arbitrary $i \in U$, it follows that

$S_{P\cup Q}(u_i) = S_P(u_i) \cap S_P(u_i)$
$= S_P(u_i) \cap (\{u_i\} \cup (U - S_P(u_i))) = \{u_i\}$

Therefore,

$U/SIM(P \cup Q) = \{S_{P\cup Q}(u_i) = \{u_i\}, i \le |U|\} = \omega$

This completes the proof. ■

Corollary 3.1 *The following properties hold*

1) $\sim (\sim (U/SIM(P))) = U/SIM(P)$,
2) $\omega = \sim \delta, \delta = \sim \omega$.

Definition 3.2 Let $S = (U, A)$ be an information system, $P,Q \subseteq A$, and $U/SIM(P)$, $U/SIM(Q)$ two knowledge on the universe U , where

$U/SIM(P) = \{S_P(u_1), S_P(u_2), \cdots, S_P(u_{|U|})\}$,

$U/SIM(Q) = \{S_Q(u_1), S_Q(u_2), \cdots, S_Q(u_{|U|})\}$.

Knowledge closeness between the knowledge $U/SIM(P)$ and the knowledge $U/SIM(Q)$ is defined as

$$H(P,Q) = \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i) \cap S_Q(u_i)|}{|S_P(u_i) \cup S_Q(u_i)|} , \qquad (2)$$

where $\frac{1}{|U|} \le H(P,Q) \le 1$.

The knowledge closeness represents the measure of the similarity between two knowledge on U. The more the overlap between knowledge is, the larger knowledge closeness H

is, and vice versa.

Proposition 3.2 *(Maximum) Let* $S=(U,A)$ *be an information system,* $P,Q\subseteq A$, *and* $U/SIM(P)$, $U/SIM(Q)$ *two knowledge on the universe* U. *If* $U/SIM(P)=U/SIM(Q)$, *then the knowledge closeness between the knowledge* $U/SIM(P)$ *and* $U/SIM(Q)$ *achieves its maximum value 1.*

Proof. It is straightforward.

Proposition 3.3 *(Minimum) Let* $S=(U,A)$ *be an information system,* $P,Q\subseteq A$, *and* $U/SIM(P)$, $U/SIM(Q)$ *two knowledge on the universe U. If* $U/SIM(P)=\sim(U/SIM(Q))$, *then the knowledge closeness between the knowledge* $U/SIM(P)$ *and the knowledge* $U/SIM(Q)$ *achieves minimum value* $\frac{1}{|U|}$.

Proof. From the definition of tolerance relation, we have that for arbitrary $i\in U$, the tolerance classes induced by u_i in $U/SIM(P)$ and $U/SIM(Q)$ are $S_P(u_i)$ and $S_Q(u_i)$, respectively. Since $U/SIM(P)=\sim(U/SIM(Q))$ we have

$$S_Q(u_i)=\{u_i\}\cup(U-S_P(u_i))(i\le|U|).$$

Hence, we have that

$$H(P,Q)=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}$$

$$=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)\cap(\{u_i\}\cup(U-S_P(u_i)))|}{|S_P(u_i)\cup\{u_i\}\cup(U-S_P(u_i))|}$$

$$=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|(S_P(u_i)\cap\{u_i\})\cup(S_P(u_i)\cap(U-S_P(u_i)))|}{|S_P(u_i)\cup(U-S_P(u_i))|}$$

$$=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|\{u_i\}\cup\phi|}{|S_P(u_i)\cup(U-S_P(u_i))|}$$

$$=\frac{1}{|U|}.$$

This completes the proof. ■

Corollary 3.2 $H(\omega,\delta)=\frac{1}{|U|}$.

4. Knowledge Distance

In this section, we first introduce the concept of knowledge distance to measure the difference between two knowledge on the same universe. Then, its some important mathematical properties are obtained. Finally, experimental analyses on two public data sets are performed for verifying the validity of this knowledge distance.

Definition 4.1 Let $S=(U,A)$ be an information system, $P,Q\subseteq A$, and $U/SIM(P)$, $U/SIM(Q)$ two knowledge on the universe U, where

$U/SIM(P)=\{S_P(u_1),S_P(u_2),\cdots,S_P(u_{|U|})\}$ and
$U/SIM(Q)=\{S_Q(u_1),S_Q(u_2),\cdots,S_Q(u_{|U|})\}$.

