The Reduction Algorithm of Concept Lattice Based on the Improved Discernibility Matrix

Qiu-na Zhang[1], Li Li[1], Dong-mei Li[1], Jing Zhang[2]

[1] Qinggong College, North China University of Science and Technology, Tangshan, China
[2] Hengshui engineering school of science and technology, Hengshui , China
zhangqiuna1980@163.com
563611609@qq.com
472479909@qq.com

***Abstract* — This article proposes the concept lattice reduction based on the improved discernibility matrix using discernibility matrix of rough set, and with examples, proves that the improved reduction algorithm costs shorter time, makes reduction speed faster, and also saves storage space in comparison with the original discernibility matrix reduction algorithm.**

Keywords- formal context, concept lattice, discernibility matrix, reduct, core

I. INTRODUCTION

Formal concept analysis (FCA) was proposed by Wille R. in 1982, which is a data analysis tool used in concept discovery, sorting and display. As a kind of efficient and potential knowledge discovery tool, FCA receives much attention from artificial intelligence field. FCA is widely used in many engineering fields, such as machine learning, pattern recognition, expert system, computer network, decision analysis and data mining [1−4]. One important task of knowledge discovery and data analysis is knowledge reduction. Concept lattice reduction provides a powerful knowledge and data reduction tool for artificial intelligence, knowledge discovery and data mining based on FCA.

We want to gain more concise knowledges from Concept lattice, the core data structure in formal concept analysis. Therefore, the attribute reduction in concept lattice is developed. The attribute reduction in concept lattice is an important topic of knowledge discovery. Attribute reduction is to cause the knowledge to be expressed simply by deleting the non-correlated or unimportant attributes. At the same time the basic information cannot be loused in view of different goals. Meanwhile, the attribute reduction also becomes the important job in extracting rules based on concept lattice. So the study of the attribute reduction in concept lattice has been a hot problem in the field of concept lattice.

Z.Pawlak, a Poland mathematician, put forward Rough Set Theory, a kind of new mathematics data analysis theory, in 1982. It deals with the un-accuracy and uncertain problem by the strict mathematics formula, and has three kinds of ability of deducing, summing up and general knowledge reasoning. So, Rough Set is widely used in machine study, knowledge finding, knowledge acquisition, decision analysis, data mining, pattern-recognition, and so on, So far, Rough Set Theory has already become the focus-studied in international artificial intelligence.

Rough Set Theory is a kind of tool that treats un-accurate and fuzzy question. Its main thought is that on the premise of keeping categorized ability un-changed, it leads the decision or categorized rule through reduction of knowledge. Simplification of decision tables in Rough Set is to lead decision rule or categorized rule at least on the basis of the least condition attribute and minimum redundant attribute value.

Concept lattice describes concept’s relationship between the connotation and extension, which is a kind of perfect knowledge expression form. Concept lattice and rough set are all based on equivalence class as the basic building blocks. Thus there exists close inner link. In this paper, on the basis of the relationship between concept lattice and rough set, the reduction algorithm of concept lattice based on the improved discernibility matrix in rough set is proposed.

II. THE DEFINITION AND PROPOSITION OF CONCEPT LATTICE

Concept lattice analyzes the data by using formal context which is defined as follows:

Difinition1 [5] We call (U, A, I) a formal context, $U = \{x_1, x_2, \cdots, x_n\}$ is a object set, each $x_i \,(i \le n)$ is a object. $A = \{a_1, a_2, \cdots, a_m\}$ is an attribute set , each $a_j \,(j \le m)$ is an attribute.
I is a binary relation between U and A , $I \subseteq U \times A$ if $(x, a) \in I$, then we say x holds attribute a, which is denoted as xIa .

In this article we use 1 expresses $(x, a) \in I$, use 0 expresses $(x, a) \notin I$, then the formal context can be expressed as a form which only contains 0 and 1.

