

Granular Computing Based Ontology Learning Model and Its Applications

Hongcan Yan, Feng Zhang, Baoxiang Liu

College of Science, North China University of Science and Technology, 46 Xinhua Road, Tangshan, Hebei, China

Emails: yanhongcan@ncst.edu.cn

605313106@qq.com

liubx5888@126.com

Abstract: *The cores of granular computing are the granule, the granule layer and the granule structure. In this paper the concepts “Ontology granule” and “Compatible granule” were defined, applying the granular computing ideas and an ontology model, an Ontology granular set; and thus ontology tree generation algorithms were proposed. These algorithms produce an initial ontology granular set with a compatible class, and extend the other ontology granules by the vector of intension IG, building a lattice hierarchy and a conception tree model of ontology with the vector of relation RG. The empirical research of the traditional Chinese medicine ontology shows that these algorithms are correct and efficient, and provide a good technical way for ontology learning.*

Keywords: *Granular computing, ontology learning, ontology granule, compatible granule, ontology tree.*

1. Introduction

In the paper “Fuzzy Sets and Information Granularity” Zadeh has considered that the concept of information granule exists in many fields, but it is different in form in the various fields. Then “granular mathematics” was renamed to “granular computing” by Lin, and thus it was abbreviated as “GrC”. Lin published a series of papers on granular computing and neighbourhood systems. In recent years, with the development of granular computing research, many local and international scholars have joined the field of granular computing research.

Fuzzy sets, rough sets and quotient space theory are the three calculation models in granular computing theory [1], and they all discuss the granularity computing problem within the framework of set theory. The quotient space theoretical model defines the hierarchical quotient space chain structure by natural projection. The hierarchical structure in the fuzzy quotient space establishes a

corresponding transformation by cutting the relations in the fuzzy equivalence relation. Granularity partition in the rough set theory can be controlled through an increase or reduction of attributes. The mutual transformation of the concept granules in the concept lattice theoretical model can be achieved altering the intension of the concept. Although these methods of transformation are different, there is still a common feature they share – they all emphasize the necessity to demonstrate the important properties of problem solving in the process of transformation on different granule levels. The main characteristic of granular computing is that the solution of the same problem can be freely communicated between the different granular layers. Due to this advantage, the people can use the granular computing method to efficiently solve complex problems.

In the field of artificial intelligence, ontology was first defined by Neches and others, as the basic terms and relations comprise the vocabulary of a topic area, as well as the rules for combining the terms and relations to define an extension to the vocabulary. It is clear that ontology is the set of related concepts in a specific domain and the various relations within the set. In a rich semantic ontology, a conceptual model of the tree is formed on the basis of the concept and its hierarchy, defined by the semantic relation.

Therefore, with the help of the concept hierarchy, concepts with similar or related definitions constitute sets, which is the perspective of a granular computing model.

Ontology similarity calculation has been a key problem in the application of ontology. Its theoretical basis are mainly matching the technology of the figure and tree [2], similar processing technology of a pattern in the database, and similar processing technology in the first-order predicate logic, such as machine learning technology [3]. The purpose of the research into ontology similarity or its application is mainly reflected in four aspects:

(1) Ontology merging (Merge) is a process that merges multiple ontologies into one ontology.

(2) Ontology mapping (Mapping) extends the content of ontology or communicates information among various ontologies.

(3) Modular ontology (Modula) can be split into small modules according to the paradigm of software engineering, with each module being an independent ontology.

(4) Ontology building (Building) requires the participation of domain experts, and since the manual construction of ontology takes time and energy, it is a complicated systematic work of engineering. During the process of building ontology, the experts tend to violate the bias errors, and a time dynamic update is particularly difficult to achieve. Therefore, the researchers have put forward ontology learning, an automatic or semi-automatic building ontology method. If the existing similar ontology is used to build new ontology, the workload can be greatly reduced. Thus the first three kinds of technology are often applied to building ontology.

This paper uses a granular computing model to a layer and transforms the granules under the formal context of domain knowledge in order to form a lattice

structure of an ontology granular set. The lattice structure is then transformed into a tree structure. At the end of this process, the paper constructs an ontology model that provides a new method and strategy for auto-building and mapping of ontology.