Knowledge distance between the knowledge $U/SIM(P)$ and the knowledge $U/SIM(Q)$ is defined as

$$D(P,Q)=\frac{1}{|U|}\sum_{i=1}^{|u|}(1-\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}),\qquad(3)$$

where $0\le D(P,Q)\le 1-\frac{1}{|U|}$.

The knowledge distance denotes the measure of difference between two knowledge on the same universe. The more the overlap between knowledge is, the smaller knowledge distance H is, and vice versa.

Proposition 4.1 *(Maximum) Let* $S=(U,A)$ *be an information system,* $P,Q \subseteq A$*, and* $U/SIM(P)$, $U/SIM(Q)$ *two knowledge on the universe* U. *If* $U/SIM(P) = \sim(U/SIM(Q))$, *then the knowledge distance between the knowledge* $U/SIM(P)$ *and the knowledge* $U/SIM(Q)$ *achieves maximum value* $1-\frac{1}{|U|}$.

Proof. It is straightforward.

Corollary 4.1 $D(\omega,\delta)=1-\frac{1}{|U|}$.

Proof. This proof is similar to that of Proposition 3.3.

Proposition 4.2 *(Minimum) Let* $S=(U,A)$ *be an information system,* $P,Q \subseteq A$*, and* $U/SIM(P)$, $U/SIM(Q)$ *two knowledge on the universe U. If* $U/SIM(P)=U/SIM(Q)$, *then the knowledge closeness between the knowledge* $U/SIM(P)$ *and the knowledge* $U/SIM(Q)$ *achieves minimum value 0.*

Proof. It is straightforward.

Proposition 4.3 *Knowledge distance* D *has the following properties*

1) $D(P,Q)\ge 0$ *(non-negative),*

2) $D(P,Q)=D(Q,P)$ *(symmetrical).*

Proof. They are straightforward.

Proposition 4.4 *Let* $S=(U,A)$ *be an information system,* $P,Q \subseteq A$*, and* $U/SIM(P)$, $U/SIM(Q)$ *two knowledge on the universe* U. *Then,* $H(P,Q)+D(P,Q)=1$.

Proof. From Definition 3.2 and 4.1, we have that

$$D(P,Q)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}\right)$$
$$=1-\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}$$
$$=1-H(P,Q).$$

Hence, $H(P,Q)+D(P,Q)=1$. ■

Obviously, there is a strictly mutual complement relation between the knowledge distance and the knowledge closeness in terms of Definition 3.2 and 4.1.

Example 4.1 For Table 1, let P={ Price } and Q ={Max-speed}. Compute the knowledge distance between P and Q.

By computing, we have that

$$U/IND(P)=\{\{u_1,u_4,u_5\},\{u_2,u_3,u_6\}\},$$
$$U/IND(Q)=\{\{u_1,u_2,u_3,u_6\},\{u_4,u_5\}\}.$$

If we regard Table 1 as a special incomplete information system, we can obtain the following

$$U/SIM(P)=\{\{u_1,u_4,u_5\},\{u_2,u_3,u_6\},\{u_2,u_3,u_6\},\{u_1,u_4,u_5\},\{u_1,u_4,u_5\},\{u_2,u_3,u_6\}\},$$
$$U/SIM(Q)=\{\{u_1,u_2,u_3,u_6\},\{u_1,u_2,u_3,u_6\},.\{u_1,u_2,u_3,u_6\},\{u_4,u_5\},\{u_4,u_5\},\{u_1,u_2,u_3,u_6\}\}.$$

By computing, the knowledge distance between P and Q is

$$D(P,Q)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap S_P(u_i)|}{|S_P(u_i)\cup S_P(u_i)|}\right)$$
$$=\frac{1}{6}[(1-\frac{1}{6})+(1-\frac{3}{4})+(1-\frac{3}{4})+(1-\frac{2}{3})+(1-\frac{2}{3})+(1-\frac{3}{4})]$$
$$=\frac{3}{8},$$

and the knowledge closeness between P and Q is

$$H(P,Q)=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)\cap S_P(u_i)|}{|S_P(u_i)\cup S_P(u_i)|}$$
$$=\frac{1}{6}(\frac{1}{6}+\frac{3}{4}+\frac{3}{4}+\frac{2}{3}+\frac{2}{3}+\frac{3}{4})$$
$$=\frac{5}{8}.$$

Therefore,

$$H(P,Q)+D(P,Q)=\frac{5}{8}+\frac{3}{8}=1.$$

Proposition 4.5 *Let* $S=(U,A)$ *be an information system,* $P,Q,R\subseteq A$ *with* $P\preceq Q\preceq R$*. Then* $D(P,R)\geq D(P,Q)$ *and* $D(P,R)\geq D(Q,R)$.

