A new formal representation of granules based on features

Hong Li[1, 2], Xiaoping Ma[1], Zhenghua Xin[2*]

[1]*School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221008, Jiangsu*

[2]*Laboratory of Intelligent Information processing, School of Information Engineering, Suzhou University, Suzhou 234000, Anhui*

Received 1 May 2014, www.tsi.lv

Abstract

In order to provide the unified representation of granules under Granular Computing, this paper studied the granules' basic meaning, descriptions and relationships between the features of granules, and described that the four elements of granules were closely related to granular features. Then the paper presented four-tuples formal representation on data levels based on granular features. The representation includes the object set, feature set, relationship sets and constraints set. Then the representations of several special granules were presented. They showed the unity and advantage of the representation. At last, this paper gave a specific example. By the example, the formal representation has important significance in methodology to solve the granular representation on data levels based on granular features well. And this method is conductive to solve problems and study of granular computing theory.

Keywords: granule, feature, formalization, representation, granular computing

1 Introduction

The world is composed of granules. People understand the world in the way of the granule. In this sense, the granule is the concept applied to practice accompanied with human's existence. It is a natural way to deal with the problem of human daily life [1]. The granule concept should be devoted to L.A.Zadeh who published a paper in 1979 [2]. In that paper, information granule is described to be a proposition.

In the fuzzy sets, quotient space, rough sets, concept lattice or other different models of granular computing, the meaning of granule is more and more clear but different. The different meanings are not conducive to the development of granular computing. The granular computing is based on the formation of granule which includes granule formalization and granule representation [3]. Due to different representations of granule, there is no unified basis and premise for the theory of granular computing. Therefore, a unified form and meaning interpretation of the granule will greatly facilitate the study in granule computing. And it is beneficial to establish the granular computing in a unified theory framework.

The granular computing theory has a long history. The granular computing theory formed a variety of granular representations relative to the specific background. In 1979, L. A. Zadeh studied the size of classified category or the granules [2]. In 1990, Bo Zhang studied the quotient space granular computing. He took quotient set as granules and used granules to calculate. In 1998, T. Y. Lin defined the granule through the binary relations at the view of the neighbourhood [5] in the subsequent series of literatures. In 2001, Skowron described the information granule [6, 7] and its calculation. In 2002, Y. Y. Yao defined a fundamental granule through logical language [8]. In 2004, Qing Liu took the binary symmetric theory as the fundamental theory [9]. In 2006, Liang Zhang used Galois connection to describe granule [10]. In 2009, Hong Li proposed the four-tuples formal representation of granules [11].

From the paper [1] to the paper [10], all took the fuzzy sets, quotient space, rough set or concept lattice as background in an information system. And they have significant limitations. The formal four-tuples representation of the granule has a certain abstract meaning in [11]. But its application is not convenient. This paper gives the meaning, feature, and the other formal representation of granules based on the above work. In this paper, there is not the background of fuzzy sets, quotient space, rough set or concept lattice. Therefore, this method is conducive to solve the problems.

2 The feature of granules

2.1 DESCRIPTION OF THE GRANULAR FEATURES

Any object has a large number of basic features, which are independent of other objects. People abstract the common feature of some object and call the feature is part of this class object. Any granule has three basic properties. They are the intrinsic, external and environmental properties [12]. It is believed that granular features are the combination of these three basic properties and the formed values. These three basic

[*] ***Corresponding author*** e-mail: begin0000@qq.com

properties include internal feature, external feature and environmental feature. The formed values are in the form of the property-value's binary pair. In addition to basic features, there are other features. The granular identification can be increased. There is Definition 1 and Definition 2 to understand the attribution characteristics of granule.

Definition 1 (Granular Feature): A special logo of a granule is called granular feature to distinguish it from others. It describes the unit granular ingredients, which include both property and property value. The property is called feature items and the property value is called feature values.

