Information Sciences 222 (2013) 611–625

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Rough set model based on formal concept analysis

Xiangping Kang, Deyu Li *, Suge Wang, Kaishe Qu

School of Computer and Information Technology, Shanxi University, Shanxi, Taiyuan 030006, China
Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Taiyuan 030006, China

A R T I C L E I N F O

Article history:
Received 25 June 2010
Received in revised form 6 March 2012
Accepted 24 July 2012
Available online 8 August 2012

Keywords:
FCA
Rough set theory
Rough concept lattice
Rugh concept
Decision dependency

A B S T R A C T

This paper proposes a rough set model based on formal concept analysis. In this model, a solution to an algebraic structure problem is first provided in an information system: a lattice structure is inferred from the information system and corresponding nodes are called rough concepts. How to deal with common problems in rough set theory based on rough concepts is then explored, such as upper and lower approximation operators, reducts and cores. Decision dependency has become a common form of knowledge representation owing to its properties of expressiveness and ease of understanding, so it has been widely used in practice. Finally, application of rough concepts to the extraction of decision dependencies from a decision table is studied; a complete and non-redundant set of decision dependencies can be obtained from a decision table. Examples demonstrate that application of the method presented in this paper is valid and practicable. The results not only provide a better understanding of rough set theory from the perspective of formal concept analysis, but also demonstrate a new way of combining rough set theory and formal concept analysis.

1. Introduction

Rough set theory, proposed by Pawlak in 1982, is a theory used to study information systems characterized by inexact, uncertain or vague information [26]. One obvious advantage is that rough set theory does not need any preliminary or additional information about data. Because it is an effective tool with vast potential for knowledge acquisition, rough set theory has been widely investigated in the field of artificial intelligence [2,4,5,30,31,38,56].

On the basis of the philosophical understanding of a concept as a unit of thought constituted by its extent and intent, Wille proposed formal concept analysis (FCA) in 1982 [44]. The concept lattice with a complete structure and solid theory is an effective tool in FCA and is very suitable for mining potential concepts from data. FCA has been widely studied and applied to machine learning, software engineering and information retrieval [8,13,20,23,28].

Both FCA and rough set theory are complementary tools for data modeling and data analysis [1,22,26,39,44,50,52,57], and relations between them have attracted much research attention. Some achievements have been made in combining and comparing the two theories to improve our understanding of their similarities and differences. Existing studies are summarized below [54].

By investigating similarities and differences between two theories, comparative studies can provide a more general data analysis framework. Recently, more emphasis has been placed on integration of the two theories into a unified form. Kent argued that the two theories have much in common in terms of both goals and methodologies, and a new theory of rough concept analysis was introduced that can be viewed as a synthesis of rough set theory and FCA [14]. Wu et al. proposed an

* Corresponding author at: School of Computer and Information Technology, Shanxi University, Shanxi, Taiyuan 030006, China.
E-mail addresses: kangxiangping@yahoo.cn (X. Kang), lidysxu@yahoo.cn (D. Li), wsg@sxu.edu.cn (S. Wang), quks@sxu.edu.cn (K. Qu).

0020-0255/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2012.07.052

accuracy computational approach to characterize the rough formal concept numerically and studied basic relationships between indiscernibility relations and accuracies of rough concepts [48]. Concept lattices and approximation spaces were combined using a Heyting algebra structure [24]. Wolski investigated Galois connections and their relations to rough set theory [45]. Wasilewski demonstrated formal contexts and information systems, and described general approximation spaces [42]. Ho developed a method for acquiring concepts with lower and upper approximations in the framework of rough concept analysis [9]. Qi et al. discussed basic connections between FCA and rough set theory, and analyzed relationships between a concept lattice and the power set of a partition [29]. Wei and Qi studied the reduct theory from the viewpoint of rough set theory and concept lattice theory, and discussed their relations [43]. Lai and Zhang argued that each complete fuzzy lattice can be represented as the concept lattice of a fuzzy context based on rough set theory if and only if the residual lattice satisfies the law of double negation; they also proved that the expressive power of concept lattices based on rough set theory is weaker than that of concept lattices based on FCA [15]. Wang and Zhang revealed some basic relationships between extensions of concepts and equivalence classes in rough set theory, and studied relations between the reduct of formal context in FCA and the attribute reduct in rough set theory [41]. Yao conducted a comparative study of rough set theory and FCA based on the notion of definability [50]. Wolski defined operators of FCA and rough set theory using the specialization order for elements of a topological space, and further proved that FCA and rough set theory together provide a semantics for tense logic s4.t [46].