2. Ontology description based on an information granule

The basic component of granular computing is mainly comprised of three parts: a granule, a granule layer and a granule structure, with the granule being the basic element of the granular computing model. A granule can be considered as a set of individual elements described by internal attributes, and as a whole described by its external attributes. It is essential that the granule exists only in specific environment. All the granules together constitute a granule layer according to an actual demand of the granulation rule. Granules of the same layer inside often have similar natures or functions. The internal structure of a granule layer, namely the relation between the granules, refers to the domain structure constituted by all granules on the granule layer. The relations of all granule layers make a relational structure, called a granule structure [4]. In [5] the authors prove that the granule in the rough set, quotient space and concept lattice can be described by an abstract granule of the granular space, allowing the granules to be unified in different ways, and eventually additional ideas will be provided to formulate a formal definition of granular computing.

2.1. Information granule description of ontology

Nicola Guarino defined the ontology conceptualization as $C = \langle D, W, R \rangle$, where D is a domain; W is a state set of relevant things in the domain; R is the concept relation set in the domain space $\langle D, W \rangle$. The relation here refers to the extension relation, being a certain state of the elements in the domain D . If D is represented by the related core concepts and instance objects in the domain, and W – by the attribute set of concepts, ontology can be described as the formal context of domains. Thus, the formal context of the Traditional Chinese Medicine dyspnea cases can be shown in Table 1.

Table 1. Formal context of the Traditional Chinese Medicine dyspnea cases

Case	Dyspnea cases	Phlegm (color)	Chills and fever	Body of tongue	Tongue coating (color)
1	Dyspnea and tachypnea	yellow	high grade fever	yellow tongue	yellow
2	Cough and dyspnea	yellow	low grade fever	pale tongue	white
3	Weakness dyspnea	non-phlegm	low grade fever	yellow tongue	red
4	Dyspnea and tachypnea	yellow	aversion to cold	pale tongue	yellow
5	Asthma	white	low grade fever	red tongue	white
6	Cough and dyspnea	yellow	tidal fever	yellow tongue	yellow
7	Cough and dyspnea	yellow	high grade fever	red tongue	white
8	Dyspnea and tachypnea	white	tidal fever	pale tongue	white
9	Weakness dyspnea	white	aversion to cold	red tongue	red
10	Weakness dyspnea	white	low grade fever	pale tongue	white

Ontology is more concerned with the intension of the concept relation itself, namely semantics. Therefore, for better description of the ontology information and

implicit semantic information, ontology is divided into two parts: the description of the external part EO (External Ontology) and the description of the internal part IO (Internal Ontology). EO is the ontology instance corresponding to the formal context, as shown in Table 1. IO is the ontology concept corresponding to the formal context, and it provides a detailed description of various kinds of relations, as shown in Table 2.

Table 2. Formal context of the ontology IO

Concept	a_1	a_2	...	a_n
x_1	1	2	...	4
x_2	3	1	...	1
...
x_n	2	4	...	1

In Table 2: 1 represents the relation of “Part of”; 2 represents the relation of “Kind of”; 3 represents the relation of “Instance of”; 4 represents the relation of “Attribute of”, and we can also define these figures in other relations; $\{a_1, a_2, \dots, a_n\}$ represents an attribute set.

Definition 1. Ontology Information Table (OIT). Quaternion $\langle U, A, V, f \rangle$: U is the set of domain concepts and objects instance, also called theory of the domain; $A = \{a_1, a_2, \dots, a_n\}$ is the set of attributes; $V = \{v_{a_1}, \dots, v_{a_n}\}$ is the set of attribute values; $f: U \times A \rightarrow V$ is an information function, and an attribute value of each object in U . After reprocessing Table 1, we obtain Table 3. Obviously, the IO and EO formal context of ontology is a table of multi-valued attribute.

Table 3. The new formal context

Instance	a_1	a_2	a_3	a_4	a_5
x_1	4	2	3	3	3
x_2	2	2	1	1	1
x_3	1	0	1	3	2
x_4	4	2	4	1	3
x_5	3	1	1	2	1
x_6	2	2	2	3	3
x_7	2	2	3	2	1
x_8	4	1	2	1	1
x_9	1	1	4	2	2
x_{10}	1	1	1	1	1

As shown in Table 3: $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$, $A = \{\text{dyspnea, cases, phlegm, chills and fever, body of tongue, tongue coating}\} = \{a_1, a_2, a_3, a_4, a_5\}$, $V = \{0, 1, 2, 3, 4\}$.