Definition 2 (Granular feature's description): The description of the formal granular feature is a three tuples: $GF=<id, A, V>$. The *id* represents the granular unique identifier. The *A* represents the attribute sets or feature set of the granule. It is an n-dimensional vector (a_1, a_2, a_3, …, a_n) and n is a natural number. The *V* represents the set of attribute value or the feature value. They are *n m*-dimensional vectors (b_{i1}, b_{i2}, b_{i3}, …, b_{im}). The *m* and n are both natural numbers. And m is feature value with the maximum number in n attributes (i=1,2,…,*n*).

In general, the description *GF* of granular feature can be reduced to tuple form. If *F* is the feature set of granule, i.e. $F=<A, V>$, then $GF=<id, F>$.

Any granular feature can be divided into the common and unique feature. The common feature contains all common properties in problem domain. It is common feature of all granules in problem domain. The unique feature is the property of part granules in problem domain.

Any feature of granules can be divided into two classes which are the essential and phenomena feature. The essential feature belongs to a particular granule which is the granule that contains all the inherent feature of individual objects. The inherent properties include the structure properties, behaviour, functions, etc. They are common features set of these class individual objects. The phenomenon feature is the external feature shown by the appearance of the class individual objects or the individuals.

Similarly, granular feature can be divided into atomic and complex feature. Atomic fee221q`q23ature is no longer divided to other basic feature. The type of atom is real, character and Boolean and other basic data types. A complex feature is constructed from the other defined complex features based on atomic features.

In the decision-making problem, the granular feature can be divided into condition and decision-making feature. The feature describing the problem condition is called the conditional feature. And the feature describing the problem of making decisions is called the decision-making feature.

Feature value is an instance of feature item. The possible values reflect the variability of feature item itself in a different environment. Feature items with the feature value are cases of the relationship. They describe the general and special relationship between the feature item and feature value.

Definition 3 (The feature space of problem domain): All the characteristics are given by the related problem domain and structural forms. The structural forms are constructed by all the relationships between features. All the characteristics are named as the feature space of problem domain. The feature space of problem domain is expressed as $\Omega=<FS, FR>$. The FS is the set of problem domain feature. And FR is a collection of feature relationship.

The feature space in the problem area usually uses the form of the tree structure to represent. It is shown as in the Figure 1. Each node of feature tree in the feature space represents the feature item. It can be identified by a unique feature mark. The feature leaf node represents atom feature. The non-leaf node represents complex features. The feature tree has only one root node to express the root of the problem domain feature space. The root node represents a concept or system. In the feature tree, the parent-child tree nodes are connecting directly. That represents the parent feature and child feature. In this way, the granule can be divided into atomic granule and complex granule corresponding to the granular feature level.

2.2 THE RELATIONSHIP AND EXPRESSION OF GRANULAR FEATURE

There are relations between any two things. There are various relationships between feature and feature of granules. The relationship between a pair of features usually has different type. A type of relationship can be applied to a different feature set, and format different semantic information. Therefore, it is particularly important to determine the type of relationship and study the association between various types of mechanisms.

In general, there are four types of relationships between two features. The first is parent-child relationship. The feature item of the parent relationship decides which sub-feature is necessary item to constitute the feature. For example, the "age" feature is such feature. And it decides "juvenile", "youth", "middle", "old" and so on. The second is separation-integration relationship. It is feature aggregation and decomposition. The complex feature can be formed through aggregating features. Simple feature can be formed through decomposing the feature and ultimately form atomic characteristics. The third is dependency relationship. That is an instance value of a set of features. On the other hand, it determines the unique or multiple instance values of another set of features. The "Age" and "title" instance features can decide "professional", "health" and other features instance value. The fourth is association relationship. It exists between the features of the semantic association. The timing relationship between operating features, the combination relationship between describing features, the mapping relationship between operating and describing features and so on are the examples. At some view, the parent-child relationship and separation-integration relationship together form the combination relationship. This type of relationship between features can be divided into the combination, dependency and association categories.

