International Journal of Grid and Distributed Computing
Vol. 9, No. 1 (2016), pp.79-88
http://dx.doi.org/10.14257/ijgdc.2016.9.1.08

Multi-Value Attribute Concept Lattice Reduction Based on Granular Computing

Hongcan Yan Feng Zhang, Baoxiang Liu

College of Science, North China University of Science and Technology, Tangshan 063000
yanhongcan@ncst.edu.cn

Abstract

Concept lattice essentially describes the relationship between objects and attributes. The reduction of multi-value attribute concept lattice is a hot topic in the fields of information retrieval, knowledge discovery and data mining etc., while the granular computing emphasizes observing and analyzing the same problem from different granular worlds. It makes the complex problems around us be mapped to an easy to handle and more simple theory of calculation. The paper gives definitions of the concept granule and the compatible concept granular set by application of information granular and the granular of layered theory, and provides an algorithm to compute concept granular set through calculation of the compatible relationship. The paper further constructs the concept granule lattice, and then deletes the attribute of smaller contribution to concept granule. Through the comparison of the concept granule lattices, the multi-value attribute reduction could be achieved and the core attribute set in the formal context could be obtained. Instances could demonstrate the high efficiency and accuracy of this algorithm that is easier to realize through programming. Through the resolution of attribute, the calculation complexity could be reduced and the efficiency of calculation could be improved.

Keywords*: Granular computing, Concept granule, Multi-value attribute reduction, Concept granule lattice, Concept granular resolution*

1. Introduction

In the current field of computational intelligence research, granular computing is a new method of simulating human thinking and solving the complex problem. It covers all the theories, methods and technology about granularity which is an effective tool to solve complex problem, enlarges data mining and processes fuzzy information [1]. Granular computing is a new term for the problem solving paradigm and may be viewed more on a philosophical rather than technical level [2]. Yao views granular computing as a triangle: structured thinking in the philosophical perspective, structured problem solving in the methodological perspective and structured information processing in the computational perspective [3]. Information granule [4-6] is an essential part of granular computing. It refers to a collection of elements. These elements have the indistinguishability, similarity, or function on cohesion *etc*. Information granulation aims to establish a user-oriented effective concept based on the outside world.

Support by The National Natural Science Foundation of China (61370168) and The National Natural Science Foundation of Hebei（F2014209238）

ISSN: 2005-4262 IJGDC

Concept lattice [7] is also named Galois lattice or formal concept analysis. It is firstly proposed by WILLER in 1982. The concept lattice can clearly show the hierarchical structure of the relationship between concepts. It can be visualized the knowledge structure of database. It is a method of data mining. Each node of concept lattice is a formal concept consisting of two parts: extension, it covered with the instances of concept, and intension, it is the common features of instances. Therefore, the concept lattice has been widely applied to an information retrieval, data mining, knowledge discovery and other fields. With the rapid development of Internet, data objects that the application processes in this field are more and more complex, so the numbers of concept lattice nodes grow, especially when the attribute of concept lattice is not binary attributes (0,1) but the multi-valued attributes. So the multi-valued attribute concept lattice reduction is the main point of the applying research.

So far, most of the multi-valued attribute reduction uses the multi-value attributes uniformizable method [8,9], which increases the complexity of data processing. Reference [10] gives the multi-value attributes related definitions and properties of the formal context, then in the formal context where the multi-value attribute is not converted to uniformizable value, basing on the Galois connection, Reference [10] put forward multi-valued attribute reduction method under the multi-value attribute formal context, and multi-value attributes are classified. Multi-value attribute formal context reduction method could be obtained on the basis of discernibility matrix in multi-value attribute set. According to this idea, this paper introduces granular computing theory and gives the concept granule, concept granular resolution and other core definition and property, and then this paper gets the concept granular set through the granulation calculated of compatibility relation and the compound algorithm of concept granule. In this way, the multi-value attributes reduction of concept lattice could be achieved. Instances could demonstrate that based on granular computing, multi-value attribute concept lattice reduction algorithm is not only efficient and easier to realize through programming, but also more direct to get the core attribute sets of formal context, and provides a new method and way for applications in the field of information retrieval and knowledge discovery.