For a formal context (U, A, I), we give the definitions of object set $X \subseteq U$ and attribute set $B \subseteq A$:

$$X^* = \{a | a \in A, \forall x \in X, xIa\}$$
$$B^* = \{x | x \in U, \forall a \in B, xIa\}$$

Definition 2[5] Let (U, A, I) be a formal context, if binary array (X, B) satisfies $X^* = B$ and $X = B^*$, then we call (X, B) a formal concept, *X* is called the conceptive extension, *B* is called conceptive connotation.

As to the concept of formal context (U, A, I), we can use the relationship between super concept and sub-concept to define their order relation:

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

All concepts partially ordered set of formal context (U, A, I) is denoted by $L(U, A, I)$, we call it concept lattice.

Definition 3[5] Let $L(U, A_1, I_1)$ and $L(U, A_2, I_2)$ be two concept lattice. If any $(X, B) \in L(U, A_2, I_2)$, exists $(X', B') \in L(U, A_1, I_1)$, then $X' = X$, call $L(U, A_1, I_1)$ finer than $L(U, A_2, I_2)$, which is denoted as

$$L(U, A_1, I_1) \le L(U, A_2, I_2)$$

If $L(U, A_1, I_1) \le L(U, A_2, I_2)$ and $L(U, A_2, I_2) \le L(U, A_1, I_1)$, we say the two concept lattice isomorphism, which is denoted as

$$L(U, A_1, I_1) \cong L(U, A_2, I_2)$$

Under the formal context (U, A, I), $\forall D \subseteq A$, denoted as $I_D = I \cap (U \times D)$, then (U, D, I_D) is also a formal context.

Definition 4[5] For a formal context (U, A, I), if there exists a attribute set $D \subseteq A$, satisfies $L(U, D, I_D) \cong L(U, A, I)$, then we call *D a* consistent set of (U, A, I), if $\forall d \in D$, $L(U, D-\{d\}, I_{D-\{d\}})$ is not isomorphism to $L(U, A, I)$, then *D* is a reduction of (U, A, I). The intersections of all reductions are called core.

Lemma 1[5] For any formal context (U, A, I), reduction must exist.

Example 1 For a formal context (U, A, I) of table I, $U = \{1,2,3,4\}$, $A = \{a, b, c, d, e\}$, write the concept lattice and reduction.

TABLE I. A FORMAL CONTEXT

U	*a*	*b*	*c*	*d*	*e*
1	1	0	1	1	1
2	1	1	1	0	0
3	0	0	0	0	1
4	1	1	1	0	0

concept lattice is given by figure 1, reductions $D_1 = \{a, b, e\}$, $D_1 = \{b, c, e\}$

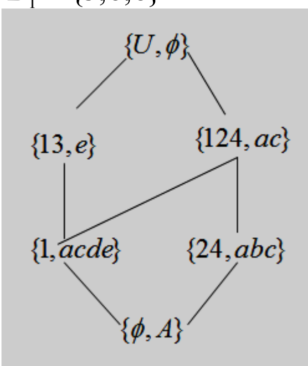

Fig.1example 1concept lattice

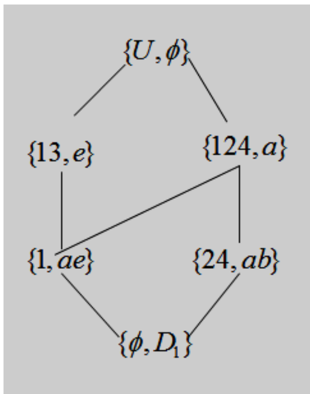

Fig.2 example 1 reduction concept lattice

III. REDUCTION ALGORITHM OF ROUGH SET BY IMPROVED DISCERNIBILITY MATRIX

Proposition 1[6] For a decision table $S=(U,A,C,D)$, if for any $y \neq x$, $d_x|C=d_y|C \rightarrow d_x|D=d_y|D$ 。 Then we call the regulation d_x uniform，otherwise we call it not uniform.

