Concept Granule-Based Granular Lattice and Application in Knowledge Retrieval

Lin Sun
College of Computer & Information Engineering, Henan Normal University, Xinxiang 453007, P. R. China
Engineering Technology Research Center for Computing Intelligence and Data Mining, Henan Province, China
International WIC Institute, Beijing University of Technology, Beijing 100124, P. R. China
Email: linsunok@gmail.com

Jiucheng Xu, Yangyang Liu
College of Computer & Information Engineering, Henan Normal University, Xinxiang 453007, P. R. China
Engineering Technology Research Center for Computing Intelligence and Data Mining, Henan Province, China
Email: xjc@htu.cn

***Abstract*—In this paper, from the viewpoint of granular computing, information granule and concept granule are introduced to information systems, and then structure of concept granules and their distance, concept granular entropy and significance degree are presented. The concept of concept granule-based granular lattice is proposed in information systems. The representation and operation rules of complete granular lattice are given. Moreover, some important properties and propositions about these concepts presented above are discussed as well, whose performances are shown through an illustrative example. Finally, a conceptual framework of knowledge retrieval with granular lattice using concept granules may help to understand the nature of granules and granular hierarchical structures, and further enlarge the application areas of granular computing.**

***Index Terms*—granular computing, information system, concept granule, granular lattice, knowledge retrieval**

I. Introduction

The notation of fuzzy information granules was first introduced by Zadeh [1]. From then on, the idea of information granules has been widely researched and applied to many fields, such as rough sets, fuzzy sets, evidence reasoning, image retrieval, and so on [2-4]. Granular computing was proposed as a formal terminology by Lin [5]. Granular computing, as a new and rapidly growing paradigm of information processing, has attracted many researchers and practitioners [6-8]. Granular computing is an umbrella term to cover any theories, methodologies, techniques, and tools that make use of information granules in complex problem solving [9-12]. Granular computing as an important tool has found many applications in soft computing, knowledge discovery and data mining, and has obtained some good results [13-17]. Moreover, granular computing is also a new intelligent computing method based on the problem concept space partition [18, 19]. There are usually two aspects to study granular computing: the structure of particles and the calculation of particles [20]. The former is mainly used to solve formation, representation and interpretation of particles. The latter is mainly used to solve the use of particles. At present, granular intension is very clear, but it is diverse in different granular computing models such as fuzzy sets, quotient space, rough sets, concept lattice, etc [8, 9, 20-24]. Granular formation including granulation and representation of granules is the foundation of granular computing. At present, however, there is not a unified framework in granular representation among various fields, which causes that granular computing is difficult to be formally represented and complicated to be computed in applications of intelligent systems. Then, the development of granular computing is largely restricted by representation of granules. Therefore, it is necessary to present new representation and operations of granules and establish efficient granular spaces and granular structure models in information systems.

After a long-term research, granular computing theory has formed a variety of granular representation forms that are closely related to concrete backgrounds. In 1979, Zadeh studied the size of divided class or granule [1]. In 1990, Zhang and Zhang [21] studied granular computing of quotient space, and used quotient sets as granules to compute. In 1998, Lin defined granules by binary relation in the view of neighbourhood in [22] and a series of papers about granular computing. In 2001, Skowron et al. [23] described information granules and its computing. In 2002, Yao and Yao [24] defined a basic granule by logic language. In 2004, Liu and Huang [25] called binary symmetry as basic granules. In 2006, Zhang [26] used Galois connection to describe granule. In 2009, Li [27] gave a four-group formalization representation of granule. The syntactic representation of granules in [28] is a compound formula, so granule is more complicated to be represented, and it is difficult to describe the hierarchical relationship among granules. Because granule has its certain syntax and semantics, when we consider hierarchical structure among granules, if we only take semantics into consideration, its hierarchical relationship cannot be expressed accurately, so we must also consider

doi:10.4304/jsw.9.7.1976-1983

its syntactic hierarchical structure. However the method of grammar representation of granules in [28] is difficult to reflect syntactic levels. Wille [29] established the theory of formal concept analysis. Formal concept analysis is usually called concept lattice. It is a mathematical framework for discovery and design of concept hierarchies from an information system called a formal context [30]. All concepts in a formal context form a concept lattice and can be depicted by a Hasse diagram, where each node expresses a formal concept. The concept lattice is the core structure of data in formal concept analysis, from which a set of objects and a set of attributes are uniquely reflected from each other and the relations between the generalized concept and the specialized concept are described. Recently there are many achievements in the field of formal concept analysis, such as formal concept analysis combined with the theory of rough set [30]. Zhang et al. [31] set up the theory and the approach to attribute reduction of concept lattice in which the formal context is regarded as a special information system, and proposed the judgment theorems of attribute reduction. Yao and Chen [32] studied the relations between FCA and the theory of rough set. Belohlavek [33] extended the concept lattices to fuzzy concept lattices in fuzzy formal context. Fan et al. [34] studied various fuzzy concept lattices and discussed fuzzy inference methods. Ma et al. [35] built the relations between fuzzy concept lattices and granular computing.