2. Concept Granule and Granular Layer on the Domain

Definition1. Formal context: *(U, A, R)* is a formal context, which $U=\{u_1,u_2,\cdots,u_n\}$ is the set of objects, each $u_i\,(i\leq n)$ called an object; $A=\{a_1,a_2,\cdots,a_m\}$ is the set of attributes, each $a_j\,(j\leq m)$ called an attribute; *R* is binary.

Definition2. Core attribute: Formal context *(U,A,R)* all the reduction sets are $\{D_i \mid i\in t\}$, the set of attributes *A* can be divided into the following three categories:

(1)The set of core attributes: $C=\bigcap_{i\in t} D_i$

(2) The set of relative necessary attributes: $K=\bigcup_{i\in t} D_i-\bigcap_{i\in t} D_i$

(3) The set of unnecessary attributes: $I=A-\bigcup_{i\in t} D_i$

Among them, *t* is the set of indexes.

Theory1. For any formal context *(U, A, R)*, there must be reduction sets [11].

Proof: If for any $a\in A$, has $L(U,A-\{a\},R_{A-\{a\}})\neq L(U,A,R)$ then A is a reduction set. If there is $a\in A$, then $L(U,A-\{a\},R_{A-\{a\}})\cong L(U,A,R)$, study $B_1=A-\{a\}$. If for any $b_1\in B_1$, there is $L(U,A-\{a\},R_{A-\{a\}})$, then B_1 is a reduction set; otherwise study $B_2=B_1-\{b_1\}$, repeat the above process, due to A is finite set, we can always find at least one reduction set. So there must be reduction sets of formal context *(U, A, R)*

Definition3. Compatible function: For $\forall x,y\in U$, if mapping $\tau: x\times y\rightarrow[0,1]$ content the condition:

(1) $\tau(x,x)=1$ (Reflexivity)

(2) $\tau(x,y)=\tau(y,x)$ (Symmetry)

Then $\tau(x,y)$ is called the compatible degree on set *U* of *x* and *y*. τ is compatible function.

Definition4. Threshold function: Set $\lambda \in [0,1]$ is threshold value, denoted by: $\tau_\lambda = \{(x,y) | \tau(x,y) \geq \lambda\}$, said τ_λ is under the threshold of compatible relationship.

Definition5. Concept granule: Concept granule binary group *CG= (EG, IG)*, *EG* is the concept granular extension and *IG* is the concept granular intension.

Definition6. Rough-fine of concept granule: The extension of concept granule CG_1 and CG_2 is respectively eg_1 and eg_2, content: $eg_1 \subseteq eg_2$. That means concept granule CG_1 is finer than CG_2, and concept granule CG_2 is rougher than CG_1. Concept granule CG_1 divides finer in the set of objects *U*. denoted by: $CG_1 \leq CG_2$. "≤ " is partial order relation of concept granular set.

Definition7. Parent granule and child granule: If concept granule $CG_1 \leq CG_2$ and there do not have concept granule CG_3 between them content: $CG_1 \leq CG_3 \leq CG_2$, then called CG_1 is direct child granule of concept granule CG_2, also CG_2 is direct parent granule of concept granule CG_1.

Definition8. Concept granular layer: If the cardinality of concept granular extension is equal: $|EG_i| = |EG_k|$, then concept granule CG_i and CG_k are in the same layer of concept granule.

Definition9. Concept granular set: All the concept granules constitute concept granular set: $CGS = \{CG_1, CG_2, \cdots, CG_n\}$, where containing extension is domain of concept granule and extension is null set of concept granule.

Definition10. Concept granule lattice: According to the rough-fine relationship between concept granules, we can use the form of lattice to show concept granular set. The root node of the lattice is concept granule, whose extension is domain, and frontier node is concept granule whose extension is null set of concept granule. The lattice of parent and child nodes is content rough-fine relationship of concept granule.