Lemma 2[6] According to the equivalence class decision, for the same decision rules of equivalence class, discernibility of the condition attribute is not needed. What shoud be done is to compare the individuals of different equivalence class to estimate whether the condition attribute is discernible, and whether the decision rules is uniform.

Proposition 2[6] Obtained attribute values from information system are distinct attributes. After reduction, if there is a single attribute as a conjunction, then the set of these attributes is a core set, which does not contain more than one attribute of nuclear conjunction items, each of which may be reducible.

Proposition 3[6] Obtained attribute values from information system is distinct attributes, and after reduction, we get discernible conjunctive normal form, and they can be equivalent to transformation as the minimum discernibility disjunctive normal form, and each item disjunction is corresponding to a reduction of the information system.

IV. THE REDUCTION OF CONCEPT LATTICE

For a formal context $L(U,A,I)$, $U=\{x_1,x_2,\cdots,x_n\}$, we can hold discernibility matrix by the reduction of rough set, it is a $n \times n$ symmetric matrix.

$c_{ij}=\{a \in A | a(x_i) \neq a(x_j) \quad i,j=1,2,\cdots,n\}$, gets discernibility function

$f(a_1,a_2,\cdots,a_k)=\wedge\{\vee_{c_{ij}} | 1 \le j \le i \le n, c_{ij} \neq \phi$, By using the Boolean operation, we can get all attribute reductions $D_1,D_2,\cdots,D_n$ of concept lattice.

Example 2 Let $S=(U,A,C,D)$ be a decision table, and $C=\{a,b,c,d\}$ is a condition attribute set, $D=\{e\}$ is a decision attribute set. The decision table is followed as table II..

TABLE II. DECISION TABLE

U	*a*	*b*	*c*	*d*	*e*
1	1	0	0	1	1
2	0	0	0	0	0
3	1	1	0	2	2
4	1	0	0	0	1
5	2	1	0	2	2
6	1	1	0	1	0
7	2	2	2	2	2

The concept lattice of this decision table is followed as figure 3（All the concepts of extension are omitted）

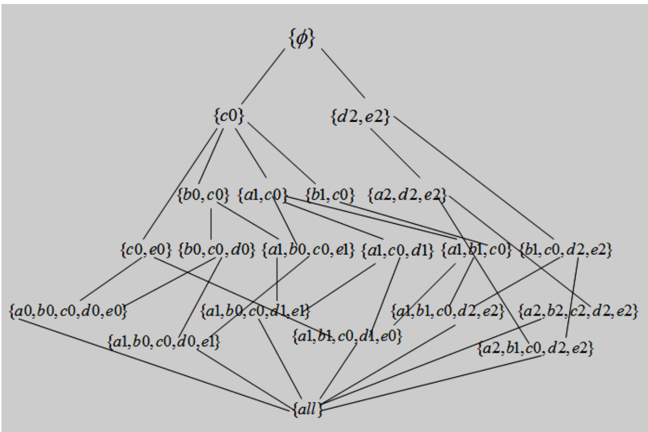

Fig.3 example 2 concept lattice

The discernibility matrix of this formal context is

	1	2	3	4	5	6	7
1							
2	*acd*						
3	*bd*	*abd*					
4	*d*	*a*	*bd*				
5	*abd*	*abd*	*a*	*abd*			
6	*b*	*abd*	*d*	*bd*	*ad*		
7	*abcd*	*abcd*	*abc*	*abcd*	*bc*	*abcd*	

discernibility function

$$f = (a \vee c \vee d) \wedge (b \vee d) \wedge (a \vee b \vee d) \wedge d \wedge a \wedge b \wedge (a \vee d) \wedge (b \vee c) \wedge (a \vee b \vee c \vee d) \wedge (a \vee b \vee c)$$
$$= abd$$

Therefore, the reduction of this concept lattice is $\{a,b,d\}$.