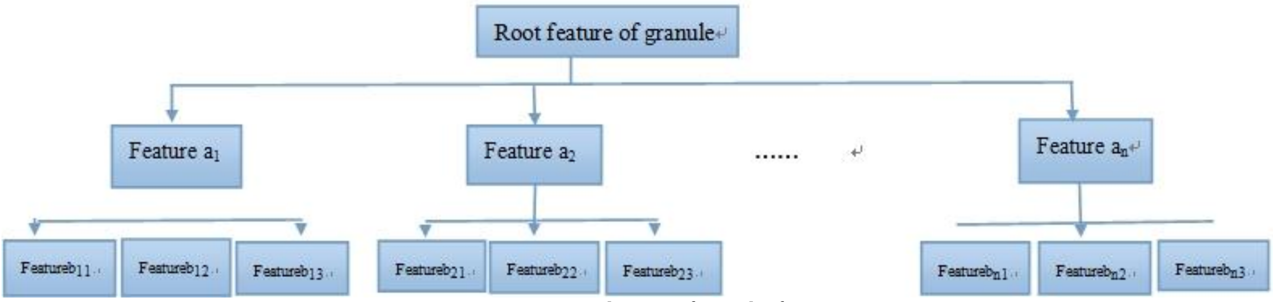


FIGURE 1 Tree diagram of granular features

Here study the expressions of these three type relationships.

Definition 4 (Feature set is derivable): Suppose that the two granules' features are described as $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. If all the granules' features in F_2 can be defined by some granules' feature in F_1 and the features' existing condition in F_1 has no direct relationship whether the features in F_2 exists or not , it is called that feature set F_2 is derivable from the feather sets F_1. It is denoted as $F_1 \subseteq F_2$.

Proposition 1: suppose that F_1 and F_2 are the two features of the same problem domain Ω set in feature space, and the feature set F_2 is derivable from the feature set F_1. Any feature in F_2 is the parent feature of all features that is used to define the feature self in F_1. And all the features in F_1 are the sub feature of F_2.

Definition 5 (Feature set is dependent): Suppose that the two granular features are $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. If $F_1 \subseteq F_2$, we call that there exist dependent relationships between the feature sets F_2 and F_1. F_2 exists dependent of F_1. It is expressed as $F_1 \Rightarrow F_2$.

Definition 6 (Feature sets are equivalent): Suppose that the description of the two granules feature are $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. If $F_1 \Rightarrow F_2$ and $F_2 \Rightarrow F_1$, we say that the feature set F_2 and the feature set F_1 are equivalent. Or it is interdependent between the feature set F_2 and the feature set F_1. It is denoted as $F_1 \equiv F_2$.

Definition 7 (Feature set is independent): Suppose that the description of the two granules feature are $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. If there not exists $F_1 \Rightarrow F_2$ and also not exists $F_2 \Rightarrow F_1$, then the feature sets F_2 and F_1 are independent each other. It is denoted as $F_1 \neq F_2$.

Proposition 2: Suppose F_1 and F_1 are two independent feature sets in the same problem domain feature space Ω. Then any one feature in F_1 and any one feature in F_2 are characteristic independent on each other.

Definition 8 (Feature relationships): Suppose that the description of the two granules feature are $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. The formula $R_F=<GF_1, GF_2>=\{<a_1,a_2>| a_1 \in F_1, a_2 \in F_2\}$ denotes the relations between GF_1 and GF_2.

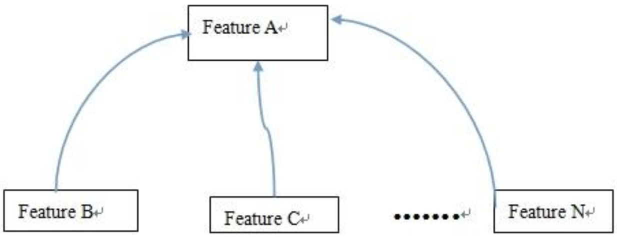
If the feature sets F_2 and F_1 are independent separately, then there are no relations or exists empty relation between GF_1 and GF_2. It is denoted as $R_F=\Phi$.

If there is dependent relation between the feature sets F_2 and F_1, then there exists not empty relation between GF_1 and GF_2. It is denoted as $R_F \neq \Phi$.