The traditional concept lattice theory that analyzes formal concept from extension and intension can well solve the hierarchical relationship among concepts, but it is limited by the binary limitations of the formal context and difficult to be used in the multi-valued intelligent systems. A concept which is the fruit of human thinking includes the essential information about some objects and distinguishes among many other objects. Meanwhile, new concepts are often deduced from accepted concepts, and that is the hierarchical structure. Thus a concept is regarded as an information granule that plays an important role during perception and recognition of people [30]. In formal concept analysis, granulation of the universe of discourse, description of granules, relationship between granules, and computing with granules are issues that need further scrutiny. Then, formal concept analysis naturally has the thoughts of granular computing and shows these thoughts from formation of concept to comparison of relations among concepts and then to exhibition of concept lattice. Since the basic structure of a concept lattice induced from a formal context is the set of object concepts and every formal concept in the concept lattice can be represented as a join of some object concepts, each object concept can be viewed as an information granule in the concept lattice [12]. Granular computing consists of granule, granular layer and granular structure. Granule is the basic unit that realizes knowledge and discovers knowledge from the cognitive perspective [18]. Granular layer distinguishes the exiting problems or knowledge according to a certain division criterion. Granular structure is the set of correlations between granulation grain and granular layer. It reflects the process that people observe, perceive and solve problems from perspective of different granular layers and granules. Concept discovery process of the formal concept analysis reflects the thought of granular computing, and the concept lattice generated from it intuitively reflects the three components of granular computing. Different formal concepts can be called granules, different concept layers can be perceived as granule layers, and the structure of the whole concept lattice reflects structure of its granules. In this paper, based on traditional granular computing theory and combination with the formal concept analysis of the classical concept lattice theory, concept granule is introduced to use information granules to represent concepts, and the uncertainty relationship among granules by using the structure and analytical methods of concept is described at the same time. Then, the concept hierarchical relationship of information systems is obtained by extracting the data structure-granular lattice. This representation method of granules dose not only resolve the hierarchical structure description problem which is difficult to be achieved by traditional granular computing methods, but also for classical concept lattice theory, this method that represents concept by using granules introduces formal concept analysis into information systems and breaks the binary limitations in original formal contexts, gathers the rough set, the concept granule and the granular computing close together, and obtains an extension of the concept lattice theory, thereby expands the application scope of concept lattice and improves its data analysis ability. In this paper, we aim at creating such a solution to solve this problem.

The rest part of this paper is organized as follows. In Section II, some basic concepts are reviewed. In Section III, some concepts of concept granules including their distance, concept granular entropy and significance degree are presented, and then the representation of granular lattice and operation rules are given. Some important propositions and properties are educed as well. In Section IV, a conceptual framework of knowledge retrieval with granular lattice using concept granules is described. Finally, the conclusions and future work are described in Section V.

II. MAIN CONCEPTS

In this section, we will review several basic concepts that are relevant to this paper, such as rough logical formula, granule, formal context, partial ordering relation and concept lattice. Detailed description and formal definitions of the theories can be found in [20, 28, 36-38].

An information system (IS) is a quadruple (U, A, V, f), where U is a finite non-empty set of objects indicating a given universe, A is a finite non-empty set of attributes, V is a union of attribute value domains such that $V = \bigcup_{a\in A} V_a$ for V_a denoting the value domain of attribute a, and f: $U\times A \rightarrow V$ is an information function which associates a unique value of each attribute with every object belonging to U, such that for any $a\in A$ and $u\in U$, $f(u, a)\in V_a$.

Let $IS = (U, A, V, f)$ be an information system, and $a \in A$, $v \in V_a$, (a, v) is an atomic formula, simply denoted by a_v, the rough logical formula can be defined as follows:

(1) a_v is an atomic formula, and atomic formulas are formulas;

(2) If A is an atomic formula, then $\sim A$ is a formula;

(3) If A and B are atomic formulas, then $\sim A$, $A \vee B$, (A), $A \rightarrow B$ are formulas;

(4) According to formulas defined in (1) and (2), formulas consisting of limited number of operations with the connectives $\sim$, $\vee$, $\wedge$, $\rightarrow$, and $\leftrightarrow$ are formulas.

According to above definitions, it is known that in equivalence relation, the objects which have equal attribute values in a certain attribute set can constitute a set, and these objects meet equivalence relation in the attribute set, so we can use this set to define granules.

A formal context can be expressed as (U, A, I), where $U = \{x_1, x_2, \ldots, x_n\}$ indicates the object set (extension), $A = \{a_1, a_2, \ldots, a_m\}$ indicates the attribute set (intension), I is the subset of the binary relation $U \times A$ between objects and attributes. If $(x, a) \in I$, then x has an attribute a, and if $(x, a) \notin I$, then x does not have an attribute a. For a formal context (U, A, I), the following operations [36] on object subset $X \subseteq U$ and attribute subset $B \subseteq A$ are denoted by

$$X^* = \{a \in A \mid \forall x \in X, (x, a) \in I\},$$

$$B^* = \{x \in U \mid \forall a \in A, (x, a) \in I\}.$$

If (U, A, I) is a formal context, A, A_1, $A_2 \subseteq G$, and B, B_1, $B_2 \subseteq M$, then one can obtain the following properties:

(1) $A_1 \subseteq A_2 \Rightarrow A_2^* \subseteq A_1^*$, $B_1 \subseteq B_2 \Rightarrow B_2^* \subseteq B_1^*$;

(2) $A \subseteq A^{**}$, $B \subseteq B^{**}$;

(3) $A^* = A^{***}$, $B^* = B^{***}$;

(4) $A \subseteq B^* \Leftrightarrow B \subseteq A^* \Leftrightarrow A \times B \subseteq I$.