The improved discernibility matrix of this formal context is

	1	4	2	6	3	5	7
1							
4							
2	*ad*	*a*					
6	*b*	*bd*					
3	*bd*	*bd*	*abd*	*d*			
5	*abd*	*abd*	*abd*	*ad*			
7	*abcd*	*abcd*	*abcd*	*abcd*			

The discernibility function is

$$f^* = (a \vee d) \wedge a \wedge b \wedge (b \vee d) \wedge d \wedge (a \vee b \vee d) \wedge (a \vee b \vee c \vee d)$$
$$= abd$$

The reduction of this concept lattice is also $\{a,b,d\}$.

V. CONCLUSION

On the basis of the relationship between concept lattice and rough set, the reduction algorithm of concept lattice based on the improved discernibility matrix in rough set is proposed. The improving discernibility matrix is compared with the former discernibility matrix, the advatages are as follows: using shorter time, making reduction speed faster, and also saving the storage space.

ACKNOWLEDGEMENTS

All authors would like to thank Hebei Province Tangshan Science Project funding (14130249B).Hebei Province Youth Science Funding: (A2014209083).

REFERENCE

[1] WilleR.Restrueturinglattieehteory:AnPaProaehbasedonhierarehiesof eoneePts.In:Rivall,ed.Ordered Sets. Dordercht-Boston:Reidel,445~470,1982.

[2] Oosthuizen G D . The Application of Concept Lattice Machine Learning. Technical Report, University of Pretoria,South Afriea,1996

[3] HoTB. Incremental conceptual clustering in the framework of Galois lattice. In: Lu H, Motoda H, Liu H, eds. KDD: Techniques and Applications. Singapore: World Scientific,49~64,1997.

[4] Kent R E. Bowman C M. Digital Libraries, Conceptual Knowledge Systems and the Nebula Interface. Technical Report, University of Arknasas,1995

[5] Wenxiu ZHANG, Yiyu YAO, Yi LIANG. The rough set and concept lattice [M] xi'an: xi'an jiaotong university press ,2006

[6] Qing LIU. Rough set and Rough inference [M]. Beijing: Science press, 2001.8

[7] Ganter B, Wllle R,Fomral Concept Analysis, Mathematical Foundations. Berlin:Springer,1999

[8] Wenxiu ZHANG, Yi LIANG, Weizhi WU. Information systems and knowledge discovery. Beijing: Science press, 2003

[9] Shaobo DENG. Research simplification of decision tables of rough set based on binary scale [D].Nanchang University, 2005

[10] Zhongzhi SHI. Advanced artificial intelligence [M]. Beijing: Science press, 2006.9.

[11] Yao Y Y Concept lattices in rough set theory. In: Dic kS, Kurgan L, Pedrycz W, eds. Proceedings of 2004 Annual Meeting of the North American Fuzzy Information processing Society(NAFIpS2004), IEEE Catalog Number: 04TH8736, June 27(30):796~801,2004.

[12] Ooshtuizen G D. Rough sets and concept lattices. In: Ziarko W P, ed. Rough Sets, and Fuzzy Sets and Knowledge Discovery(RSKD,93). London: Springer-Verlag,24~31,1994.

[13] [13] Deogun J S, Saquer J. Concept approximations for formal concept analysis. In: Stumme G. ed. Working with Conceptual

Structures. Contributions to ICCS 2000.Verlag Shaker Aachen:73~83

[14] [14] Tonella P. Using a concept lattice of decomposition slices for Program understanding and impact analysis·IEEE Transactions on Software Engineering,29(6):495~509,2003.

[15] [15] Grigoriev P A, Yevtushenko S A. Elements of an agile discovery environment. In: Grieser G, Tanaka Y, Yamamoto A, eds. Proe6th International Conference on Discovery Seience (DS2003). Lecture Notes in Artifical Intelligence,2843:309~316,2003.