In many studies, application of the results from one theory to the other has been proposed, leading to different ways of combining rough set theory and FCA. Some new concept lattices have been constructed using more modal-style operators, and the properties of these lattices have been discussed extensively [7,45,50,51]. Hu et al. used the extended concept lattice obtained by introducing an equivalence class into a Galois concept lattice to describe the implementation of rough set theory [11]. Liu et al. applied the multi-step attribute reduct method for concept lattices based on rough set theory to the reduct of the redundant premises of the multiple rules used to solve JSSP [18]. Shao et al. investigated rough set approximations within FCA in a fuzzy environment, and two new pairs of rough fuzzy set approximations within fuzzy formal contexts were defined based on both lattice-theoretic and fuzzy set-theoretic operators [36]. Wang and Liu proposed an axiomatic fuzzy set formal concept that could be applied to represent the logic operations of queries in information retrieval [40]. On the basis of grey-rough set theory, Wu and Liu proposed an extension of the notion of Galois connection in a real binary relation, as well as notions of a formal concept and a Galois lattice [47]. Inspired by the reduct method in rough set theory, a Boolean approach proposed by Mi et al. to calculate all reducts of a context was formulated via the discernibility function [21]. Yang et al. constructed discernibility matrices and functions to compute all attribute reducts of real decision formal contexts that did not affect the results of the s rules or l rules acquired [49]. Li et al. described an associated reduct method in which the discernibility matrix and Boolean functions were used to compute all the reducts of a decision formal context [16]. Other proposals [3,10,17,19,25,27,33–35,37,55] have been described briefly by Yao and Chen [54].

From the studies described above, it is clear that three research directions exist for the integration of FCA and rough set theory [54]: integration of rough set theory into FCA, integration of FCA into rough set theory, integration of both into a unified framework. Although some achievements have been made, more detailed studies are required to obtain a more general data analysis framework. The present study introduces FCA into rough set theory and proposes a rough set model based on FCA. The model provides an interesting formulation of rough sets. In particular, it expresses the indiscernibility matrix as a formal context. Thus, it combines subsets of attributes and indiscernibility relations defined by subsets of attributes in terms of formal concept operators. Thus, the model provides a re-interpretation of many results of rough set theory by FCA, which can be viewed as a new attempt to combine FCA and rough set theory.

This paper is organized as follows: Section 2 briefly recalls some basic notions of rough set theory and FCA; Section 3 builds a rough concept lattice in the information system based on FCA; Section 4 presents some applications of rough concepts in an information system. That is, on the basis of rough concepts some common problems can be solved in rough set theory, such as attribute reducts, cores; Section 5 researches on decision dependencies in a decision table and finally gets a $\alpha\beta$-complete and $\alpha\beta$-non-redundant set Σ of decision dependencies based on rough concepts; Section 6 discusses perspectives for further works.

2. Basic notions of rough set theory and FCA

This section provides the most basic notions and facts of FCA and rough set theory. For more extensive presentations, see books of [6,26].

Suppose $S = (U, AT, V, f)$ (It is also denoted as (U, AT, V, I) in the following section.) is an information system, each subset $B \subseteq AT$ can determine a binary indiscernibility relation

$$Ind(B) = \{(x, y) \in U \times U | \forall m \in B, f(x, m) = f(y, m)\}$$

Let $B, C \subseteq AT$, if $m \in B$ and $Ind(B) \neq Ind(B - \{m\})$, we say m is indispensable; Further if every $m \in B$ is indispensable, we say B is independent. The set of all independent sets of attributes is denoted by IND_S. If $C \subseteq B$ and C is independent and $Ind(B) = Ind(C)$, then C is called a reduct of B. The set of all reducts of B is denoted as $Red_S(B)$. The set of all indispensable attributes in B is called the core of B denoted as $Core_S(B)$. If $Ind(B) \subseteq Ind(C)$, we say $B \to C$ is a function dependency of S. If R is a binary indiscernibility relation on U and U/R is the partition induced by R, then the lower approximation of $X \subseteq U$ relative to R can be defined as

$$\underline{R}(X) = \bigcup_{P \subseteq X \text{ and } P \in U/R} P$$

Correspondingly, the upper approximation can be defined as

$$\overline{R}(X) = \bigcup_{P \in U/R \text{ and } P \cap X \neq \emptyset} P$$

A formal context is a triple $K = (G,M,I)$, where G and M are sets, and $I \subseteq G \times M$ is a binary relation. In the case, members of G are called objects and members of M are called attributes, and I is viewed as an incidence relation between objects and attributes. Accordingly, we write gIm or $(g,m) \in I$ expressing "the object g has the attribute m".