3. The Generation Algorithm of Concept Granular Set Based on Granular Computing

Multi-value formal context *(U,A,R)* of Concept lattice is given, $U = \{u_1, u_2, \cdots, u_n\}$, $A = \{a_1, a_2, \cdots, a_m\}$,The object *i* have attributes are $[a_{i1}, a_{i2}, \cdots, a_{im}]$

Concept granule binary group *(EG,IG)*, the intension *IG* generated by the function *Function_IG(EG)*.

Function_IG(EG):

Input: Concept granular extension *EG*

Output: Concept granular intension *IG*

```
Function_IG(EG)   {
  IG[1,2,…,m]=[0,0,…0]; // Initialize concept granular extension
    for all u ∈EG   {
        IG [1,2,…,m]= IG [1,2,…,m]+ [au1,au2,…aun];
        IG [1,2,…,m]= IG [1,2,…,m]/|EG| ;
            // |EG| is the number of concept granule CG extension set
    }
    return IG ;
}
```

3.1. Initial Concept Granular Set SCG0

Algorithm1 *Function_SCG0*

Input: formal context (the number of objects is *n*; the number of attributes is *m*); the compatible function τ and threshold λ

Output: Initial concept granular set *SCG0*

```
Function_SCG0() {
  SCG0={};      // Initial concept granular set is null set
  For( u=u1;u<=un;u++) do{    // calculate compatible class of objects
     EG={u};
     for(j=1,j<=n;j++){
        if(j!=u && τ(u,j)>= λ)    // content threshold τλ
              EG= EG∪{j}
        }        // calculate extension
    IG=Function_IG(EG);    // Calculate intension
    CG=(EG,IG);               //    Constitute a concept granule
    SCG0=SCG0 ∪CG;
   }
    Return    SCG0;
   }
```

3.2. Extend Concept Granular Set

Definition11. Attribute domain *VD:* Set the extension of Concept granule *CG* is $EG=\{u_1,u_2,\dots, u_k\}$ if $minV=min_{1\le i\le k}(a_{i1}+a_{i2}+\cdots+a_{in})$, $maxV=max_{1\le i\le k}(a_{i1}+a_{i2}+\cdots+a_{in})$, a_{ij} express object *i* possesses attribute *j*, then set pair (*minV, maxV*) is the attributes domain *VD* of Concept granule *CG*.

Property1. Concept granular coupling: If the bigger interval value of concept granular attributes domain *VD*; the smaller coupling of concept granular extension *EG*, then the looser of concept granule *CG,* the structure of *CG* is instability and can be decomposed finer concept granule; whereas, if the bigger coupling of concept granular extension *EG*; the tighter of concept granule *CG*, the structure of *CG* is stability and can be compound rougher concept granule.

Property2. Attribute resolution *AV*: the ratio of each concept granular intension *IG* and its attribute domain *VD* is called attribute resolution *AV*, denoted by: *AV=IG/VD*. If the value of attribute resolution is larger, then the contribution of attribute to the concept granule is greater; whereas, if the smaller value of the attribute resolution, then the smaller contribution of attribute to the concept granule. Attribute resolution is distinguish parameters of concept granular layer. When attribute reduction, we priority to delete the small value of attribute resolution.

According to the property1, we can carry out on the concept granule decompose and merge operations.

Definition12. Decompose and merge on concept granule: Under the same granular layer. Set: $\forall CG_i, CG_j \in CGS\ (i \ne j)$

(1) $CG_{i\cap j} = CG_i \cap CG_j$, the extension of concept granule $CG_{i\cap j}$ is $EG_i \cap EG_j$, through function *Function_IG(EG)* obtain the intension of concept granule $CG_{i\cap j}$. Concept

granule $CG_{i\cap j}$ is finer than CG_i or CG_j. Concept granule $CG_{i\cap j}$ is the direct child granule of concept granule CG_i or CG_j.