Let (U, A, I) be a formal context, where $X \subseteq U$, $B \subseteq A$. If a dualistic group (X, B) meets $X^* = B$ and $B^* = X$, then we can say (X, B) is a formal concept, simply a concept, where X is called the extension of concept and B is called the intension of concept.

The binary relationship R between set M and N is a set of the ordered binary group (m, n), where, $m \in M$, $n \in N$, namely, $R \subseteq M \times N$. Here "$\times$" is a cartesian product and $(m, n) \in R$, or mRn.

If the binary relation R defined on set H can satisfy the following conditions for any element x, y, $z \in H$, then we say R is a partial ordering relation: reflexivity (xRx), antisymmetry (xRy and $yRx \rightarrow x = y$), and transitivity (xRy and $yRz \rightarrow xRz$).

For a partial ordering relation R, we use "$\preceq$" to express this relation. If xRy, $x \preceq y$ and $x \neq y$, then $x \prec y$. If aRb, $a \prec b$, and there does not exists an element c that meets $a \prec c \prec b$, then a is a metadata of b, that is denoted as $a \prec b$.

Every partial ordering set can be expressed by *a* Hasse diagram. Elements in M are expressed by circles in Hasse diagram. If x, $y \in M$, and they meet $x \prec y$, then the circle of element y will be above the circle of element x, and this two circles are be connected by a line segment.

The concept of formal context (U, A, I) can define the ordering relation among them by relations between hypernotions and sub-concepts:

$$(X_1, B_1) \preceq (X_2, B_2) \Leftrightarrow X_1 \subseteq X_2 (\Leftrightarrow B_1 \supseteq B_2).$$

The partial ordering set of all concepts on (U, A, I) is denoted as $L(U, A, I)$ and called concept lattice. Then "$\preceq$" is a partial ordering relation on $L(U, A, I)$. The definitions of supremum and infimum are as follows:

$$(X_1, B_1) \wedge (X_2, B_2) = (X_1 \cap X_2, (B_1 \cup B_2)^{**}),$$

$$(X_1, B_1) \vee (X_2, B_2) = ((X_1 \cup X_2)^{**}, B_1 \cap B_2).$$

III. Concept Granule-Based Granular Lattice

Though one can gain a better understanding of granular computing within the rough set framework, granular computing within formal concept analysis has not been thoroughly investigated [12]. Rough set theory and formal concept analysis are actually related and often complementary approaches to data analysis. The discovery process in formal concept analysis exactly reflects the thought of granular computing, and the concept lattice generated by it reflects the three components of granular computing. The different formal concepts can be called granules, the different concept layers can be perceived as granular layers, and the whole structure of concept lattice reflects the structure of granules. Once a concept lattice from a formal context is formed, a major task is to find a minimal concept lattice so that it can avoid redundancy but at the same time maintain structure consistency.

Let $IS = (U, A, V, f)$ be an information system. An information granule is denoted by the tuple $(\varphi, G(\varphi))$, where φ refers to the intension of information granule, and $G(\varphi)$ represents the extension of information granule. For any attribute subset $B = \{a_1, a_2, \ldots, a_k\} \subseteq C$. Suppose that $V_{a_i} = \{V_{a_{i,1}}, V_{a_{i,2}}, \ldots, V_{a_{i,k}}\}$ is the domain of feature a_i, and each $V_{a_{i,j}}$ may be viewed as a concept. Then, there must exist $\varphi = \{I_1, I_2, \ldots, I_k\}$ such that $I_i \in V_{a_i}$ is a set of feature values corresponding to B. Then, the intension of an information granule can be denoted by $\varphi = \{I_1, I_2, \ldots, I_k\}$, and the extension can be denoted by $G(\varphi) = \{u \in U' | f(u, a_1) = I_1 \wedge f(u, a_2) = I_2 \wedge \ldots \wedge f(u, a_k) = I_k, a_i \in B, i \in \{1, 2, \ldots, k\}\}$. Here, $G(\varphi)$ describes the internal structure of the information granule.

In an information system $IS = (U, A, V, f)$, U is a finite set of objects, P is a finite set of atomic formulas, atomic formula $a_v \in P$, X, X_1, $X_2 \subseteq U$, B, B_1, $B_2 \subseteq P$, f:$G(P) \rightarrow G(U)$ is the mapping operator from a formulary set to an object set, and g: $G(U) \rightarrow G(P)$ means the mapping operator from an object set to a formulary set. If f meets $f(\emptyset) = U$ and $f(B_1 \cup B_2) = f(B_1) \cap f(B_2)$, and g meets $g(\emptyset) = P$ and $g(X_1 \cup X_2) = g(X_1) \cap g(X_2)$. Then $CGS = (U, P, f, g)$ is called a concept granular space of the IS.

Definition 1. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. If $B \subseteq P$ and $f(B) \subseteq U$, then a concept granule (CG) can be denoted by $CG = (B, f(B))$, where B is the intension of concept granule, and $f(B)$ is the object set that determined by intension B and called the extension of concept granule. If B is a non-empty singleton set, then CG is called an atomic concept granule.