For a set $A \subseteq G$ of objects we define

$$A^* = \{m \in M | gIm \text{ for all } g \in A\}$$

Correspondingly, for a set $B \subseteq M$ of attributes we define

$$B^* = \{g \in G | gIm, \quad \text{for all } m \in B\}$$

If $A^* = B$ and $B^* = A$, then (A,B) is called a formal concept of the context. $\mathcal{B}(K)$ denotes the set of all concepts of K. Then we have following simple facts [6].

Proposition 1. *If $K = (G,M,I)$ is a formal context, A, A_1, $A_2 \subseteq G$ are sets of objects and B, B_1, $B_2 \subseteq M$ are sets of attributes, then*

$$(1)\ A_1 \subseteq A_2 \Rightarrow A_2^* \subseteq A_1^* \quad (2)\ B_1 \subseteq B_2 \Rightarrow B_2^* \subseteq B_1^*$$
$$(3)\ A \subseteq A^{**}; B \subseteq B^{**} \quad (4)\ A^* = A^{***}; B^* = B^{***}$$

3. The rough concept lattice of an information system

Ganter and Wille essentially viewed an information system as a many-valued context and they provided a detailed description of how to assign concepts to the information system based on notions of scales and the technique of scaling [6]. In other words, the concept system of a many-valued context depends on scales and scaling. The corresponding process for one-valued contexts derived from an information system is shown in Fig. 1.

In Fig. 1, notions of scales and the technique of scaling are interpreted as follows. Each attribute of an information system can be interpreted by means of a one-valued context, and this context is the so-called scale. Choice of the scale for an attribute is essentially a matter of interpretation and is not mathematically compelling. Scaling can be viewed as a process of joining together of scales to make a one-valued context. The simplest scaling can be achieved by putting together individual scales without connecting them.

In Fig. 1, even though the first three steps are the same, derived contexts may still be different. Although not only do steps 2 and 3 determine one derived context, derived contexts based on these same key steps are also closely connected to each other. For example, in Fig. 2, in accordance with scaling (putting together individual scales without connecting them), two derived one-valued contexts are obtained from the information system in Table 1 based on the scales in Fig. 3 (scales for attributes a, b, c, d are denoted as S_a, S_b, S_c and S_d, respectively). The corresponding contexts are shown in Tables 2 and 3 (for all u_i, $u_j \in U$ (u_i,u_j) is simplified as u_{ij}. To save space, row heading for attributes are AT and column heading for elements are given as $U \times U$). Particular descriptions are given below.

Let (U,AT,V,I) be an information system, $\dot{AT}_m := \{m\} \times AT_m$ and $S_m = (U_m,AT_m,I_m)$ with $m \in AT$ are scale contexts, then Table 2 is the context (U,N,J_1) with

$$N = \bigcup_{m \in AT} \dot{AT}_m,$$

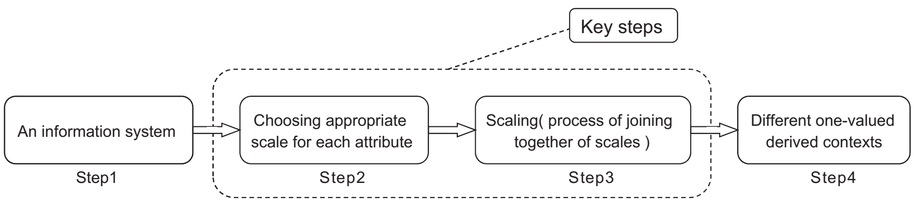

Fig. 1. The process of one-valued contexts derived from an information system.

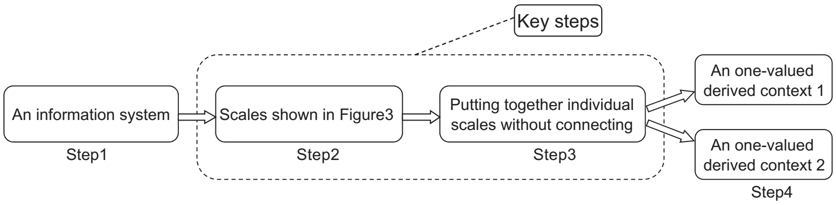

Fig. 2. The process of one-valued contexts derived from Table 1.