Definition 2. Intuitive granule G of ontology. In granular computing, $M(\psi)$ and $m(\varphi)$ are respectively the extension and intension of the information granule [6, 7]. In this paper the binary vectors $\langle M(\psi), m(\varphi) \rangle$ are defined as intuitive granule of ontology, also called Concept Granules and Information Granules. As shown in Table 3, they are called intuitive granules of ontology. $G_1 = (\{x_1, x_4, x_8\}, \{a_1\})$, $G_2 = (\{x_2, x_6, x_7\}, \{a_1\})$, $G_3 = (\{x_1, x_2, x_4, x_6, x_7\}, \{a_2\})$, $G_4 = (\{x_1, x_4\}, \{a_1, a_2\})$, $G_5 = (\{x_2, x_6, x_7\}, \{a_1, a_2\})$.

2.2. Ontology granule and ontology tree

Intuitive granule is the basic characteristic of the elements and concepts of

ontology. The concept of ontology is not isolated; the relation parameters must be attached in graining ontology.

Definition 3. Ontology Granule (OG). The triple $\langle EG, IG, RG \rangle$ represents an ontology granule. EG is the set of concepts or instances and reflects the extension of the ontology granule; IG is the granule centre corresponding to the set of each attribute value and it reflects the intension of the ontology granule; RG is the presentation of the semantic relations of an intuitive granule and reflects a certain property of classification.

Definition 4. Compatible ontology granule. If two objects, or intuitive granules and ontology granules, conform to a compatible relation τ , i.e., the close degree meets the threshold ρ given by compatible parameters, they are called a compatible object or compatible granule. The set constituted by the extension of all compatible objects and compatible granules is called a compatible ontology granule, therein $\rho \in [0,1]$.

Definition 5. Ontology granular set. All ontology granular sets conforming to a certain compatible relation τ are called ontology granular sets.

Definition 6. The relation of the ontology granule. Assuming two ontology granules: $OG1=(eg1, ig1, rg1)$, $OG2=(eg2, ig2, rg2)$, if $eg1 \subseteq eg2$, then we consider that the ontology granule OG2 includes OG1 on concept extension, and the granularity of the ontology granule OG1 is smaller than the granularity of the ontology granule OG2, or the granularity of the ontology granule OG1 is finer, denoted by $OG1 \leq OG2$.

The rough-fine of intuitive granular intension can also reflect the inclusion relation existing between the concept granules. Definition 2, the intension of ontology granular OG5 includes that of ontology granular OG2, which demonstrates that the granularity of OG2 and OG3 is rougher than that of OG5.

Definition 7. Ontology Tree (OT). The ontology tree is the concept hierarchy model of the ontology, and it is also a tree representation of the ontology granular set. Assuming two ontology granules: $OG1=(eg1, ig1, rg1)$, $OG2=(eg2, ig2, rg2)$, if $OG1 \leq OG2$, we call OG2 the super-granule of the concept granule OG1. If there is no ontology granule OG3, and the relation conforms to $OG1 \leq OG3 \leq OG2$, and $OG1 \neq OG3$, $OG2 \neq OG3$, then we consider OG2 to be a direct super-granule of the concept granule OG1, namely the parent granule, and OG1 is the child granule of the concept granule OG2. All granules conforming to that relation constitute the tree-form ontology granular lattice; i.e., the ontology tree.

The root node of the ontology tree is the ontology granule whose extension is the domain. The other child-nodes are the ontology granules formed through a compatibility relation and in the coherent space of the power set. The parent and child nodes of the tree conform to a rough-fine relation of the ontology granule.

3. Generation of an ontology tree based on granular computing

OIT is expressed as $\langle U, A, V, f \rangle$, and $|U|=m$, $|A|=n$, the u -th individual has an attribute value $[v_{u_1}, v_{u_2}, \dots, v_{u_n}]$; the extension vector of the ontology granule is EG $[u_1, u_2, \dots, u_m]$; the intension vector of the ontology granule is IG $[v_1, v_2, \dots, v_n]$, and

$u_i \in U, i=1, 2, \dots, m, v_j \in V, j = 1, 2, \dots, n$. The characteristics of the numeric information table define a compatible relation with respect to how close the degree is calculated. Among the results of the last step, all the ontology granules conforming to the threshold ρ constitute the ontology granular set.

3.1. Calculation of the ontology granule

The ontology granule triples $\langle EG, IG, RG \rangle$, the intension IG and the semantic relations RG are calculated from the extension EG of the ontology granule. The calculating process is respectively performed by the functions Fun_IG(EG) and Fun_RG(EG).