Definition 2. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. If $B_1, B_2 \subseteq P$, $CG_i = (B_i, f(B_i))$ and $CG_j = (B_j, f(B_j))$ are two arbitrary concept granules, and $B_1 \subseteq B_2$, then the CG_i is a parent granule of the CG_j, or the CG_j is a child granule of the CG_i, denoted by $CG_j \prec CG_i$, where "$\prec$" means that they satisfy a partial ordering relation.

Definition 3. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B_i, B_j \subseteq P$, $CG_i = (B_i, f(B_i))$ and $CG_j = (B_j, f(B_j))$ are two arbitrary concept granules. The distance between CG_i and CG_j is defined as

$$DIS(CG_i, CG_j) = | f(B_i \wedge B_j) - f(B_i \vee B_j) |,$$

where $|\cdot|$ denotes the cardinality of a set.

From Definition 3, it can be obtained the following properties immediately.

Proposition 1. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B_i, B_j \subseteq P$, $CG_i = (B_i, f(B_i))$ and $CG_j = (B_j, f(B_j))$ are two arbitrary concept granules. Then the following properties hold

(1) $0 \le DIS(CG_i, CG_j) \le |U|$;

(2) $DIS(CG_i, CG_j) = 0$ if and only if $CG_i = CG_j$;

(3) $DIS(CG_i, CG_j) = |U|$ if and only if $B_i \wedge B_j = \varnothing$;

(4) $DIS(CG_i, CG_j) = DIS(CG_j, CG_i)$.

Proposition 2. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B_i, B_j, B_k \subseteq P$, $CG_i = (B_i, f(B_i))$, $CG_j = (B_j, f(B_j))$ and $CG_k = (B_k, f(B_k))$ are three arbitrary concept granules. If $B_i \prec B_j \prec B_k$, then $DIS(B_i, B_k) = DIS(B_i, B_j) + DIS(B_j, B_k)$.

Proof. The proof is similar to that of Theorem 2 in [39].

Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. It follows from Definition 1 that $\gamma = (\varnothing, U)$ can be called the most rough concept granule. In what follows, how to measure the distance between CG and γ should be considered. It is known that the greater the distance between CG and γ is, the finer the granule is. The smaller the distance between CG and γ is, the coarser the granule is. Then, a concept of concept granule entropy is introduced to measure the distance between CG and γ as follows.

Definition 4. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B \subseteq P$, the concept granular entropy of $CG = (B, f(B))$ is denoted by CGE_{CG}, defined as

$$CGE_{CG} = \frac{1}{|U|} DIS(CG, \gamma) = 1 - \frac{|f(B)|}{|U|}.$$

According Definitions 3 and 4, it can be known that the distance of concept granules can not only measure the difference degree between concept granules in the same universe, but also describe the thickness-variational degree of concept granules caused by the change of intension in the concept granular space. For instance, in a concept granular space $CGS = (U, P, f, g)$, $B \subseteq P$, removing an atomic formula a_v from B of the concept granule $CG = (B, f(B))$, a new concept granule $CG' = (B - \{a_v\}, f(B - \{a_v\}))$ can be obtained. The distance of CG and CG' is used to measure the variation degree of granules that is caused by the intension of concept granule CG becoming small. So the importance degree of atomic formula a_v to concept granule CG can be measured by this distance. It follows that the greater the distance is, the more important a_v is to CG.

Definition 5. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B \subseteq P$ and $a_v \in B$, if $CG = (B, f(B))$ and $CG' = (B - \{a_v\}, f(B - \{a_v\}))$, then the significance degree of a_v with respect to CG is denoted by $SIG(CG, a_v)$, defined as

$$SIG(CG, a_v) = | CGE_{CG} - CGE_{CG'} | = \frac{| f(B) - f(B - a_v) |}{|U|}.$$

From Definition 5, it is easy to obtain the following properties and proposition.

Property 1. $0 \le SIG(CG, a_v) \le 1$.

Property 2. Any atomic formula a_v is necessary to a concept granule CG if and only if $SIG(CG, a_v) > 0$.

Proposition 3. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B \subseteq P$, $CG = (B, f(B))$, to make $a_v \in B$, and a_v in B is unnecessary for CG, if $f(B) = f(B - \{a_v\})$. Otherwise a_v is necessary for CG.

Corollary 1. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. Then $SIG(CG, a_v) = 0$ if and only if $f(B) = f(B - \{a_v\})$ for any $B \subseteq P$ and $a_v \in B$.

Proof. It can be derived directly from Definition 5 and Proposition 3.

Definition 6. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B, R \subseteq P$, $a_v \in R$, $CG = (B, f(B))$, and $CG' = (R, f(R))$, the CG' is called a reduct of CG if and only if $f(B) = f(R)$ and $SIG(CG, a_v) > 0$.

Definition 6 states an approach to reduct in concept granular space using concept granules, which helps to simplify representation of granule in information systems.

When an intelligent information system is expressed by $IS = (U, A, V, f)$, the knowledge concept of the intelligent information system is formalized by its concept granules, and granular hierarchical structure chart of concept granules. That is to say, the data structure that is produced by a partially ordered set $(CGS, \prec)$ can be called a granular lattice. In the structure of granular lattice, the operation rules between concept granules are strictly defined in order to make the operations among concepts, and can be moved according to standardized rules. This formalized method can describe structure of intelligent information system and its relationships more clearly and intuitively.