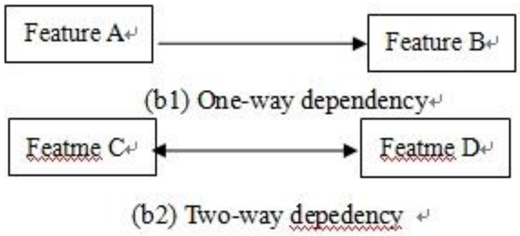
Definition 9 (Feature relationship's representation): Suppose that the description of the two granules feature are $GF_1=(id_1, F_1)$ and $GF_2=(id_2, F_2)$ in the same problem domain feature space Ω. And there is no empty relation between GF_1 and GF_2. There are the following inferences.

- If $F_1 \subseteq F_2$, there exist a_1, $a_2 \in F_1$ and $a_3 \in F_2$. And there are the features of father and son between a_3 and a_1, a_2. However, there are brother characteristics between a_1 and a_2. There exists $a_3=a_1 \cup a_2$, then we say that there exists combination relations between GF_1 and GF_2. That is denoted as $GF_1 \rightarrow cGF_2$.
- If there exist $X \subseteq F_1$ and $Y \subseteq F_2$ to make $X \Rightarrow Y$ established in Ω, then there is one-way dependent relation between GF_1 and GF_2. It is expressed as $GF_1 \rightarrow iGF_2$. Especially, if there exist $X \Rightarrow F_1$ and $Y \Rightarrow F_2$, and $X \Rightarrow Y$ and $Y \Rightarrow X$ are simultaneously established in Ω, then there are bidirectional dependent relations:
- or interdependent relations. It is expressed into $GF_1 \leftrightarrow iGF_2$.
- If there exist $X \subseteq F_1$ and $Y \subseteq F_2$, and there is a constrained condition $J=f(X, Y)$ between the feature combination X and the feature combination Y, we call there is a link between GF_1 and GF_2. It is expressed as $GF_1 \rightarrow a\ GF_2$. Generally, suppose that there is multiple granules features' description $GF_i=(id_i, F_i)$ and $X_i \subseteq F_i (i=1,2,\ldots)$ in the same problem domain feature space Ω. If there is a constrained condition $J=f(X_1,X_2,\ldots)$ among many granules feature combination $J=f(X_1,X_2,\ldots)$, then there is multiple correlation between $GF_i(i=1,2,\ldots)$. It is denoted as $\rightarrow a(GF_1, GF_2, \ldots)$.

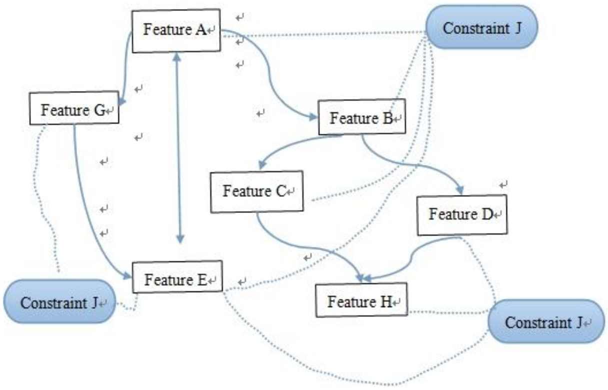
The combination, dependency and association relationships between granule features are shown in the Figure 2. The granule feature and relationships between granule features provide an important prerequisite for their formal representations.



(a) combination relationship



(b) dependency relationship



(c) association relationship

FIGURE 2 Relationships of granular features

3 The formal representation of granules

Since the granule is the model which is a shared objective of the various components [13]. In the process of the applying granular computing to solve real problems, the complex relationships of large original problem must be formed. Therefore, the large original problem needs to be decomposed into a number with complete information granule. The necessary merging operation and other treatments have been done according to the specific requirements and background. This requires the definition of internal restraint mechanism. The mechanism used the same granule structure of formal representation. The definition must include the complete information of granules.

Some references analysed the granular objects, features, relationships, and the state of the four basic elements [13]. Describing a granule should be used the four elements. The four elements are the object, the feature, the relationship and the state. As objects, relations and states three elements are closely related with the feature elements, using these four elements as formal representation of granule is essentially feature-based representation.