Table 1
An information system.

	a	b	c	d
u_1	Excellent	Very low	Excellent	0.5
u_2	Very poor	Low	Poor	0.7
u_3	Good	Medium	Very poor	0.3
u_4	Good	High	Good	0.6

S_a =

	++	+	~~
++	×	×	
+		×	
~~			×

S_b =

	&&	&	#	$
&&	×	×		
&		×		
#			×	
$				×

S_c =

	++	+	~	~~
++	×	×		
+		×		
~			×	
~~			×	×

S_d =

	>= 0.5	>= 0.7	>= 0.3	>= 0.6
0.5	×		×	
0.7	×	×	×	×
0.3			×	
0.6	×		×	×

++:=excellent, +:=good, ~~:=very poor, ~:=poor, &&:=very low, &:=low, #:=medium, $:=high

Fig. 3. Assigning scales S_a, S_b, S_c, S_d to attributes a, b, c, d in Table 1 respectively.

Table 2
The one-valued context 1 derived from Table 1.

	a			b				c				d			
	++	+	– –	& &	&	#	$	++	+	–	– –	0.5	0.7	0.3	0.6
u_1	×	×		×	×			×	×			×		×	
u_2			×		×					×		×	×	×	×
u_3		×				×				×	×			×	
u_4		×					×		×			×		×	×

Table 3
The one-valued context 2 derived from Table 1.

	u_{11}	u_{12}	u_{13}	u_{14}	u_{21}	u_{22}	u_{23}	u_{24}	u_{31}	u_{32}	u_{33}	u_{34}	u_{41}	u_{42}	u_{43}	u_{44}
a	×		×	×		×					×	×			×	×
b	×	×				×					×					×
c	×			×		×				×	×					×
d	×		×		×	×	×	×			×		×		×	×

and

$$uJ_1(m,n) \Longleftrightarrow m(u) = v \text{ and } vI_m n.$$

And Table 3 is the context $(U \times U, AT, J_2)$ with

$$(u_1, u_2)J_2 m \Longleftrightarrow m(u_1) = v_1, m(u_2) = v_2 \text{ and } v_1 I_m v_2.$$

The model proposed in this paper is inspired by the above discussions. In other words, in accordance with the simplest scaling, a derived one-valued context can be obtained from an information system based on the simplest scales in Table 4 (the scale for attribute m_i is denoted as S_{m_i}, where $m_i(U) = \{v_{i_1}, v_{i_2}, \ldots, v_{i_n}\}$). The model is described in detail below.

Suppose $S = (U, AT, V, f)$ is an information system, then for $\forall m \in AT$ and $\forall x, y \in U$, by the following rule

$$(x,y)I_S m \Longleftrightarrow f(x,m) = f(y,m)$$

S can be transformed to a one-valued context

$$K_S = (G, AT, I_S)$$

where $G = \{(x,y) | x, y \in U\}$, we say K_S is deduced from S.

As an example, an information system about cars is given in Table 5, where $U = \{u_1, u_2, u_3, u_4, u_5\}$ is the set of objects and $AT = \{a_1, a_2, a_3, a_4, a_5\}$ is the set of attributes with a_1 = price, a_2 = size, a_3 = engine, a_4 = maximum speed, and a_5 = performance/price ratio. Table 6 shows the one-valued context deduced from Table 5.

Table 4
The scale S_{m_i} of attribute $m_i \in AT$.

	v_{i_1}	v_{i_2}	v_{i_3}	$\cdots$	v_{i_n}
v_{i_1}	×				
v_{i_2}		×			
v_{i_3}			×		
$\vdots$				$\ddots$	
v_{i_n}					×

Table 5
An information system about cars.

	a_1	a_2	a_3	a_4	a_5
u_1	Low	Full	Diesel	Low	Low
u_2	Low	Full	Gasoline	High	High
u_3	High	Full	Diesel	Medium	Low
u_4	Low	Compact	Diesel	Low	Low
u_5	Low	Full	Diesel	High	Low

Table 6
The one-valued context deduced from Table 5.

	a_1	a_2	a_3	a_4	a_5
(u_1, u_1)	×	×	×	×	×
(u_1, u_2)	×	×			
(u_1, u_3)		×	×		×
(u_1, u_4)	×		×	×	×
(u_1, u_5)	×	×	×		×
(u_2, u_2)	×	×	×	×	×
(u_2, u_3)		×			
(u_2, u_4)	×				
(u_2, u_5)	×	×		×	
(u_3, u_3)	×	×	×	×	×
(u_3, u_4)			×		×
(u_3, u_5)		×	×		×
(u_4, u_4)	×	×	×	×	×
(u_4, u_5)	×				
(u_5, u_5)	×	×	×	×	×
(u_2, u_1)	×	×			
(u_3, u_1)		×	×		×
(u_4, u_1)	×		×	×	×
(u_5, u_1)	×	×	×		×
(u_3, u_2)		×			
(u_4, u_2)	×				
(u_5, u_2)	×	×		×	
(u_4, u_3)			×		×
(u_5, u_3)		×	×		×
(u_5, u_4)	×				