Function Fun_IG(EG)

Input: The extension vector EG of ontology granule

Output: The corresponding intension vector IG

```

Fun_IG(EG) {
  IG[v1, v2, ..., vn] = [0, 0, ..., 0]; // Initialize the intension of the ontology
  granules
  for all u ∈ EG do
    IG[v1, v2, ..., vn] = IG[v1, v2, ..., vn] + vu1, vu2, ..., vun];
    // Accumulate the corresponding attribute values
  IG[v1, v2, ..., vn] = IG[v1, v2, ..., vn] / |EG|;
  return IG; }

```

Function Fun_RG(EG)

Input: The extension vector EG of the ontology granule

Output: The corresponding semantic relations RG

```

Fun_RG(EG) {
  min = max = 0;
  for all u ∈ EG do
    { if (vu1 + vu2 + ... + vun) > max
      then max = vu1 + vu2 + ... + vun;
      if (vu1 + vu2 + ... + vun) < min
      then min = vu1 + vu2 + ... + vun;
    }
  RG = [min, max];
  Return RG; }

```

3.2. Calculation of a compatible ontology granular set

The function of a close degree τ and the threshold ρ of the compatible degree are given. The initialized ontology granular set SG0 is calculated by Algorithm 1, and then the other compatible ontology granules are determined by Algorithm 2.

Algorithm 1. Fun_SG0(τ, ρ)

Input: Ontology information table OIT (the number of objects is m , the number of attributes is n). The compatible function is τ and the threshold ρ .

Output: Initialized ontology granular set SG0

Algorithm 2. Fun_SG(SG0)

Input: Initialized ontology granular set SG0

Output: Ontology granular set SG3.3. The generation algorithm of an ontology tree based on a rough-fine granule and weighted tree.

Ontology is a tree structure made of a concept and instance. This paper firstly regards the second and third parameters of the ontology triples as weights of tree nodes, and secondly combines the rough-fine relation of a granule, the construction algorithm of the concept lattice hierarchy [8-10] and the construction method of the weighted tree to construct an ontology granule lattice, and eventually generates the ontology tree.

```

Fun_SG(SG0) {
  E is the set constituted by extension of all the ontology granules in SG0;
  CG={};
  Do {
    For all  $u \in E$  do {
      New_EG is the object that meet compatible function  $\tau$  and the
      threshold  $\rho$ ;
      If New_EG  $\notin E$  then
        CG=CG $\cup$ New_EG;
      Else New_EG={};
    } while New_EG!={};
    For all EG  $\notin$  CG do { // Extend a new granule
      New_IG=Fun_IG(EG);
      New_RG=Fun_RG(EG);
      OG=( New_EG, New_IG, New_RG)
      SG0=SG0 $\cup$ OG;
    }
    SG=SG0;
  }
  return SG;
}

```

Algorithm 3. Fun_OTree()

Input: Ontology granular set SG

Output: The set T of the root node and the child node of the ontology tree

```

Fun_OTree() {
   $T = \{t\}$ ; // Ontology granule  $t$  whose extension is the whole domain.  $U$ 
  is the root node of the tree.
  Do {
    For all OG  $\leq t$  do {
      // Find out all the child granules of  $t$  in ontology granular set SG
      if (OG.RG  $\subseteq$   $t$ .RG) then
        Arrange the left and right sub-tree from small to large according to
        OG.RG and OG.IG,
        SG=SG-OG;
        T=T $\cup$ OG;}
    }
  } while SG={};
  return T; }

```

4. The application of an automated construct ontology model

This paper takes the data in Tables 1 and 3 as an application example. In order to simplify the calculation, the paper deletes the last two objects and the last attribute, and keeps eight objects and four attributes, namely,

$$m = |U| = 8, n = |A| = 4, |V| = 16,$$

then the compatible relation used is as follows:

$$\tau(u_1, u_2) = 1 - \frac{\sum_{i=1}^n |x_i - y_i|}{16}, \quad u_1, u_2 \in U, \quad x_i \text{ is the } i\text{-th attribute value of the object } U_1, \text{ and } y_i \text{ is the } i\text{-th attribute value of the object } U_2.$$

This paper uses compatible parameters $\rho = 0.8$ and determines the compatible ontology granular extension by Algorithm 1:

$$\begin{aligned} EG(1) &= \{X_1, X_4, X_6, X_7\}, EG(2) = \{X_2, X_5, X_6, X_7\}, EG(3) = \{X_3\}, EG(4) = \{X_1, X_4, X_8\}, \\ EG(5) &= \{x_2, x_5, x_8\}, EG(6) = \{X_1, X_2, X_6, X_7\} = EG(7) = \{X_1, X_2, X_6, X_7\}, \\ EG(8) &= \{X_4, X_5, X_8\} \end{aligned}$$