3.1 THE FORMAL REPRESENTATION OF GRANULE

Information science uses four levels [14, 15]. These levels are in the following order, from the data, the information, the knowledge to the wisdom. They describe the different abstraction levels of the human-centred information procession. Each level needs to use the appropriate computer system to achieve. The information is commonly referred to general information. It includes the form, the carrier or the content [16]. In addition, general information and data is usually confused. Data is the information carrier. And the performance of information is data. The information in this paper indicates the general information with the indiscriminate data.

In the data layer, people use database management system (DBMS) to store and retrieve the structured data. The basic unit of storage and retrieval is the entity containing some values of the object attribute. The information storage and retrieval is the specific form of granule. The paper only discusses the formal representation of the granule in the data level here.

Definition 10 (Granule formal representation): Suppose the given feature space problem domain is Ω. A granule of this problem domain can be formally defined as the following four bytes in the following quadruple form: $G=(O, F, R, J)$

The G is used to describe a granule. The O represents those studied individual objects set of granules. It is called the granules' extension or the domain. It can be a common set and also be a fuzzy set. In some aspects, it can be limited. It also can be infinite. For example, in a certain interval, the F is all the feature set of granules. It is named as the granules' connotation, $F=(F_1, F_2, \ldots, F_n)$. $F_i=\{ F_{i1}, F_{i2}, \ldots, F_{im} \}$ $(1 \leqq i \leqq n, m \geqq 0)$ is the i-feature of each individual object in O. The F_{ij} is the j-feature of each individual object of granules in the i-feature. The F embodies the internal, the external and environmental features, etc. The R represents a collection of all relationships in granules, including the sequence structure, topology structure, and graph structure of various elements or individuals and so on. It also includes all kinds of relationships between features and various relations. The various relations first establishes between individual objects and the characteristics. Then they and features output all kinds of relationships. Of course, the output relationships are included in the F. The J is a set of constraints, which regulates the legal status of granules and timing sequence and so on. And it is used to represent the static, dynamic and unity properties of granules. J includes the static constraints and dynamic constraints. The static constraints refer to the determining formulas of granules. They provide all the legal status of this kind of granules. Dynamic constraints include the timing constraints and real-time constraints. These constraints are the time constraints, distance constraints or the rules limitation.

Definition 11 (Relationships between individual objects): Suppose the feature space of given problem domain is Ω, a granule of the problem domain can be formally defined as $G=(O, F, R, J)$, let $R_O=\{<o_1,o_2>| o_1, o_2 \in O\}$ indicate the relationship between the individual object o_1 and the individual object o_2. If there is a feature $a \in F$ to make the establishment of $a(o_1) = a(o_2)$, then we say that there is not distinguishable relationship between o_1 and o_2 on the feature a. It is denoted as $R_O \neq \Phi$. Otherwise, we call there exist distinguishable relationship between o_1 and o_2 on the feature a. It is denoted as $R_O = \Phi$.

According to definition 8, definition 10 and definition 11, there exists $R = R_F \cup R_O$. On the basis of definition 10, there are the following three propositions with the hypothesis that (O, F, R, J) is the primary granule.

Proposition 3: The granule $(O', F', R', J')$is atomic granule, if and only if $O' \subseteq O \wedge |O'|=1$, $R'=\Phi$, $J'=\Phi$.

Atomic granules have the smallest size. Because atomic granules do not contain any constraints, they can use every feature item of features at utmost. This makes the absolute multiplexing degree reach the maximum. But as the feature's degree is the smallest. The efficiency of reusing is the lowest.

Proposition 4: The granule (O', F', R', J') is elementary granule, if and only if $F' \subset F$, $R' \subset R$, $J' \subset J$.

The basic granules are implemented by non atomic features, while atomic granules can also be considered as a special kind of basic particle ($R'=\Phi$). The two are essentially the same. They both are mapping that a feature directly to an independent particle. The basic granule's size is larger than the size of atomic granules. Thus the mining efficiency is improved. But because of too many demanding characteristics, the cost of mining massive data increases.

Proposition 5: The granule (O', F', R', J') is original granule, if and only if $O' \subseteqq O$, $F' \subseteqq F$, $R' \subseteqq R$, $J' \subseteqq J$.