Proof. If we regard S as an incomplete information system, then $U/SIM(P)=\{S_P(u_1), S_P(u_2),\cdots,S_P(u_{|U|})\}$, $U/IND(P)=\{P_1,P_2,\cdots, P_m\}$ and $P_i=\{u_{i1},u_{i2},\cdots,u_{im_i}\}$, where $|P_i|=s_i$ and $\sum_{i=1}^{m}|X_i|=|U|$. In other words, complete information systems and incomplete information systems can be consistently represented. Then, the relationship between the elements in $U/IND(P)$ and the elements in $U/SIM(P)$ is as follows

$$P_i=S_P(u_{i1})=S_P(u_{i2})=\cdots=S_P(u_{im_i}).$$

Similarly,

$$Q_j=S_Q(u_{j1})=S_Q(u_{j2})=\cdots=S_Q(u_{jm_j}),$$
$$R_k=S_R(u_{k1})=S_R(u_{k2})=\cdots=S_R(u_{km_k}).$$

Since $P\preceq Q\preceq R$, we have

$S_P(u)\subseteq S_Q(u)\subseteq S_R(u)$ for arbitrary $u\in U$. So,

$S_P(u)\cup S_R(u)\supseteq S_P(u)\cup S_Q(u)$ and

$S_P(u)\cap S_Q(u)=S_P(u)\cup S_R(u)$.

Hence, one can obtain that

$|S_P(u)\cap S_R(u)|\geq|S_P(u)\cap S_Q(u)|$ and

$|S_P(u)\cap S_Q(u)|=|S_P(u)\cup S_R(u)|$.

Therefore, we have that

$$\begin{aligned}
&D(P,R)-D(P,Q)\\
&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap S_R(u_i)|}{|S_P(u_i)\cup S_R(u_i)|}\right)\\
&\quad-\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}\right)\\
&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}-\frac{|S_P(u_i)\cap S_R(u_i)|}{|S_P(u_i)\cup S_R(u_i)|}\right)\\
&\geq 0.
\end{aligned}$$

Similarly, we have $D(P,R)-D(P,Q)\geq 0$. Therefore, $D(P,R)\geq D(P,Q)$ and $D(P,R)\geq D(Q,R)$ hold. ■

Proposition 4.6 *Let* $S=(U,A)$ *be an information system,* $P,Q,R\subseteq A$ *with* $P\preceq Q\preceq R$*. Then*

$D(P,R)+D(Q,R)\geq D(P,Q)$,
$D(P,R)+D(P,Q)\geq D(Q,R)$ *and*
$D(P,Q)+D(Q,R)\geq D(P,R)$.

Proof. From Proposition 4.4, one can know that $D(P,R)\geq D(P,Q)$ and $D(P,R)\geq D(Q,R)$ if $P\preceq Q\preceq R$. It is clear that

$D(P,R)+D(Q,R)\geq D(P,Q)$ and
$D(P,R)+D(P,Q)\geq D(Q,R)$ hold.

Therefore, we just need to prove

$$D(P,Q)+D(Q,R)\geq D(P,R).$$

Similar to the proof of Proposition 4.4, we can get that $S_P(u)\subseteq S_Q(u)\subseteq S_R(u)$ if $P\preceq Q\preceq R$
That is to say,

$$S_P(u)\cap S_Q(u)=S_P(u),$$
$$S_P(u)\cup S_Q(u)=S_Q(u);$$

$$S_P(u)\cap S_R(u)=S_P(u),$$
$$S_P(u)\cup S_R(u)=S_R(u);$$

and $S_Q(u)\cap S_R(u)=S_Q(u)$,

$$S_Q(u)\cup S_R(u)=S_R(u).$$

So $|S_P(u)|\leq|S_Q(u)|\leq|S_R(u)|$. Therefore, we have that

$$\begin{aligned}
&D(P,Q)+D(Q,R)-D(P,R)\\
&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}\right)
\end{aligned}$$