Proposition 2. *In $K_S = (G, AT, I_S)$, let $B \subseteq AT$, then*

$$B^* = Ind(B)$$

From Proposition 2, we can see that B^* is an indiscernibility relation on U. K_S is deduced from S on the basis of whether two objects have the same value for one attribute or not. It is closely related to the notion of a discernibility matrix in rough set theory in a slightly different form. In other words, K_S is a different representation of a discernibility matrix. For example, suppose that $(c_{ij})_{n\times n}$ is the discernibility matrix of the information system S, with $c_{ij} = \{a \in AT | a(x_i) \neq a(x_j)\}$ and $B \subseteq AT$. Then, for all $1 \leqslant i, j \leqslant n$ $(x_i, x_j) \in B^*$ if and only if $c_{ij} \cap B = \emptyset$.

In $S = (U, AT, V, f)$, let $\widetilde{A}_1$ and $\widetilde{A}_2$ be partitions of U, if every block of $\widetilde{A}_1$ is contained in some block of $\widetilde{A}_2$, we say $\widetilde{A}_1$ is a refinement of $\widetilde{A}_2$ and $\widetilde{A}_2$ is a coarsening of $\widetilde{A}_1$, which is denoted by $\widetilde{A}_1 \tilde{\subseteq} \widetilde{A}_2$. Operations $\tilde{\cup}$ and $\tilde{\cap}$ of sets $\widetilde{A}_1$ and $\widetilde{A}_2$ are defined as follows [53].

$$\widetilde{A}_1 \tilde{\cup} \widetilde{A}_2 = U/(R_1 \cup R_2)^\star$$
$$\widetilde{A}_1 \tilde{\cap} \widetilde{A}_2 = \{P \cap Q | P \in \widetilde{A}_1, Q \in \widetilde{A}_2, P \cap Q \neq \emptyset\}$$

where $\widetilde{A}_1 = U/R_1, \widetilde{A}_2 = U/R_2$ and $(R_1 \cup R_2)^\star$ is the transitive closure of $R_1 \cup R_2$. $\widetilde{A}_1 \tilde{\cap} \widetilde{A}_2$ is the largest partition that is a refinement of both $\widetilde{A}_1$ and $\widetilde{A}_2$, $\widetilde{A}_1 \tilde{\cup} \widetilde{A}_2$ is the smallest partition that is a coarsening of both $\widetilde{A}_1$ and $\widetilde{A}_2$. In the same way we can define operations $\tilde{\cup}$ and $\tilde{\cap}$ of $\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_n$.

In $K_S = (G, AT, I_S)$, let $\widetilde{A}$ be a partition of U and $B \subseteq AT$. For $\widetilde{A}$ we define

$$\widetilde{A}' = \left(\bigcup_{P \in \widetilde{A}} P \times P \right)^*$$

Correspondingly, for $B \subseteq AT$ we define

$$B' = U/B^*$$

From the above discussions, the following conclusions can be obtained immediately from Propositions 1 and 2.

Theorem 1. *In $K_S = (G, AT, I_S)$, let B, $C \subseteq AT$, then*

$$(1)\ B' \tilde{\subseteq} C' \Longleftrightarrow B^* \subseteq C^* \quad (2)\ B'' = B^{**}$$
$$(3)\ B' = C' \Longleftrightarrow B^* = C^* \quad (4)\ B \Rightarrow C \Longleftrightarrow B' \tilde{\subseteq} C'$$

Theorem 2. *In $K_S = (G, AT, I_S)$, let $\widetilde{A}, \widetilde{A}_1, \widetilde{A}_2$ be partitions of U and B, B_1, $B_2 \subseteq AT$, then*

$$(1)\ \widetilde{A}_1 \tilde{\subseteq} \widetilde{A}_2 \Rightarrow \widetilde{A}'_2 \subseteq \widetilde{A}'_1 \quad (2)\ B_1 \subseteq B_2 \Rightarrow B'_2 \tilde{\subseteq} B'_1$$
$$(3)\ \widetilde{A}\ \tilde{\subseteq}\ \widetilde{A}'';\ B \subseteq B'' \quad (4)\ \widetilde{A}' = \widetilde{A}''';\ B' = B'''$$