3.2 THE UNITY ANALYSIS OF GRANULAR FOUR-TUPLES REPRESENTATION

The four-tuples (O, F, R, J) of definition 10 can be well unified fuzzy sets, quotient space and rough set etc. It can be used as a model of the granular computing

When taking the granules (O, F, R, J) as the whole study corresponding to the original problem space is the original granule. At this time, the element of granule can be in the elements of the original problem. And it can also be the sub-granule composed of elements of the original problem space.

The relationship R is an equivalence relation at the macro view of granule. A raw granule is composed of all subsets or sub-granules corresponding to a certain equivalent division of the original problem. This granule description corresponds to a quotient space of the original problem space.

When the relationship R is an equivalence relation, a sub-granule is a subset of the original problem. This granule description corresponds to the granule of a rough set.

The relationship R is fuzzy relation at the fuzzy view of granule. It naturally corresponds to the granule of a fuzzy set.

Therefore, the granule representation form with four-tuples (O, F, R, J) of the definition 10, is convenient to converse at different views, such as macro and micro, the overall and partial, precise and vague. It is consistent with the features of human intelligence. It is also convenient to absorb some relatively mature theory, extract their common characteristic, and benefit to establish a unified theoretical framework for granular computing.

3.3 THE ADVANTAGES OF GRANULAR FOUR-TUPLES REPRESENTATION

Compared with the definition of granule in the literature [4-11], using the four-tuples (O, F, R, J) of definition 10 to describe the granule has strong advantages.

1) the granule representation of this paper reflects the unity of granule. The granules defined in literature [5-10] take fuzzy sets, quotient space, rough set or concept lattice as background in information systems. It has significant limitations. In addition, the formal representation of the four-tuples for granule has a certain abstract meaning in [11], but its application is not convenient. According to the above section B, the granular representation of this paper gets rid of the background of fuzzy sets, quotient space, rough set or concept lattice. It can be well unified the existing quotient space, rough sets and fuzzy sets etc. It is a typical granular computing model.

2) it reflects the concrete and abstract unity characteristic. In the four-tuples (O, F, R, J), some specific individual objects, the specific feature and conditions together form a granule. Therefore, the granule is the set of entities, which reflects the specific nature of granules. At the same time, this granule representation solves problems based on the user's needs, and establishes a way to perceive the objective world. In this computing model, the objective world is granulated into granules. Because the virtual world in users' opinion is composed of the granules, it can support and help users understand the objective world. And this reflects the abstract characteristic of granule.

3) it reflects the unification of connotation and extension well. In the four-tuples (O, F, R, J), O indicates the individual object set of the granules. It is the extension of granule. It is also a universe term. It can be a normal set or a fuzzy set. F is the set of all the granule

features. It is the content of granules, $F = (F_1, F_2, ..., F_n)$. $F_i = \{F_{i1}, F_{i2}, ..., F_{im}\}$ ($1 \leqq i \leqq n, m \geqq 0$) is the i-feature of each individual object in O. F_{ij} is the i-j-feature of each individual object of the granule ($1 \leqq i \leqq n, 0 \leqq j \leqq m$). F reflects the internal and external, environmental and other basic features of the granule. Compared with the granule definitions in the literature [4-11], the granule representation of this paper reflects preferable unification of the connotation and extension.

- Fourth, it reflects the unification of static and dynamic feature. In the four-tuples (O, F, R, J), J is a set of constraints. It provides the legal statuses of granules and the timing between the legal statuses with such constraints. It includes the static constraints and dynamic constraints. Static constraints determine the granule style. They provide all the legal status of such granules. Dynamic constraints include granule timing constraints, real-time constraints between the legal statuses. It may be time constraints or distance limitations, the rules limitations and so on. Compared the static feature with the definition of granule in the literature [4-10], it reflects the unification of static and dynamic feature well.
- Fifth, it reflects the unification of the original granule and the basic granules. When any problem is studied, the original granule takes the whole object as a granule. Atomic granule does not contain any constraints, with minimum size of the granule. It usually contains only a feature of each item. The basic granule is achieved by the non-atomic feature. Atomic granule is a special form of basic granules. The propositions from the proposition 3 to 5, it gets the original, the basic or atomic granule according to different selection of O, F, R, J in the four-tuples (O, F, R, J). They reflect the unification of the original granule and the basic granule. But the granular representation forms in the literature [4-11] do not have this feature.