$$+\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap S_R(u_i)|}{|S_Q(u_i)\cup S_R(u_i)|})$$
$$-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap S_R(u_i)|}{|S_P(u_i)\cup S_R(u_i)|})$$
$$=\frac{1}{|U|}\sum_{i=1}^{|U|}(\frac{|S_P(u_i)\cap S_R(u_i)|}{|S_P(u_i)\cup S_R(u_i)|}+1$$
$$-\frac{|S_P(u_i)\cap S_Q(u_i)|}{|S_P(u_i)\cup S_Q(u_i)|}-\frac{|S_Q(u_i)\cap S_R(u_i)|}{|S_Q(u_i)\cup S_R(u_i)|})$$
$$=\frac{1}{|U|}\sum_{i=1}^{|U|}(\frac{|S_P(u_i)|}{|S_R(u_i)|}+1-\frac{|S_P(u_i)|}{|S_Q(u_i)|}-\frac{|S_Q(u_i)|}{|S_R(u_i)|}).$$

Denoted by $p=|S_P(u_i)|$, $q=|S_Q(u_i)|$ and $r=|S_R(u_i)|$. From $|S_P(u)|\leq|S_Q(u)|\leq|S_R(u)|$, it follows that $0<p\leq q\leq r\leq|U|$. Suppose that the function $f(p,q,r)=1+\frac{p}{r}-\frac{p}{q}-\frac{q}{r}$. Here, we only need to prove $f(p,q,r)\geq 0$. Therefore

$$f(p,q,r)=1+\frac{p}{r}-\frac{p}{q}-\frac{q}{r}=\frac{qr+pq-pr-q^2}{qr}$$
$$=\frac{(r-q)(q-p)}{qr}\geq 0.$$

Hence, $D(P,Q)+D(Q,R)-D(P,R)$
$=\frac{1}{|U|}\sum_{i=1}^{|U|}f(p,q,r)\geq 0$. This completes the proof. ■

In the following, through experimental analyses, we illustrate some properties of the knowledge distance in information systems. We have downloaded two public data sets with practical applications from UCI Repository of machine learning databases, which are information system dermatology with 240 objects and information system monks-problems with 432 objects. All condition attributes in the two data sets are discrete. We analyze knowledge distances between knowledge induced by all attributes of an information system and knowledge induced by various numbers of attributes. The changes of values of knowledge distances with the number of attributes in these two data sets are shown in Figure 1 and Figure 2.

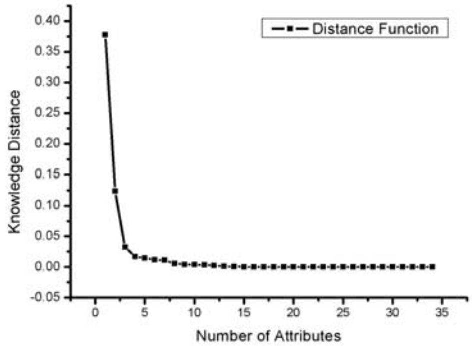

Figure 1 Knowledge distance with the number of attributes about dermatology

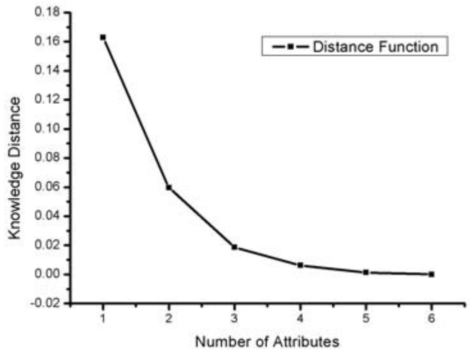

Figure 2 Knowledge distance with the number of attributes about monks-problems

It can be seen from Figure 1 and Figure 2 that the value of knowledge distance decreases as the number of selected attributes becomes bigger in the same data set. In other words, through adding number of attributes, the knowledge induced by these attributes can approach to the knowledge induced by all attributes in this information system, i.e., the knowledge distance between them can approach to zero. Therefore, we can draw a conclusion that the knowledge distance can well characterize the difference between two knowledge in the same information system.