(2) $CG_{i\cup j} = CG_i \cup CG_j$, the extension of concept granule $CG_{i\cup j}$ is $EG_i \cup EG_j$. Through function *Function_IG(EG),* the intension of concept granule $CG_{i\cup j}$ is obtained. Concept granule $CG_{i\cup j}$ is rougher than CG_i or CG_j. Concept granule $CG_{i\cap j}$ is the direct parent granule of concept granule CG_i or CG_j.

The Algorithm1 can calculate the initial concept granular set of the formal context. According to the definition12, we can extend other concept granules through "intersection" and "union" operation of concept granule.

Algorithm2 *Function_SCG*

Input: Initial concept granular set *SCG0*

Output: Concept granular set *SCG*

```
Function_SCG(SCG0)   {
  E is the extension of all concept granules in SCG0
  SCG={};
    Do {
      For all CGi ∈SCG0 do {
     NEW_EG= (EGi ∪EG) ||   (EGi∩EGk) ;
          // Decompose and merge operation on concept granule
            If(NEW_EG ∉E)
              {   New_IG=Function_IG(EG);
                  CG=( New_EG, New_IG)
                  SCG0=SCG0 ∪CG}
              }
       }
      Return SCG;
    }
      SCG=SCG0;
    Return SCG;
  }
```

When we obtain concept granular set SCG, and construct the lattice of concept granular set based on layered construct lattice algorithm of concept lattice, definition 8 and 10. For example, SCG={CG_1,CG_2, ⋯,CG_n}, if $CG_1 \leqslant CG_2$, then we call concept granule CG_1 is the child node of CG_2 ,and concept granule CG_2 is the parent node of CG_1, if $|CG_2| = |CG_3|$. Then concept granule CG_2 and CG_3 are in the same layer of concept granule lattice.

4. Multi-value Attribute Reduction of Concept Lattice

In order to fully illustrate the efficiency and accuracy of the algorithm, and compare with reduction attribute results of Reference [10], the paper adds four objects on the premise of Reference [10] given formal context and creates the formal context shown in table1

Table 1. Formal Context

	a1	a2	a3	a4
u1	2	3	1	1
u2	3	4	2	2

u3	1	1	4	2
u4	1	1	3	2
u5	3	4	1	2
u6	1	1	2	1
u7	2	2	4	1
u8	2	4	1	1

Based on characteristics of date, using compatible relationship:

$$\tau(u_j,u_k)=\frac{1}{1+\sqrt{\frac{\sum_{i=1}^{m}(x_i-y_i)^2}{n}}}$$

$u_j, u_k \in U$, x_i is *i-th* attribute value of object u_j, as well y_i is i-th attribute value of object u_k, this paper using compatible parameters $\lambda = 0.5$

We can obtain initial concept granular set composed of four concept granules through algorithm1.

SCG0=*{({u1,u2,u5,u8},{2.5,3.75,1.25,1.5}),({ u3,u4,u7},{1.3,1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,1.25,3.25,1.5},({ u6,u4 },{1,1,2.5,1})}*

We can obtain concept granular set through algorithm2 expend concept granules. as follows:

SCG=*{({u1,u2,u5,u8},{2.5,3.75,1.25,1.5}),({ u3,u4,u7},{1.3,1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,1.25,3.25,1.5},({ u6,u4 },{1,1,2.5,1}),({u1,u2,u3,u4,u5,u6,u7,u8},{1.88,2.5,2.3,1.5})}*

The concept granule lattice is shown in Figure 1.