Definition 7. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. For any $B \subseteq P$, $GC = (B, f(B))$, if $B = g(f(B))$, then the CG is called a complete concept granule, denoted by $\tilde{C}G$.

Definition 8. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$. If $G = \cup\{(B, f(B)) \mid B \in P, B = g(f(B))\}$, then there is an unique partially ordered set (CGS, $\prec$) corresponded with it, and the partially ordered set has a unique supremum and infimum. The lattice structure produced by this partially ordered set is called granular lattice, denoted by $\tilde{G}(U, P, f, g)$.

Thus, it follows from Definition 8 that there is always a unique maximum sub-concept (infimum) and a unique minimum parent concept (supremum). In the following, for an information system $IS = (U, A, V, f)$ with its concept granular space $CGS = (U, P, f, g)$, we give two operation rules in its granular lattice.

Definition 9. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$ and granular lattice $\tilde{G}(U, P, f, g)$. For any B_i, $B_j \subseteq P$, $\tilde{C}G_i = (B_i, f(B_i))$, and $\tilde{C}G_j = (B_j, f(B_j))$, the upper bound operation denoted by "$\wedge$" and the lower bound operation denoted by "$\vee$" are defined respectively as

$$\tilde{C}G_i \wedge \tilde{C}G_j = (g(f(B_i \cup B_j)), f(B_i) \cap f(B_j)),$$

$$\tilde{C}G_i \vee \tilde{C}G_j = (B_i \cap B_j, f(g(f(B_i) \cup f(B_j)))).$$

From Definitions 7, 8 and 9, it is easy to obtain the following property.

Property 3. Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$ and granular lattice $\tilde{G}(U, P, f, g)$. For any B_i, $B_j \subseteq P$, $\tilde{C}G_i = (B_i, f(B_i))$, and $\tilde{C}G_j = (B_j, f(B_j))$, then the following properties hold

(1) $\tilde{C}G_i \in \tilde{G}$;

(2) $\tilde{C}G_j \in \tilde{G}$;

(3) $\tilde{C}G_i \wedge \tilde{C}G_j \in \tilde{G}$;

(4) $\tilde{C}G_i \vee \tilde{C}G_j \in \tilde{G}$.

(5) The concept granule obtained by limited number of operations among the complete concept granules by using "$\vee$" and "$\wedge$" is still a granular lattice.

Property 3 states actually an approach of granule lattice extraction. For any concept granule, one can find the corresponding complete concept granule by iteration approach. For an information system, all complete concept granules are obtained, and then the parent and the child concept are linked by a segment, then the granular lattice Hasse diagram corresponding to an information system can be got according to the principle that parent concept is on the above and sub-concept is on the below.

Example 1. Let $IS = (U, A, V, f)$ be an information system, shown in Table I, where $U = \{x_1, x_2, x_3, x_4\}$, $A = \{a, b, c\}$.

TABLE I.
AN INFORMATION SYSTEM

U	a	b	c
x_1	1	0	0
x_2	0	1	2
x_3	0	0	1
x_4	1	1	2

For the IS in Table I, we assume that its concept granular space $CGS = (U, P, f, g)$. Let $P = \{a_v | a \in A, v \in V_a\} = \{a_0, a_1, b_0, b_1, c_0, c_1, c_2\}$, and $CG = (B, f(B)) = \{\{b_1\}, \{x_2, x_4\}\}$ be the atomic concept granule that meets the value 1 on the attribute b. The complete concept granules corresponding to it can be calculated easily. According to the above operation rules, let $U_1 = f(B)$, then $U_2 = U_1 \cup f(B) = f(B) = \{x_2, x_4\}$, $B_1 = g(U_2) = g(\{x_2, x_4\}) = \{b_1, c_2\}$, $B_2 = B_1 \cup g(U_2) = \{b_1, c_2\}$, $U_3 = f(B_2) = f(\{b_1, c_2\}) = \{x_2, x_4\}$. Because $U_3 = f(B_2)$, $B_2 = \{b_1, c_2\} = g(U_3)$, (B_2, U_3) is the corresponding complete concept of the CG. Thus, all of the complete concept granules of IS are obtained according to a certain order. That is, $\{\varnothing, U\}$, $\{\{a_1, b_0, c_0\}, \{x_1\}\}$, $\{\{a_0, b_1, c_2\}, \{x_2\}\}$, $\{\{a_0, b_0, c_1\}, \{x_3\}\}$, $\{\{a_1, b_1, c_2\}, \{x_4\}\}$, $\{\{b_1, c_2\}, \{x_2, x_4\}\}$, $\{\{a_0\}, \{x_2, x_3\}\}$, $\{\{a_1\}, \{x_1, x_4\}\}$, $\{\{b_0\}, \{x_1, x_3\}\}$, $\{P, \varnothing\}$. Hence, a granular lattice Hasse diagram of IS is shown in Fig. 1.

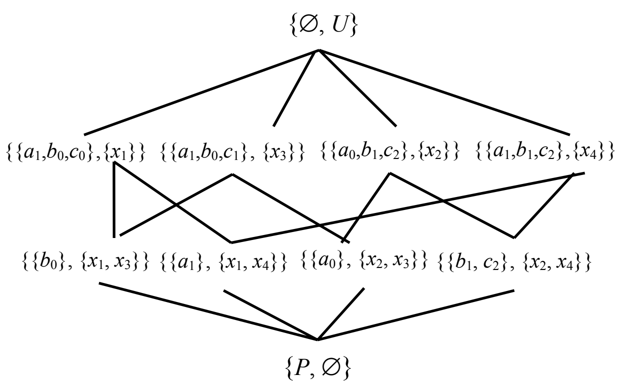

Figure 1. Granular lattice Hasse diagram of IS in Table I.