5. Relationship between Knowledge Distance and Knowledge Granulation

Knowledge granulation is an important concept of granular computing proposed by Zadeh (1997). The knowledge granulation of an information system gives a measure of uncertainty about its actual structure. In general, knowledge granulation can represent the discernibility ability of knowledge in information systems. Especially, several measures in an information system closely associated with granular computing such as granulation measure, information entropy, rough entropy and knowledge granulation and their relationships were discussed (Liang et a. 2004, Liang et al. 2006). In the literature (Qian and Liang 2006b), we introduced two concepts so-called combination entropy and combination granulation to measure the uncertainty of an information system. In the literature (Liang and Qian 2006), an axiom definition of knowledge granulation was given, which gives a unified description for knowledge granulation. In this section, we will discuss the relationship between knowledge distance and knowledge granulation.

Let $S=(U,A)$ be an information system, $P,Q \subseteq A$. $K(P)=\{S_P(x_i) \mid x_i \in U\}$, $K(Q)=\{S_Q(x_i) \mid x_i \in U\}$. We define a partial relation $\preceq$ with set size character as follows (Liang and Qian 2006):

$P \preceq' Q$ if and only if, for $K(P)=\{S_P(x_1), S_P(x_2),\cdots,S_P(x_{|U|})\}$, there exists a sequence $K'(Q)$ of $K(Q)$, where

$$K'(P)=\{S_P(x'_1),S_P(x'_2),\cdots,S_P(x'_{|U|})\},$$

such that $|S_P(x_i)| \le |S_Q(x'_i)|$. If there exists a sequence $K'(Q)$ of $K(Q)$ such that $|S_P(x_i)| \le |S_Q(x'_i)|$, then we will call that P is strict granulation finer than Q, and denote it by $P \prec' Q$.

Definition 5.1 (Liang and Qian 2006) Let $S=(U,A)$ be an information system and G be a mapping from the power set of A to the set of real numbers. We say that G is a knowledge granulation in an information system if G satisfies the following conditions:

1) $G(P) \ge 0$ for any $P \subseteq A$ (Non-negativity);

2) $G(P)=G(Q)$ for any $P,Q \in A$ if there is a bijective mapping function $f: K(P) \to K(Q)$ such that

$$|S_P(u_i)| = |f(S_P(u_i))| (\forall i \in \{1,2,\cdots,|U|\}),$$

where $K(P)=\{S_P(u_i) \mid x_i \in U\}$ and $K(Q)=\{S_Q(u_i) \mid x_i \in U\}$ (Invariability);

3) $G(P) < G(Q)$ for any $P,Q \in A$ with $P \prec' Q$ (Monotonicity).

Corollary 5.1 If $P \preceq Q$, then $D(P,\omega) \le D(Q,\omega)$.

Proof. Since knowledge $\omega = \{\{u_i\} \mid u_i \in U\}$ and $P \preceq Q$. So for u_i we have that

$$\{u_i\} \subseteq S_P(u_i) \subseteq S_Q(u_i).$$

Thus, $1 \le |S_P(u_i)| \le |S_Q(u_i)|$.

Hence, we have that

$$D(P,\omega)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(1-\frac{|S_P(u_i)\cap\{u_i\}|}{|S_P(u_i)\cup\{u_i\}|}\right)$$

$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)|-1}{|S_P(u_i)|}$$
$$\le \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(u_i)|-1}{|S_Q(u_i)|}$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap\{u_i\}|}{|S_Q(u_i)\cup\{u_i\}|})$$
$$= D(Q,\omega),$$

i.e., $D(P,\omega)\le D(Q,\omega)$. This completes the proof. ■

Proposition 5.1 $G_1(P)=D(P,\omega)$ *is a knowledge granulation in Definition 5.1.*

Proof. 1) Obviously, it is non-negative;

2) Let $P,Q\subseteq A$, then

$U/IND(P)=\{P_1,P_2,\cdots,P_m\}$ and $U/IND(Q)=\{Q_1,Q_2,\cdots,Q_n\}$ in a complete information system can be denoted by

$U/SIM(P)=\{S_P(u_1),S_P(u_2),\cdots,S_P(u_{|U|})\}$

and

$U/SIM(Q)=\{S_Q(u_1),S_Q(u_2),\cdots,S_Q(u_{|U|})\}$.