There are seven granules in all. The initialized ontology granular set is as follows:

$$\begin{aligned} SG_0 &= \{(\{X_1, X_4, X_6, X_7\}, \{3, 2, 3, 2.24\}, \{9, 12\}), (\{X_2, X_5, X_6, X_7\}, \\ &\{2.5, 2, 2.5, 1.75\}, \{6, 11\}), (\{X_3\}, \{1, 0, 1, 3\}, \{5, 5\}), (\{X_1, X_4, X_8\}, \\ &\{4, 1.67, 3, 1.67\}, \{8, 12\}), (\{x_2, x_5, x_8\}, \{3, 1.33, 1.33, 1.33\}, \{6, 8\}), \\ &(\{X_1, X_2, X_6, X_7\}, \{2.5, 2, 2.25, 2.25\}, \{6, 12\}), (\{X_4, X_5, X_8\}, \\ &\{3.67, 1, 2.33, 1.33\}, \{7, 11\})\} \end{aligned}$$

The ontology granular set can be determined through the expansion operation of Algorithm 2:

$$\begin{aligned} SG &= \{(\{X_1, X_4, X_6, X_7\}, \{3, 2, 3, 2.24\}, \{9, 12\}), (\{X_2, X_5, X_6, X_7\}, \{2.5, 2, 2.5, \\ &1.75\}, \{6, 11\}), (\{X_3\}, \{1, 0, 1, 3\}, \{5, 5\}), (\{X_1, X_4, X_8\}, \{4, 1.67, 3, 1.67\}, \{8, \\ &12\}), (\{x_2, x_5, x_8\}, \{3, 1.33, 1.33, 1.33\}, \{6, 8\}), (\{X_1, X_2, X_6, X_7\}, \{2.5, 2, 2.25, \\ &2.25\}, \{6, 12\}), (\{X_4, X_5, X_8\}, \{3.67, 1, 2.33, 1.33\}, \{7, 11\}), (\{x_1, x_2, x_4, x_5, x_8\}, \\ &\{3.25, 1.5, 1.75, 1.75\}, \{6, 12\}), (\{x_1, x_2, x_4, x_6, x_7, x_8\}, \{3, 1.83, 2.5, 1.83\}, \{6, \\ &12\}), (\{x_1, x_2, x_4, x_5, x_6, x_7\}, \{2.67, 1.67, 2.17, 1.83\}, \{6, 12\}), (\{x_1, x_2, x_3, x_4, x_5, x_6, \\ &x_7, x_8\}, \{4.89, 1.5, 2.13, 2\}, \{5, 12\})\}. \end{aligned}$$

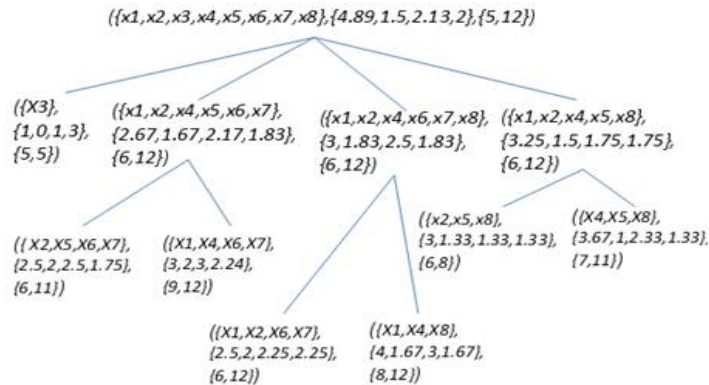


Fig.1. Ontology tree corresponding to ontology granular set

The ontology tree shown in Fig. 1 can be obtained through the application of Algorithm 3.

Based on the granular computing theory, this paper carries out concept classification by the compatibility relations. With the third parameter RG of the ontology granule, the paper implements a granule layer. With the second parameter IG the paper achieves arrangement of the ontology granule in the same layer from left to right, and at the end the paper successfully completes the learning process of the ontology concept hierarchy model. Through empirical study of the ontology on dyspnea of TCM, an ontology tree is exported using Protégé tools to obtain the concept fragment of ontology as shown in Fig. 2. The method in the paper is proved to be scientific and efficient and it enriches the auto-building theory of ontology.