IV. KNOWLEDGE RETRIEVAL WITH GRANULAR LATTICE USING CONCEPT GRANULES

Knowledge representation in a structured way is consistent with human thoughts and is easily understandable [18, 20]. Knowledge retrieval system for similar characteristics from a large number of databases is a process of dealing with uncertain information. One of the key features of knowledge retrieval is that knowledge are visualized in a structured way so that users could get contextual awareness of related knowledge and make further retrieval [40]. Then, our knowledge retrieval method based on hierarchical structure of granular lattice using concept granules concentrates on how to provide and use knowledge in more convenient ways. Therefore, this objective of our proposed method of problem solving is to match query requirements of users with selected relevant characteristic concept granules of information systems using corresponding knowledge selection algorithms, and then return the search results which are required knowledge to the users.

Here, the problem to be solved is how to similarly match user's query characteristics and data in user characteristic granule base according to a certain algorithm, and a group of candidate results that meet a certain similarity will be sorted according to optimal and returned to users. Query characteristics are formalized into a concept granule, feature granules are formalized into an information system, and then the similarly matched process of characteristics and data in user characteristic granule base will be converted into a traversal process of granular hierarchical structure of feature granules of granular lattice.

Let $IS = (U, A, V, f)$ be an information system with its concept granular space $CGS = (U, P, f, g)$ and granular lattice $\tilde{G}(U, P, f, g)$. From Definitions 1 and 8, it is known that user's query characteristic granule $\tilde{C}G = (B, f(B))$ for any $B \subseteq P$ be projected to user characteristic granule base and partition user characteristic granule base into three parts as follows:

(1) The first part is the objects that fully meet query characteristics granule $\tilde{C}G$. Namely, this part is a certain query result.

(2) The second part is the objects that do not meet query characteristics granule $\tilde{C}G$. This part is not query result and do not involve the solution,

(3) The third part is the objects that do not completely meet query characteristics granule $\tilde{C}G$. This part is the approximate query results that need to be sorted, the approximate results namely associate with search target but partial knowledge form a lot of branches from coarse to fine.

The process of knowledge retrieval is going along correct branch from coarse to fine according to the user's interaction information. Then, our granular structure discovered and constructed by a knowledge retrieval model should be in multilevel and multiview. On the basis of constructing granular lattice model and improving the granular applicability, this paper gives the initial model of knowledge retrieval with granular lattice using concept granules. Fig. 2 gives the process of construction for knowledge retrieval with granular lattice using concept granules.

Here, characteristics of knowledge retrieval based on the structure of granular lattice are that before precise retrieval it formalizes user search conditions to characteristics granules by domain knowledge and user context. The characteristics granule is only granular description, has not actual connotation. When we project this characteristics granule to feature base, concept granular space will be formed, which can be comprehended as a subsystem of original information systems. All these relevant data information form a structure of concept granules of different granularity levels. When users browse concept granular space of concerned data information, main operations of knowledge retrieval system include amplification and reduction. Amplification operation is used to formalize knowledge granules of particle size smaller, provide accurately browsing for query. In view of models of concept granules into consideration, amplification operation converts granularity size larger quotient sets into granularity size smaller quotient sets. Reduction operation is used to formalize knowledge granules of particle size larger, provide coarse browsing for user query. It is convenient for users to seize integral knowledge structure with granular lattice. In specific operation process, amplification and reduction operations can be realized by increasing or decreasing constraint attributes. User main retrieval tasks are preliminary realized by selecting nodes in concerned visual granular structure by interactive way. Although the basic principle when system retrieve concrete granular knowledge nodes is similar to information retrieval system, user knowledge structure of granular lattice provided by knowledge retrieval system helps users to execute more effective retrieval when users cannot certain needed information.