In $K_S = (G, AT, I_S)$, let $\widetilde{A}$ be a partition of U and $B \subseteq AT$. If $\widetilde{A}' = B$ and $B' = \widetilde{A}$, we say $(\widetilde{A}, B)$ is a rough concept of S, B is a rough intent, and $\widetilde{A}$ is a rough extent. The set of all rough concepts of S is denoted by $\mathcal{B}(S)$. Let $(\widetilde{A}_1, B_1)$ and $(\widetilde{A}_2, B_2)$ be two rough concepts of S, we define

$$(\widetilde{A}_1, B_1) \preceq (\widetilde{A}_2, B_2) \Longleftrightarrow \widetilde{A}_1 \tilde{\subseteq} \widetilde{A}_2 \Longleftrightarrow B_1 \supseteq B_2$$

The relation "$\preceq$" is the hierarchical order of rough concepts. Obviously, a lattice structure of S can be deduced, and it is a complete lattice called rough concept lattice of S. It's still denoted by $\mathcal{B}(S)$, if there is no danger of confusion.

Theorem 3. *In $K_S = (G, AT, I_S)$, let $B \subseteq AT$, if $\widetilde{A}$ is a partition of U, then*

$$(B', B''), (\widetilde{A}'', \widetilde{A}') \in \mathcal{B}(S)$$

Theorem 4. *In $K_S = (G, AT, I_S)$, let T be an index set. $(\widetilde{A}_t, B_t)$ is a rough concept of the information system S for every $t \in T$, and the partially ordered set $\mathcal{B}(S)$ is a complete lattice, where its infimum and supremum can be defined as*

$$\bigwedge_{t \in T} (\widetilde{A}_t, B_t) = \left(\widetilde{\bigcap_{t \in T}} \widetilde{A}_t, \left(\bigcup_{t \in T} B_t \right)'' \right)$$
$$\bigvee_{t \in T} (\widetilde{A}_t, B_t) = \left(\left(\widetilde{\bigcup_{t \in T}} \widetilde{A}_t \right)'', \bigcap_{t \in T} B_t \right)$$

For an information system S, the rough concept lattice $\mathcal{B}(S)$ can be built as follows:

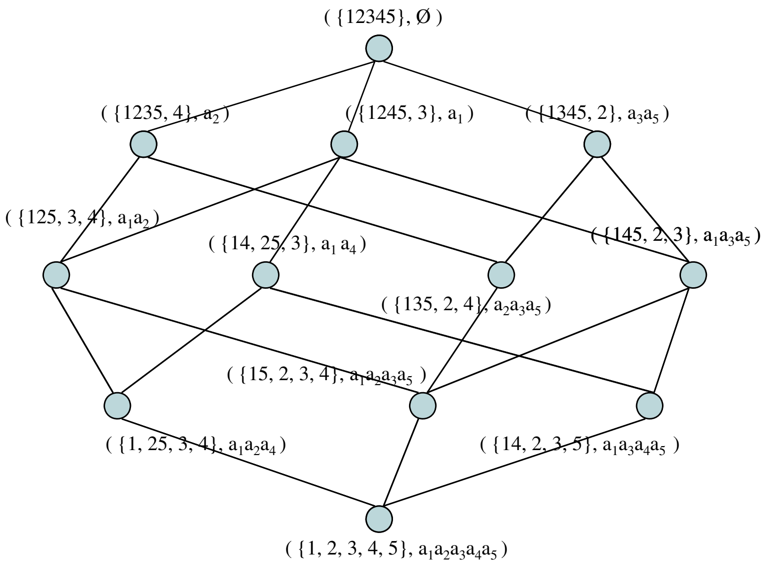

Fig. 4. The rough concept lattice with respect to Table 1.

1. Transform S to K_S.
2. Build the concept lattice $\mathcal{B}(K_S)$ of K_S using existing algorithms.
3. $\forall(A,B) \in \mathcal{B}(K_S), (A,B) \mapsto (U/A,B)$; that is, $\mathcal{B}(K_S) \to \mathcal{B}(S)$.
4. Output $\mathcal{B}(S)$.

For example, the rough concept lattice shown in Fig. 4 ($u_1,\ldots,u_5$ is simplified as 1, …, 5) with respect to Table 1 can be built using the above steps. For convenience, each set $\{P_1,P_2,\ldots,P_n\}$ is simplified as $P_1\ P_2 \cdots P_n$ in the following.

Since $B \to C \iff Ind(B) \subseteq Ind(C) \iff B^* \subseteq C^* \iff C^{**} \subseteq B^{**} \iff C \subseteq B^{**}(\iff C \subseteq B''$ or $B' \tilde{\subseteq}\ C'$), a function dependency of S is a rule for K_S in essence, as defined by Ganter and Wille [6]. Many researchers have studied rules extracted from concept lattices. Compared to other methods, rules extracted from a concept lattice have equal or better effects. In a rough concept lattice, by adopting existing methods with no or few changes, we can extract function dependencies in one-valued contexts, the process is not discussed in detail.