5. Conclusion

Based on the granular computing theory, we have successfully completed the ontology concept hierarchy model of the learning process. The model has achieved concept classification by compatible relationship, and completed a grain of the layer with the help of the ontology grain of the third parameter of RG and the second parameter IG. In the study of the provincial fund project "Based on a set pair analysis of the case-based reasoning model and the application of ontology", we screened a base of more than 1500 records, and established a relatively perfect card for the basis of the traditional Chinese medicine database. On the basis of the application of the ontology construction model based on granular computing, we have successfully constructed out the ontology of traditional Chinese medicine.

The research work of the next step is to improve the ontology description based on granular computing, to define the granularity concepts of the ontology granule and the nature of the granule relationship, to define the ontology similarity function through transformation of the granule, and ultimately to apply ontology mapping and integration.

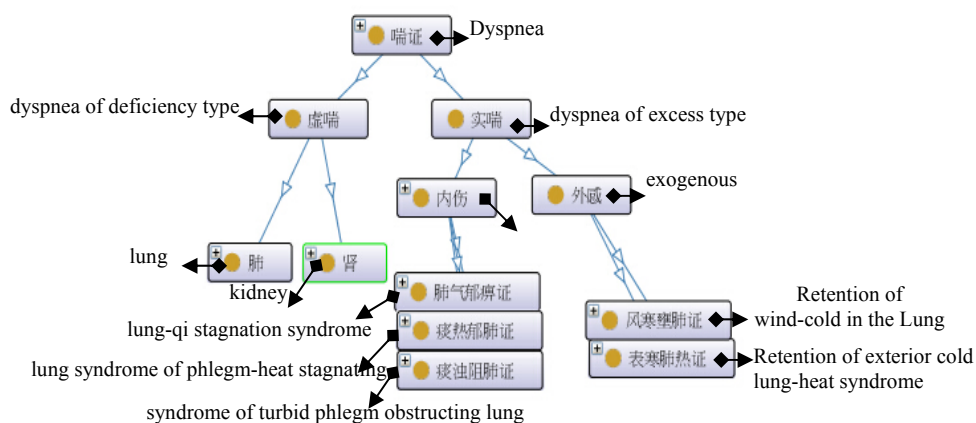


Fig. 2. Ontology fragment of TCM

Acknowledgements: This work was supported by the National Natural Science Foundation of China (61370168) and the National Natural Science Foundation of Hebei (F2014209238).

References

1. Wang, G., Q. Zhang, J. Hu. An Overview of Granular Computing. – J. CAAI Transactions on Intelligent Systems, Vol. 2, 2007, No 6, pp. 8-26.
2. Rose, D., H. Stefan. Comparison of “Personal Ontologies” Represented Through Conceptual. – In: Proc. of 13th European Conference on Artificial Intelligence (ECAI-98), Brighton, UK, 1998, pp. 341-345.
3. An-Hai, D., M. Jayant, D. Pedro, H. Alon. Ontology Matching: A Machine Learning Approach. – In: S. Steffen, S. Rudi, Eds. Handbook on Ontologies in Information Systems. Heideberg, DE, Springer-Verlag, 2003, pp. 396-416.
4. Yao, Y. Y. Granular Computing for Data Mining. – In: Proc. of SPIE Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, USA, 2006.
5. Yan, L., X. Zhang, J. He, F. Li. Uniform Description for Granules Appearing in Rough Set, Quotient Space and Concept Lattice. – Computer Engineering and Applications, Vol. 46, 2010, No 9, pp. 38-41.
6. Guo, W., X. Zhang. Description of Ontology Modules Based on Granularity. – New Technology of Library and Information Science, 2010, No 2, pp. 1-5.
7. Fan, P. Ontology and Ontology Grain Computing Research. East China, Jiaotong University, Master’s Thesis, 2008.
8. Li, J., L. Wei, Y. Huang. The Structure of Multi-Granulation Object Oriented Concept Lattice. – Journal of Northwest University (Natural Science Edition), Vol. 42, 2012, No 2, pp. 185-189.
9. Sheng, Y., Y. Li. A Method for Ontology Merging Based on the Concept Lattice. – Microelectronics and Computer, Vol. 25, 2008, No 9, pp. 34-36.
10. Ma, Y., G. Zhang, R. Zhang. On Granulation Based Concept Lattice of Formal Context. – Computer Applications and Software, Vol. 27, 2010, No 7, pp. 11-13.