Suppose that there be a bijective mapping function $f:U/SIM(P)\to U/SIM(Q)$ such that $|S_P(u_i)|=|f(S_P(u_i))|(i\in\{1,2,\cdots,|U|\})$ and $f(S_P(u_i))=S_Q(u_{j_i})(j_i\in\{1,2,\cdots,|U|\})$, then we have that

$$D(P,\omega)=\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap\{u_i\}|}{|S_P(u_i)\cup\{u_i\}|})$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)|-1}{|S_P(u_i)|}$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(u_{j_i})|-1}{|S_Q(u_{j_i})|}$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap\{u_i\}|}{|S_Q(u_i)\cup\{u_i\}|})$$
$$= D(Q,\omega),$$

i.e., $G_1(P)=G_2(Q)$.

3) Let $P,Q\subseteq A$ with $P\prec' Q$, then for arbitrary $S_P(u_i)(i\le|U|)$, there exists a sequence $\{S_Q(u_1^{'}),S_Q(u_2^{'}),\cdots,S_Q(u_{|U|}^{'})\}$ such that $|S_P(u_i)|<|S_Q(u_i)|$. Hence, we obtain that

$$D(P,\omega)=\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap\{u_i\}|}{|S_P(u_i)\cup\{u_i\}|})$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(u_i)|-1}{|S_P(u_i)|}$$
$$< \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(u_i^{'})|-1}{|S_Q(u_i^{'})|}$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap\{u_i\}|}{|S_Q(u_i)\cup\{u_i\}|})$$
$$= D(Q,\omega),$$

i.e., $G_1(P)<G_2(Q)$. This completes the proof. ■

Corollary 5.2 *If* $P\preceq Q$ *, then* $D(P,\delta)\ge D(Q,\delta)$.

Proof. Since $\delta=\{S_P(u_i)\mid S_P(u_i)=U,u_i\in U\}$ and $P\preceq Q$. So for u_i we have that

$$S_P(u_i)\subseteq S_Q(u_i)\subseteq U .$$

Thus $|S_P(u_i)|\le|S_Q(u_i)|\le|U|$.

Hence, we have that

$$D(P,\delta)=\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap U|}{|S_P(u_i)\cup U|})$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|-|S_P(u_i)|}{|U|}$$
$$\ge \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|-|S_Q(u_i)|}{|U|}$$
$$= \frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap U|}{|S_Q(u_i)\cup U|})$$
$$= D(Q,\delta),$$

i.e., $D(P,\delta)\ge D(Q,\delta)$. This completes the proof. ■

Proposition 5.2 $G_2(P)=1-\frac{1}{|U|}-D(P,\delta)$ *is a knowledge granulation in Definition 5.1.*

Proof. 1) Obviously, it is non-negative;

2) Let $P,Q\subseteq A$, then $U/IND(P)=\{P_1, P_2,\cdots,P_m\}$ and $U/IND(Q)=\{Q_1,Q_2,\cdots,Q_n\}$ in complete information system can be denoted by

$U/SIM(P)=\{S_P(u_1),S_P(u_2),\cdots,S_P(u_{|U|})\}$

and

$U/SIM(Q)=\{S_Q(u_1),S_Q(u_2),\cdots,S_Q(u_{|U|})\}$.

Suppose that there be a bijective mapping function $f:U/SIM(P)\to U/SIM(Q)$ such that $|S_P(u_i)|=|f(S_P(u_i))|(i\in\{1,2,\cdots,|U|\})$ and $f(S_P(u_i))=S_Q(u_{j_i})(j_i\in\{1,2,\cdots,|U|\})$, then we have that

$$
\begin{aligned}
G_2(P)&=1-\frac{1}{|U|}-D(P,\delta)\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap U|}{|S_P(u_i)\cup U|})\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|U|-|S_P(u_i)|}{|U|})\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|U|-|S_Q(u_{j_i})|}{|U|})\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap U|}{|S_Q(u_i)\cup U|})\\
&=1-\frac{1}{|U|}-D(Q,\delta)\\
&=G_2(Q),
\end{aligned}
$$

i.e., $G_2(P)=G_2(Q)$.