4. Applications of rough concepts in an information system

In the Pawlak rough set model, a pair of lower and upper approximation operators deduced from the approximation space play central roles. In the following we provide new definitions of the lower and upper approximation operators based on rough concepts. For convenience, we denote the set of all rough intents of S as $\mathcal{U}_S$ in the simplified form.

Theorem 5. *Let $B^+ = \bigcap \{\mathcal{C} \in \mathcal{U}_S | B \subseteq \mathcal{C}\}$ with $B \subseteq AT$, then*

(1) *$B^+ = B''$.*
(2) *$Ind(B^+) = Ind(B)$.*
(3) *If $(\mathcal{O}_B, B^+) \in \mathcal{B}(S)$, then $U/Ind(B) = \mathcal{O}_B$.*

Proof 1.

(1) Suppose $B^+ \subset B''$, then $B \subseteq B^+ \Rightarrow B'' \subseteq (B^+)''$ contradicts with $B^+ \subset B''$. Therefore $B'' \subseteq B^+$ holds. In addition, because $(B',B'') \in \mathcal{B}(S) \Rightarrow B^+ \subseteq B''$, we have $B^+ = B''$.
(2) $Ind(B^+) = Ind(B'') = (B'')^* = B^{***} = B^* = Ind(B)$.
(3) Since $(B',B''), (\mathcal{O}_B, B^+) \in \mathcal{B}(S)$ and $B'' = B^+$, there exists $\mathcal{O}_B = B'$. And then together with $U/Ind(B) = U/(B)^* = B'$, we can obtain $U/Ind(B) = \mathcal{O}_B$. □

From above discussions we know if the set $\mathcal{B}(S)$ of all rough concepts is given, then by Theorem 5 the lower approximation of $X \subseteq U$ relative to R can be obtained as follows:

$$\underline{R}(X) = \bigcup_{P \in \mathcal{O}_B \text{ and } P \subseteq X} P$$

Correspondingly, the upper approximation of X relative to R is:

$$\overline{R}(X) = \bigcup_{P \in \mathcal{O}_B \text{ and } P \bigcap X \neq \varnothing} P$$

where $R = Ind(B)$ and $(\mathcal{O}_B, B^+) \in \mathcal{B}(S)$.

In the following we discuss how to deal with some important problems in rough set theory, such as reducts, independents, cores and function dependencies based on $\mathcal{U}_S$. Some conclusions can then be obtained immediately.

- $Ind_S = \{B \subseteq AT | \forall a \in B, (B - \{a\})^+ \neq B^+\}$.
- $Core_S(B) = \{a \in B | (B - \{a\})^+ \neq B^+\}$.
- $C \in Red_S(B)$ if and only if $C^+ = B^+$ and $\nexists C_1 \subset C$ with $C_1^+ = B^+$.
- If $B, C \subseteq AT$, then $B \to C \Leftrightarrow C \subseteq B^+$.

Inspired by a previous study [12], we propose a new way of solving problems in rough set theory below.

Theorem 6. *Let B, $C \subseteq AT$, then following statements are equivalent*

(1) $C'' \subseteq B''$ *(If $B \subseteq C$, then $B'' = C''$)*
(2) *For any $L \in \mathcal{U}_S$, $B \not\subseteq L$ or $C \subseteq L$ holds*
(3) $B \to C$

Proof 2. "(1) $\leftrightarrow$ (2)": Fristly suppose (1) holds. For any $L \in \mathcal{U}_S$, if $B \not\subseteq L$, then (2) is true; for any $L \in \mathcal{U}_S$, if $B \subseteq L$, then $B'' \subseteq L'' = L$. Because $C'' \subseteq B''$, we can obtain $C'' \subseteq L \Rightarrow C \subseteq L$. Hence (2) is true. Then suppose (2) holds, this implies $\cap\{L \in \mathcal{U}_S | C \subseteq L\} \subseteq \cap\{L \in \mathcal{U}_S | B \subseteq L\} \Rightarrow C^+ \subseteq B^+ \Rightarrow C'' \subseteq B''$. In addition, if $B \subseteq C$, then there exists $B'' \subseteq C''$. Together with $C'' \subseteq B''$ we obtain $B'' = C''$. Hence (1) is true.