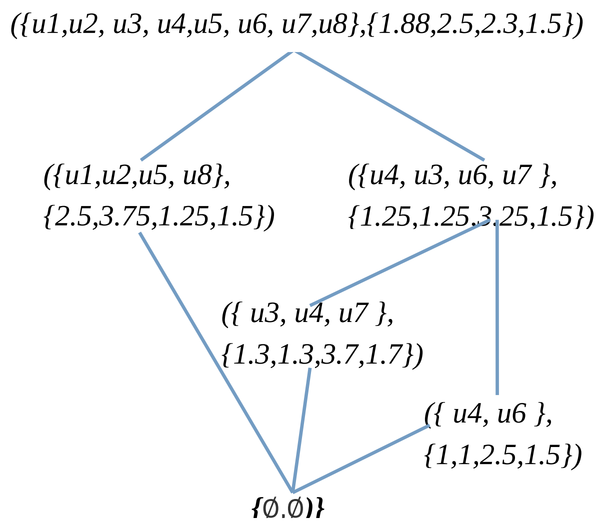

Figure 1. Concept Granule Lattice

4.1. Calculate Attributes Resolution

Based on Definition9, the attribute resolution of four concept granules are respectively *VD1=4,VD2=2,VD3=4,VD4=2.* According to Definition2, and by the second element *IG* of concept granule, attribute resolution can be calculated as follows:

$AV_{a1}=((2.5/4)+(1.3/2)+(1.25/4)+(1/2))/4=0.52$

$AV_{a2}=((3.75/4)+(1.3/2)+(1.25/4)+(1/2))/4=0.6$

$AV_{a3}=((1.25/4)+(3.7/2)+(3.25/4)+(2.5/2))/4=1.06$

$AV_{a4}=((1.5/4)+(1.7/2)+(1.5/4)+(1/2))/4=0.52$

Obviously :$AV_{a4} \leq AV_{a1} < AV_{a2} < AV_{a3}$

4.2. Attribute Reduction Based on Attribute Resolution Sequential:

1. Probe attribute a4

Based on Definition2, we priority to delete attribute *a4*, and then the new formal context could be obtained. Initial concept granular set could be obtained through algorithm1.

SCG0=*{({u1,u2,u5,u8},{2.5,3.75,1.25}),({u3,u4,u7},{1.3,1.3,3.7}),({u4,u3,u6,u7},{1.25 ,1.25,3.25},({ u6,u4 },{1,1,2.5})}*

We can obtain the concept granular set *SCG* through the expansion operation of the algorithm 2:

SCG=*{({u1,u2,u5,u8},{2.5,3.75,1.25}),({ u3,u4,u7},{1.3,1.3,3.7}),({u4,u3,u6,u7},{1.25,1.25,3.25},({ u6,u4 },{1,1,2.5}),({u1,u2,u3,u4,u5,u6,u7,u8},{1.88,2.5,2.3})}*

The concept granule lattice is shown in Figure.2.

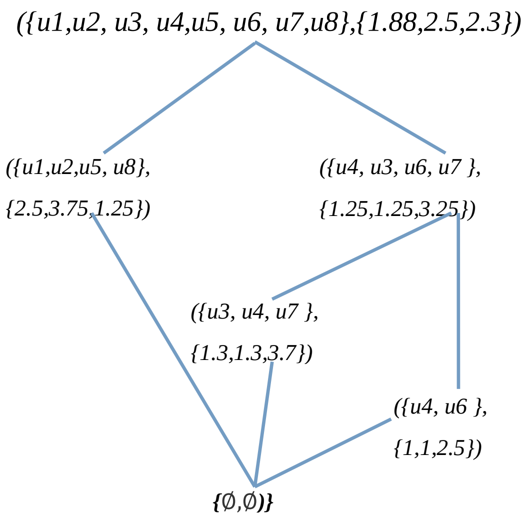

Figure 2. Concept Granule Lattice*{A-a4}*

Through comparing Figure 1 and 2: delete attribute *a4* of concept granule lattice and original concept granule lattice are isomorphism. So *a4* is redundancy attribute.

2. Probe attribute *a1*.

Next, probe the concept granules after deleting attribute *a1*. We can obtain the initial concept granular set through Algorithm1.