3) Let $P,Q\subseteq A$ with $P\prec' Q$, then for arbitrary $S_P(u_i)(i\le|U|)$, there exists a sequence $\{S_Q(u_1'),S_Q(u_2'),\cdots,S_Q(u_{|U|}')\}$ such that $|S_P(u_i)|<|S_Q(u_i')|$. Hence, we obtain that

$$
\begin{aligned}
G_2(P)&=1-\frac{1}{|U|}-D(P,\delta)\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_P(u_i)\cap U|}{|S_P(u_i)\cup U|})\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|U|-|S_P(u_i)|}{|U|})\\
&<1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|U|-|S_Q(u_i')|}{|U|})\\
&=1-\frac{1}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}(1-\frac{|S_Q(u_i)\cap U|}{|S_Q(u_i)\cup U|})\\
&=1-\frac{1}{|U|}-D(Q,\delta)\\
&=G_2(Q),
\end{aligned}
$$

i.e., $G_2(P)<G_2(Q)$. This completes the proof. ■

Proposition 5.1 and 5.2 show that $D(P,\omega)$ and $1-\frac{1}{|U|}-D(P,\delta)$ are all special forms of knowledge granulation in Definition 5.1, and can be used to measure the uncertainty of knowledge induced by attribute set $P\subseteq A$ in the view of granular computing.

6. Conclusion

In the view of granular computing, the information entropy and knowledge granulation can measure the discernibility ability of knowledge on the universe. But these two kinds of measures could not felicitously characterize the difference between two knowledge with the same value of information entropy or knowledge granulation. For this reason, a new measure so-called knowledge distance has been introduced to information systems. We have shown the mechanism how this measure characterizes the difference among knowledge by several important properties and experimental analyses on two public data sets. Furthermore, we have pointed out that the relationship between the knowledge distance and knowledge granulation. With the above the discussions, we have developed the theoretical foundation of measuring knowledge distance in information systems for its further research.

References

[1] Beaubouef, T., Petry, F.E. & Arora, G. (1998). Information-theoretic measures of uncertainty for rough sets and rough relational databases. Information Sciences, 109: 535-563

[2] Chakik, F.E., Shahine, A., Jaam, J. & Hasnah, A. (2004). An approach for constructing complex discriminating surfaces based on Bayesian interference of the maximum entropy. Information Sciences, 163: 275-291

[3] Duntsch, I. & Gediga, G. (1998). Uncertainty measures of rough set prediction. Artificial Intelligence, 106: 109-137

[4] Hartley, R.V.L. (1928). Transmission of information. The Bell Systems Technical Journal, 7: 535-563

[5] Klir, G.J. (1998). Basic issues of computing with granular computing. In: Proceedings of 1998 IEEE International Conference on Fuzzy System, 101-105

[6] Komorouski, J., Pawlak, Z., Polkowski, L. & Skowron, A. (1999). Rough sets: A tutorial. In: Pal, S.K., Skowron, A. (eds.), Rough Fuzzy Hybridization: A New Trend in Decision-Making, 3-98, Singapore, Springer

[7] Kryszkiewicz, M. (1998). Rough set approach to incomplete information systems. Information Sciences, 112: 39-49

[8] Kryszkiewicz, M. (1999). Rules in incomplete information systems. Information Sciences, 113: 271-292

[9] Liang, J.Y., Chin, K.S., Dang, C.Y. & Yam, C.M. (2002), A new method for measuring uncertainty and fuzziness in rough set theory. International Journal of General Systems, 31 (4): 331-342

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

[11] Liang, J.Y. & Shi, Z.Z. (2004). The information entropy, rough entropy and knowledge granulation in rough set theory. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 12 (1): 37-46

[12] Liang, J.Y., Shi, Z.Z., Li, D.Y. & Wierman, M.J. (2006). The information entropy, rough entropy and knowledge granulation in incomplete information system. International Journal of General Systems, 35 (6): 641-654

[13] Liang, J.Y. & Qian, Y.H. (2005). Rough set approximation under dynamic granulation. Lecture Notes in Artificial Intelligence, 3641: 701-708

[14] Liang, J.Y. & Qian, Y.H. (2006). Axiomatic approach of knowledge granulation in information system. Lecture Notes in Artificial Intelligence, 4304: 1074-1078

[15] Lin, T.Y. (1998). Granular computing on binary relations I: Data mining and neighborhood systems, II: Rough sets representations and belief functions. In: Polkowski, L, Skowron, A. (eds.), Rough Sets in Knowledge Discovery I, pp. 107-140