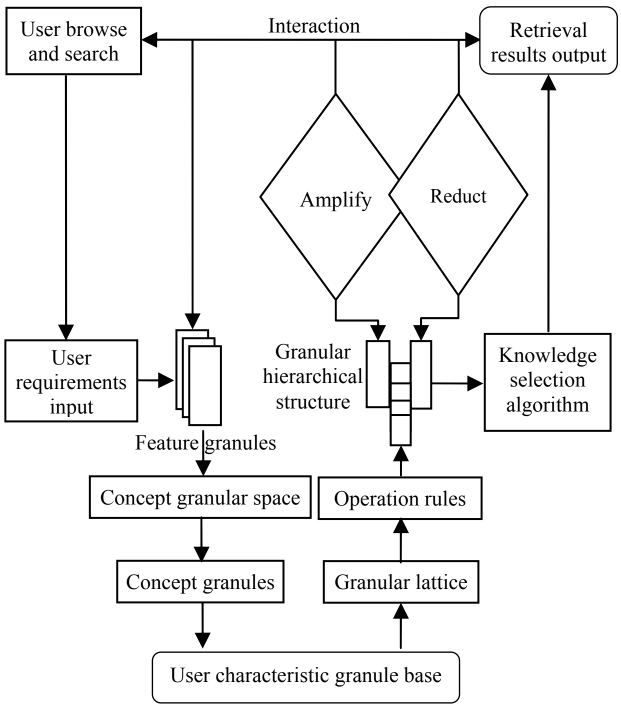

Figure 2. Knowledge retrieval with granular lattice using concept granules

V. Conclusions

At present, granular computing has become an important tool in many application fields, such as soft computing, knowledge discovery and data mining, and some good results have been obtained. However, there is not a unified framework in the representation of granules among various fields. Because granule has its certain grammar and semantics, its syntactic hierarchical structure must be considered. And formal concept analysis can completely map data sets to hierarchical structure models according to binary relation of objects and attributes in data sets. So this paper references the existing results to study concept lattice theory and granular computing theory, combines this two theories above, re-expresses granules by the concept of concept granule, and describes hierarchical structure of granule through extracting granular lattice. Then their some important properties, propositions, and operation rules are developed. Finally, from the viewpoints of information processing, a conceptual framework of knowledge retrieval system with granular lattice using concept granules and its process are described. Furthermore, the applications of granular lattice in knowledge discovery and data mining need to be studied to speed up the developing of granular computing.

ACKNOWLEDGMENT

We are highly grateful to the anonymous reviewers, referees and Editor-in-Chief for their valuable comments and hard work.

This work was supported by the National Natural Science Foundation of China (Nos. 60873104, 61370169), the Key Project of Science and Technology Department of Henan Province (No. 112102210194), the Science and Technology Research Key Project of Educational Department of Henan Province (Nos. 12A520027, 13A52 0529), and the Education Fund for Youth Key Teachers of Henan Normal University.

REFERENCES

[1] L. A. Zadeh, "Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", *Fuzzy Sets and Systems*, vol. 90, pp. 111−127, 1997.

[2] L. Sun, J. C. Xu, and Y. Tian, "Feature selection using rough entropy-based uncertainty measures in incomplete decision systems", *Knowledge-Based Systems*, vol. 36, pp. 206−216, 2012.

[3] L. Sun, J. C. Xu, S. Q. Li, X. Z. Cao, and Y. P. Gao, "New approach for feature selection by using information entropy", *Journal of Information and Computational Science*, vol. 8, pp. 2259−2268, 2011.

[4] L. Sun, J. C. Xu, Z. A. Xue, and J. Y. Ren. "Decision degree-based decision tree technology for rule extraction", *Journal of Computers*, vol. 7, pp. 1769− 1779, 2012.

[5] 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. Heidelberg: Springer, 1998, pp. 107−120.

[6] L. Sun and J. C. Xu, "A granular computing approach to gene selection", *Bio-Medical Materials and Engineering*, vol. 24, pp. 1307−1314, 2014.

[7] L. Sun, J. C. Xu, J. Y. Ren, and T. H. Xu, "Granularity partition-based feature selection and its application in decision systems", *Journal of Information and Computational Science*, vol. 9, pp. 3487−3500, 2012.

[8] L. Sun, J. C. Xu, Y. W. Hu, and L. N. Du, "Granular space-based feature selection and its applications", *Journal of Software*, vol. 8, pp. 817−826, 2013.

[9] J. T. Yao, A. V. Vasilakos, and W. Pedrycz, "Granular computing: perspectives and challenges", *IEEE Transactions on Cybernetics*, vol. 43, pp. 1977−1989, 2013.

[10] L. Sun and J. C. Xu, "Feature selection using mutual information based uncertainty measures for tumor classification", *Bio-Medical Materials and Engineering*, vol. 24, pp. 763−770, 2014.

[11] L. Sun, J. C. Xu, Z. A. Xue, and L. J. Zhang, "Rough entropy-based feature selection and its application", *Journal of Information and Computational Science*, vol. 8, pp. 1525−1532, 2011.

[12] W. Z. Wu, Y. Leung, and J. S. Mi. "Granular computing and knowledge reduction in formal contexts", *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, pp. 1461−1474, 2009.

[13] J. C. Xu and L. Sun, "Knowledge entropy and feature selection in incomplete decision systems", *Applied Mathematics & Information Sciences*, vol. 7, pp. 829−837, 2013.

[14] J. C. Xu and L. Sun, "A new knowledge reduction algorithm based on decision power in rough set", *Transactions on Rough Sets*, vol. 12, pp. 76−89, 2010.

[15] L. Sun, J. C. Xu, and Y. P. Song, "Information quantity-based decision rule acquisition from decision tables", *Journal of Convergence Information Technology*, vol. 7, pp. 57−67, 2012.

[16] L. Sun, J. C. Xu, and L. J. Zhang, "Approaches to knowledge reduction of decision systems based on conditional rough entropy", *International Journal of Advancements in Computing Technology*, vol. 3, pp. 129−139, 2011.

[17] Y. H. Chen and Y. Y. Yao, "A multiview approach for intelligent data analysis based on data operators", *Information Sciences*, vol. 178, pp. 1−20, 2008.

[18] L. Sun, N. Zhong, and J. C. Xu, "Granularity-based user-centric multi-strategies and application in knowledge retrieval", *Journal of Computer*, vol. 8, pp. 2101−2109, 2013.