SCG0=*{({u1,u2,u5,u8},{3.75,1.25,1.5}),({ u3,u4,u7},{1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,3.25,1.5},({ u6,u4 },{1,2.5,1})}*

We can obtain the concept granular set SCG through Algorithm2. as follows:

SCG=*{({u1,u2,u5,u8},{3.75,1.25,1.5}),({ u3,u4,u7},{1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,3.25,1.5},({ u6,u4 },{1,2.5,1}),({u1,u2,u3,u4,u5,u6,u7,u8},{2.5,2.3,1.5})}*

The concept granule lattice is shown in Figure.3.

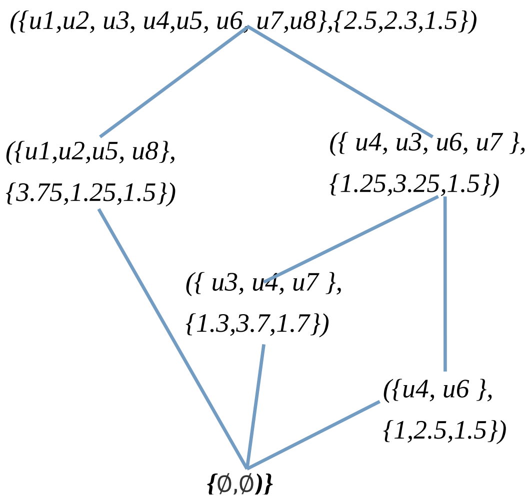

Figure 3 .Concept Granule Lattice{A-a1}

Similarly, *a1* is redundancy attribute.

3. Probe attribute *a2*

When delete attribute *a2*, then we can obtain the initial concept granular set through algorithm1.

SCG0=*{({u1,u2,u5,u6,u8},{2.2,1.4,1.4}),({u2,u1,u5,u8},{2.5,1.25,1.5}),({u3,u4,u7},(1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,3.25,1.5}),({ u6, u1,u4,u8 },{1.5,1.75,1.25})}*

We can obtain concept granular set SCG through Algorithm2.

SCG=*{({u1,u2,u5,u6,u8},{2.2,1.4,1.4}),({u2,u1,u5,u8},{2.5,1.25,1.5}),({u3,u4,u7},(1.3,3.7,1.7}),({u4,u3,u6,u7},{1.25,3.25,1.5}),({u6,u1,u4,u8},{1.5,1.75,1.25}),({u1,u2,u4,u5,u6,u8},{2,1.67,1.5}),({u1,u8},{2,1,1}),({u1,u3,u4,u6,u7,u8},{1.4,2.8,1.4}),({u4,u6},{1,2.5,1.5})}*

The concept granule lattice is shown in Figure.4.

Through concept granule lattice display we can find: deleting attribute *a2* of concept granule lattice and **o**riginal concept granule lattice is different; then *a2* is not redundant attribute. Because of $AV_{a2}<AV_{a3}$, probing attribute reduction should be ceased.

In conclusion: *a1* and *a4* are redundancy attributes, after deleting *a1* and *a4*, concept granule lattice will not change, namely a1 and *a4* are less contribution to concept granule lattice, and can be deleted. However, deleting *a2*, concept granule lattice changes a lot, and enlarges lots concept granules, so *a2* is great contribution to concept granule lattice. Because $AV_{a2}<AV_{a3}$, then *a3* is greater contribution to concept granule lattice. As with *a2*, *a3* is the core attribute. The conclusion is Consistent with reference [10]

In this case, the concept granule is generated by threshold $\lambda = 0.5$, if the threshold function is raised; then the compatibility on concept granule is batter, and its coupling increased. The structure of concept granule is more stable. Based on granular computing, multi-value attribute concept lattice reduction avoids complex pattern matching and traverses the database operation for many times, so it greatly improves the compute efficiency, and obtains the rank of important attributes precisely. Next, the attribute the relationship between classification indexes set t and the attribute resolutions *AV* should be analyzed, and the method in the application of interval concept lattice attribute reduction should be further studied.