[16] Pawlak, Z. (1991). Rough Sets, Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht

[17] Polkowski, L. & Skowron, A. (1998). Towards adaptive calculus of granules, granularity of knowledge, indiscernibility and rough sets. In: Proceedings of 1998 IEEE International Conference on Fuzzy Systems, 111-116

[18] Qian, Y.H. & Liang, J.Y. (2006a). Combination entropy and combination granulation in incomplete information system. Lecture Notes in Artificial Intelligence, 4062: 184-190

[19] Qian, Y.H. & Liang, J.Y. (2006b). Rough set method based on multi-granulations. In: Proceedings of 5th IEEE International Conference on Cognitive Informatics, I: 297-304

[20] Shannon, C.E. (1948). The mathematical theory of communication. The Bell System Technical Journal, 27 (3, 4): 373-423

[21] Skowron, A. & Rauszer, C. (1992). The discernibility matrices and functions in information systems, In: Slowinski, A (ed.), Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, pp. 331-362. Kluwer Academic, Dordrecht

[22] Wierman, M.J. (1999). Measuring uncertainty in rough set theory. International Journal of General Systems, 28 (4): 283-297

[23] Yao, Y.Y. (2001). Information granulation and rough set approximation. International Journal of Intelligent Systems, 16: 87-104

[24] Yao, Y.Y. (2006). Three perspectives of granular computing. In: Proceedings of the International Forum on Granular Computing from the Rough Set Perspectives

[25] Yao, Y.Y. (1996). Two views of the theory of rough sets in finite universes. International Journal of Approximate Reasoning, 15 (4): 291-319

[26] Yao, Y.Y. (2000). Granular computing: basic issues and possible solutions. In: Proceedings of the Fifth International Conference on Computing and Information, I: 186-189

[27] Yao, Y.Y. (1999). Stratified rough sets and granular computing. In: Proceedings of the 18th International Conference of the North American Fuzzy Information Processing Society, 800-804

[28] Zadeh, L.A. (1979). Fuzzy sets and information granularity. In: Gupta, N., Ragade, R., Yager, R., et al. (eds.) Advances in Fuzzy Set Theory and Application, pp. 3-18. Amsterdam: North-Holland

[29] Zadeh, L.A. (1996) Fuzzy logic = computing with words. IEEE Transactions on Fuzzy System, 4 (1): 103-111

[30] Zadeh, L.A. (1997). Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and System, 90: 111-127

[31] Zadeh, L.A. (1998). Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent system. Soft Computing, 2 (1): 23-25

[32] Zhang, L., Zhang, B. (2003). Theory of fuzzy quotient space (methods of fuzzy granular computing). Journal of Software (in Chinese), 14 (4): 770-776

Yuhua Qian is a doctoral student of School of Computer and Information Technology at Shanxi University, China. His research interests include rough set theory, granular computing and artificial intelligence. He received the M.S. degree in Computers with applications at Shanxi University (2005).

Jiye Liang is a professor of School of Computer and Information Technology and Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education at Shanxi University. His research interests include artificial intelligence, granular computing data mining and knowledge discovery. He received the Ph.D degree in Information Science from Xi'an Jiaotong University. He also has a B.S. in computational mathematics from Xi'an Jiaotong University.

Chuangyin Dang received a Ph.D. degree in operations research/economics from the University of Tilburg, The Netherlands, in 1991, a M.S. degree in applied mathematics from Xidian University, China, in 1986, and a B.S. degree in computational mathematics from Shanxi University, China, in 1983. He is Associate Professor at the City University of Hong Kong. He is best known for the development of the D1-triangulation of the Euclidean space and the simplicial method for integer programming. His current research interests include computational intelligence, optimization theory and techniques, applied general equilibrium modeling and computation. He is a senior member of IEEE and a member of INFORS and MPS.

Wang Feng is a postgraduate of School of Computer and Information Technology at Shanxi University. Her research interests include granular computing and rough set theory.

Wei Xu is a doctoral student of School of Management, Graduate University of Chinese Academy of Sciences Chinese Academy of Sciences, China. His research interests include rough set theory and rough prediction. He received the M.S. degree in Computers with applications at Shanxi University (2006).