4 A case study

This paper gives a case to illustrate the formalization of granule. The Table 1 is a table of building block information to study the changes of granule.

TABLE 1 Building block information table

Element	Colour	Shape	Size	Stability
x_1	red	triangle	big	stability
x_2	red	triangle	big	stability
x_3	yellow	circle	small	instability
x_4	yellow	circle	small	instability
x_5	blue	square	big	stability
x_6	red	circle	middle	instability
x_7	blue	circle	small	instability
x_8	blue	square	middle	instability

Suppose 8 blocks form a set of objects O. It is noted that $O=\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$. Each building block has four kinds of feature items. That feature set is denoted by A. Each item has different feature values. The value of each feature set is denoted by V_A. A and V_A constitute a building block feature. It is denoted that $F=A\times V_A$, $A=$ {colour, shape, size, stability}, V_A = {values of each feature item}. There is such relationship, $R: F\rightarrow O$. The objects set O is decomposed into a smaller subset according to the given constraints J to select the feature item values. The original granule problem of building blocks can be expressed as $G=(O, F, R, J)$.

For example, according to the colour feature, these building blocks can be granulated into R_1 = {red, yellow, blue} three tablets. Then all the building blocks in the red colour constitute the kernel granule $X_1=\{x_1, x_2, x_6\}$. The yellow building blocks constitute the kernel granule $X_2=\{x_3, x_4\}$. The building blocks in the blue colour constitute the kernel granule $X_3=\{x_5, x_7, x_8\}$. Then $O/R_1=\{X_1, X_2, X_3\}=\{\{x_1, x_2, x_6\},\{x_3, x_4\},\{x_5, x_7, x_8\}\}$ according to the colour granulation. Similarly, R_2= {triangle, square, circle}, R_3= {large, medium, small} indicates the size. R_4 = {stable, unstable} indicates the stability. The equations are as follows according to the shape granulation, the size granulation and the stable granulation respectively.

$O/R_2=\{Y_1, Y_2, Y_3\}=\{\{x_1, x_2\},\{x_5, x_8\},\{x_3, x_4, x_6, x_7\}\}$,

$O/R_3=\{Z_1, Z_2, Z_3\}=\{\{x_1, x_2, x_5\},\{x_6, x_8\},\{x_3, x_4, x_7\}\}$,

$O/R_4=\{T_1, T_2, T_3\}=\{\{x_1, x_2, x_5\},\{ x_3, x_4, x_6, x_7, x_8\}\}$.

If one wants to find the "large and triangular" building blocks, he/she can use $R_2\cap R_3$ to find the required result granule. Similarly, if one wants to find the "big or triangular" building blocks, he can use $R_2\cup R_3$ terms.

When mining massive data, it is complex and difficult to extract the valuable information. And it needs large calculation by operating data sets directly. This method presented in this paper can get knowledge in the higher level of abstraction. It may be more universal. It can construct the corresponding concept or identity-level tree by the abstract feature value in a given data table. Multiple abstraction levels can be achieved to solve granular computing problem. At the same time, the original data set can also be horizontal compression to reduce data size, and improve efficiency for solving problems. At this point, more abstract levels of granule-based knowledge representation can be introduced. Such granular representation can provide an important prerequisite for category theory based granular computing method.

The granular triple-tuples representation based on knowledge level would be study in other paper.

5 Conclusion

This paper provides a uniform granular representation for granular computing, which studies the feature of the granule. It illustrates that the four elements of granule are closely related to granular feature. It presents a new feature-based four-tuples representation of granule. By comparison, this paper's granular four-tuples representation reflects the unity of granule, the specific and abstraction, the connotation and extension, static and dynamic, origination and basic granule. The formal representation has important methodology. It is a good solution to the data level feature-based granular representation problem. It is conducive to solve problem and granular computing theory. Of course, this representation is only a preliminary exploration, which still needs further test for its effectiveness. In addition, the abstract representation of granules needs further study. The introduction of category theory provides an important prerequisite for the granular calculation.