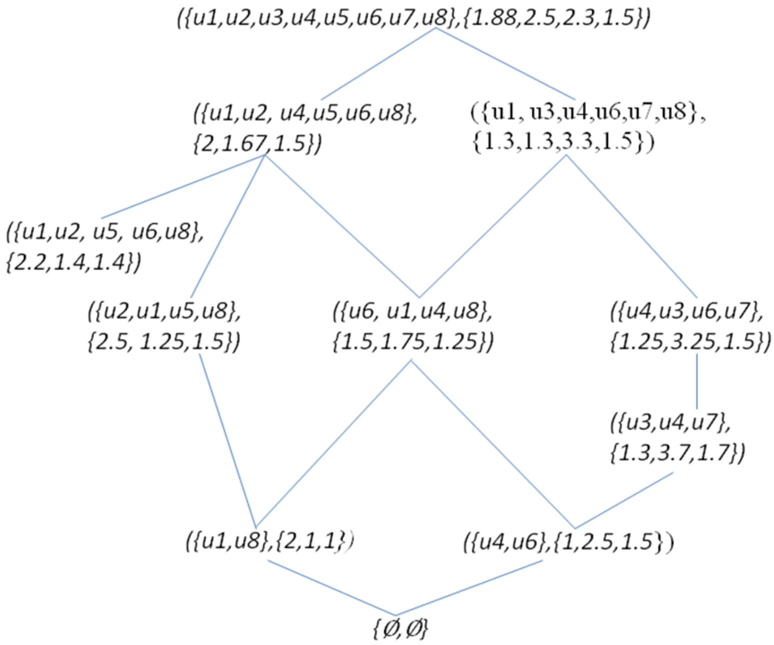

Figure 4. Concept Granule Lattice*{A-a2}*

References

[1] W. Guoyin, Z. Qinghua , H. Jun, "An overview of granular computing. Transaction on intelligent systems", vol. 2, no. 6, **(2007)**.

[2] J.T. Yao, "Information granulation and granular relationships", Proceedings of 2005 IEEE Conference on Granular Computing, **(2005)**; Bei-jing, China.

[3] Y.Y Yao, "The art of granular computing. Proceedings of International Conference on Rough Sets and Emerging Intelligent System Paradigms**", (2007)**.

[4] M. Duoqian, W. Guoyin, L. Qing, L. Zaoyang Y. Yiyu, "Granular computing in the past, present and future", Science press ,BEI Jing,**(2007)**.

[5] L. Qing, "Rough set and Rough reasoning" , Science press ,BEI Jing, **(2005)**.

A. Bargiela and W. Pedrycz, "Granular Computing：an Introduction" Kluwer Academic Publishers, **(2002)**.

[6] R. Wille, "Restructuring lattice theory: An approach based on hierarchies of concept", Rival I,ed. Ordered Sets, **(1982)**.

[7] L. Lan, "Attribute reduction based on information entropy and its application", Liao Ning: Dalian Maritime University, **(2008)**.

[8] L. Shuzhi, X. Guanghua, H. Cheng, "Decision table simplification algorithm based on attribute single–valued representation and attribute core calculation", Journal of Xi'an jiaotong university, vol. 44, no. 1, **(2010)**.

[9] Y. Jianfeng, "Multi-value concept lattice and Interval value concept lattice structure and attribute reduction", Northwestern university, **(2012)**.

Authors

Hongcan YAN, She received her M.C.A.in Computer Application(2001) and PhD in Information Sciences (2009) from Tianjin University. Now she is full professor of informatics at College of Science, North China University of Science and Technology. Since 2010 she has become a member of IEEE-CS. Her current research interests include different aspects of Web Semantic Retrieval and Knowledge discovery..

Feng ZHANG, He received B.Sc. in Mathematics (2012) from Tangshan Normal University. Now he is a postgraduate student at North China University of Science and Technology. His current research interests include different aspects of Mathematical model and its application of information processing

Baoxiang LIU, He received his B.Sc. in Mathematics (1982) from Hebei University. Now he is full professor of mathematics at College of Science, North China University of Science and Technology. Since 2004 he has become the dean of College of Science. Since 2010 he has become Senior Member of China field statistical society. His current research interests include different aspects of Rough sets theory and application.