[19] Y. Zeng, N. Zhong, Y. Wang, et al., "User-centric query refinement and processing using granularity based strategies", *Knowledge and Information Systems*, vol. 27, pp. 419−450, 2011.

[20] L. Sun, J. C. Xu, C. Wang, T. H. Xu, and J. Y. Ren, "Granular computing-based granular structure model and its application in knowledge retrieval", *Information Technology Journal*, vol. 11, pp. 1714−1721, 2012.

[21] L. Zhang and B. Zhang, "The quotient space theory of problem solving", *Fundamenta Informaticae*, vol. 59, pp. 287−298, 2004.

[22] T. Y. Lin, "Granular computing on binary relations II: rough set representations and belief functions", in *Rough Sets in Knowledge Discovery*, A. Skowron and L. Polkowski, Eds., Heidelberg: Springer, 1998, pp. 121−140.

[23] A. Skowron, J. Stepaniuk, and J. F. Peters, "Extracting patterns using information granules: A brief introduction", in *Joint JSAI 2001 Workshop Post-Proceedings of New Frontiers in Artificial Intelligence*, T. Terano, Y. Ohsawa, T. Nishida, et al., Eds. Lecture Notes in Computer Science, vol. 2253, 2001, pp. 359−363.

[24] Y. Y. Yao and J. T. Yao, "Granular computing as a basis for consistent classification problems", in *Proceedings of PAKDD'02 Workshop on Toward the Foundation of Data Mining*, 2002, pp. 101−106.

[25] Q. Liu and Z. H. Huang, "G-Logic and its resolution reasoning", *Chinese Journal of Computers*, vol. 27, pp. 865−873, 2004.

[26] L. Zhang, *The Research on Model of Knowledge Discovery Based on Granularity and Concept Lattice*, Anhui: Hefei University of Technology, 2006.

[27] H. Li, “Granule, granular set and granular system”, in *Proceedings of IEEE International Conference on Granular Computing*, Nanchang, China, 2009, pp. 340−345.
[28] Q. Liu and Q. Liu, “Granules and applications of granular computing in logical reasoning”, *Chinese Journal of Computer Research and Development*, vol. 41, pp. 546−551, 2004.
[29] R. Wille, “Concept lattices and conceptual knowledge systems”, *Computers & Mathematics with Applications*, vol. 23, pp. 493−515, 1992.
[30] G. F. Qiu, J. M. Ma, H. Z. Yang and W. X. Zhang, “A mathematical model for concept granular computing systems”, *Science China* (*Information Sciences*), vol. 53, pp. 1397−1408, 2010.
[31] W. X. Zhang, L. Wei, and J. J. Qi, “Attribute reduction theory and approach to concept lattice”, *Science in China Series F Information Sciences*, vol. 48, pp. 713−726, 2005.
[32] Y. Y. Yao and Y. H. Chen, “Rough set approximations in formal concept analysis”, *Transaction Rough Sets*, vol. 5, pp. 285−305, 2006.
[33] R. Belohlavek, “A note on variable threshold concept lattices: threshold-based operators are reducible to classical concept-forming operators”, *Information Sciences*, vol. 177, pp. 3186−3191, 2007.
[34] S. Q. Fan, W. X. Zhang, and W. Xu, “Fuzzy inference based on fuzzy concept lattice”, *Fuzzy Sets and Systems*, vol. 157, pp. 3177−3187, 2006.
[35] J. M. Ma, W. X. Zhang, Y. Leung, and X. X. Song, “Granular computing and dual Galois connection”, *Information Sciences*, vol. 177, pp. 5365−5377, 2007.
[36] W. X. Zhang, Y. Y. Yao, and Y. Leung, *Rough Set and Concept Lattice*, Shanxi: Xi'an Jiaotong University Press, 2006.
[37] Z. W. Cao, J. Qian, W. M. Zhang, and S. Deng, “A FCA-based approach for concept similarity computation”, *Fuzzy Systems and Mathematics*, vol. 22, pp. 155−162, 2008.
[38] J. C. Xu, L. Sun, and Q. Q. Zhang, *Theories and Methods of Granular Computing and its Uncertainty Information Measures*, Beijing: Academic Press, 2013.
[39] J. C. Xu, W. L. Cheng, and J. L. Shi, “Concept granular and its distance computing based on information system”, *Journal of Guangxi University* (*Nat. Sci. Ed.*), vol. 34, pp. 667−671, 2009.
[40] Y. Y. Yao, Y. Zeng, N. Zhong, and X. J. Huang, “Knowledge retrieval”, in *Proceedings of the International Conference on Web Intelligence*, Silicon Valley, USA, 2007, pp. 729−735.

Lin Sun works at College of Computer & Information Engineering, Henan Normal University. He is currently a Ph.D. Candidate at Beijing University of Technology. He received his B.S. and M.S. degree in Computer Science and Technology, Henan Normal University in 2003 and 2007, respectively. His main research interests include rough set, granular computing, bioinformatics, and data mining.

Jiucheng Xu is currently a Professor at College of Computer & Information Engineering, Henan Normal University. He received his B.S. degree in Mathematics, Henan Normal University in 1986, the M.S. degree and the Ph.D. degree in Computer Science and Technology, Xi’an Jiaotong University in 1995 and 2004, respectively. His main research interests include granular computing, rough set, data mining, intelligent information processing, image retrieval, and bioinformatics.