Acknowledgment

This work is supported by National Nature Science Foundation under Grant 60974126, Key Project of Anhui Education Department Nature Science Foundation under Grant KJ2012A263, Anhui Provincial Natural Science Foundation under Grant 10040606Q64.

References

[1] Liu Q, Sun H, Wang H F 2008 The Present Studying State of Granular Computing and Studying of Granular Computing Based on the Semantics of Rough Logic *Chinese Journal of Computers* **31**(4) 543-55
[2] Zadeh L A 1979 Fuzzy sets and information granularity *Advances in Fuzzy Set Theory and Applications North-Holland Amsterdam* 3-18
[3] Zadeh L A 1998 Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems *Soft Computing Berlin Springer-Verlag* 23-5
[4] Zhang B, Zhang L 2007 Problem solving theory and application (The 2nd Edition) *Tsinghua University Press* 4-14
[5] Lin T Y 1998 Granular Computing on Binary Relations II Rough Set Representations and Belief Functions *Rough Sets In Knowledge Discovery Physica -Verlag* 121-40
[6] Skowron A 2001 Toward Intelligent Systems Calculi of Information Granules Proceedings of International Workshop on Rough Set Theory and Granular Computing *Bulletin of International Rough Set Society* **5**(1/2) 9-30
[7] Skowron A, Stepaniuk J, Peters James F 2001 Extracting Patterns Using Information Granules *Proceedings of International Workshop on Rough Set Theory and Granular Computing Bulletin of International Rough Set Society* **5**(1/2) 135-42
[8] Yao Y Y, Yao J T 2002 Granular Computing as a Basis for Consistent Classification problems *Proceedings of PAKDD'02 Workshop on Foundations of Data Mining CIICM* 2002(5) 101-6
[9] Liu Q, Huang Z H 2004 G -Logic and Its Resolution Reachilding *Chinese Journal of Computers* **27**(7) 865-73
[10] Zhang L 2006 The Research on Model of Knowledge Discovery Based on Granularity and Concept Lattice *Hefei University of Technology*
[11] Li H 2009 Granule, Granular Set and Granular System *Proceedings of the IEEE 4th International Conference for Granular Computing Los Alamitos California IEEE Computer Society CPS* 340-5
[12] Yao Y Y 2006 Granular computing for data mining *The SPIE Conference on Data Mining, Intrusion Detection, Information Assurance and Data Networks Security Kissimmee USA* 1-12
[13] Li H 2010 Research on Granule and Its Characteristics *Journal of Suzhou University* **25**(11) 8-11 29
[14] Bellinger G, Castro D, Mills A 2010 Data, Information, Knowledge and Wisdom [EB/OL]. *www.systems-thinking.org/dikw/dikw.htm*
[15] Jing N N, Cheng J Y 2005 Data, Information, Knowledge and Wisdom *Information Science* **23**(12) 1786-90
[16] Zhu X L 2000 Fundamentals of applied information theory *Tsinghua University Press* **1-2** 4888-900

Authors

Hong Li, born in 1965, Anhui province, China

Current position, grades: professor China University of Mining and Technology, Xuzhou in Jiangsu province.
University studies: master degree in computer science and technology from Hefei University of Technology, China.
Scientific interest: intelligent computing, algorithm design and analysis, data mining.
Publications: [1] Further Study on Granulating Thinking. Journal of Suzhou university. 2013.28(5). [2] Research on four-element model of granular computing. Computer Engineering and Applications. Computer Engineering and Applications. 2013.49(4). [3] Formal Method of Granular Synthesis Based on Category Theory. Computer Engineering. 2011.37(14).

Xiao-ping Ma

Current position, grades: professor, China University of Mining and Technology, Xuzhou in Jiangsu province.
University studies: doctor degree from China University of Mining and Technology.
Scientific interest: control theory and application, computer application.
Publications: 25
Experience: The executive dean of School of information and electrical engineering, CUMT

Zhenghua Xin

Current position, grades: lecture of Suzhou University.
University studies: master degree from Hebei University of Technology, China.
Scientific interest: cloud computing, internet of things.
Publications: 17