"(3) $\leftrightarrow$ (2)": Firstly suppose (3) holds, this implies $C \subseteq B^+ \Rightarrow C \subseteq B''$. For any $L \in \mathcal{U}_S$, if $B \not\subseteq L$, then (2) holds; if $B \subseteq L$, then there exists $B'' \subseteq L'' = L$ such that $B'' \subseteq L$, and together with $C \subseteq B''$ we get $C \subseteq L$. Therefore (2) is true. Then suppose (2) holds. Especially, $B \not\subseteq B''$ or $C \subseteq B''$ for B''. Because $B \subseteq B''$ denies $B \not\subseteq B''$, $C \subseteq B''$ holds. And further we can see from $B'' = B^+$ that $B \to C$. Therefore (3) is true. □

Theorem 7. *Let $B \subseteq AT$, then following statements hold*

(1) $IND_S = \{B \subseteq AT | \forall a \in B, (B - \{a\}) \subseteq L \text{ and } B \not\subseteq L, \exists L \in \mathcal{U}_S\}$
(2) $Core_S(B) = \{a \in B | (B - \{a\}) \subseteq L \text{ and } B \not\subseteq L, \exists L \in \mathcal{U}_S\}$

Proof 3. For any $a \in B$, if there exists $L \in \mathcal{U}_S$ satisfying $(B - \{a\}) \subseteq L$ and $B \not\subseteq L$, then we can see that $B'' \neq (B - \{a\})''$ from Theorem 6. Since $B'' \neq (B - \{a\})'' \Rightarrow B''' \neq (B - \{a\})''' \Rightarrow B' \neq (B - \{a\})' \Rightarrow B^* \neq (B - \{a\})^*$, $Ind(B) \neq Ind(B - \{a\})$ holds for any $a \in B$. And further conclusions (1) and (2) can be obtain immediately. □

Theorem 8. *Let B, $C \subseteq AT$, then $C \in Red_S(B)$ if and only if C is the minimum-subset satisfying the following condition.*

$C \cap B \not\subseteq L$ or $B \subseteq L$ holds for any $L \in \mathcal{U}_S$.

Proof 4. Let $C \in Red_S(B)$, if C is not the minimum-subset involved in AT, where C satisfies the above mentioned condition, then there exists $C_1 \subset C$ such that $C_1 \cap B \not\subseteq L$ or $B \subseteq L$ for any $L \in \mathcal{U}_S$. Because $C_1 \subset C \subseteq B$, we can confirm $C_1 \bigcap B = C_1$. This implies that $C_1 \not\subseteq L$ or $B \subseteq L$ holds for any $L \in \mathcal{U}_S$. Then by Theorem 6 we obtain $C_1'' = B''$. And further $C \notin Re\, d_S(B)$ can be deduced from $C_1'' = B'' \Longleftrightarrow C_1'' = B''' \Longleftrightarrow C_1' = B' \Longleftrightarrow C_1^* = B^* \Longleftrightarrow Ind(B) = Ind(C_1)$ and $C_1 \subset C$. Hence $C \notin Red_S(B)$ contradicts with the statement that C is a reduct of B, that is, C is the minimum-subset involved in AT, where C satisfies the condition.

Conversely, suppose C is the minimum-subset involved in AT, where C satisfies the condition. If $C \not\subseteq B$, then $C_1 = (B \cap C) \subset C$ and $C_1 \cap B = (B \cap C) \cap B = C \cap B \not\subseteq L$. It is clear that $C_1 \cap B \not\subseteq L$ or $B \subseteq L$ is true for any $L \in \mathcal{U}_S$, which contradicts with the statement of C being the minimum-subset, where C satisfies the condition. Hence we can confirm $C \subseteq B$. Suppose $C \notin Red_S(B)$, then there exists $C_1 \subset C \subseteq B$ such that $C_1 \in Red_S(B)$, and further $Ind(B) = Ind(C_1) \Longleftrightarrow C_1^* = B^* \Longleftrightarrow C_1' = B' \Longleftrightarrow C_1'' = B''$. It's obvious that there exists $C_1 = C_1 \bigcap B \not\subseteq L$ or $B \subseteq L$ for any $L \in \mathcal{U}_S$ by Theorem 6. Thus it contradicts the statement that C is the minimum-set involved in AT, where C satisfies the condition. Hence $C \in Re\, d_S(B)$. □

For Table 6, $\mathcal{U}_S$ is shown in Table 7. On the basis of Theorems 6–8, we can obtain the results shown in Tables 8 and 9 ($B = a_3a_5$, $C = a_2a_4a_5$, $D = a_1a_3a_4a_5$), Table 10 ($C = a_1a_2a_4$, $D = a_1a_3a_5$) and Table 11. In addition, by rough set theory we can calculate the same results. Examples demonstrate that the methods presented in Theorems 6–8 are valid and practicable.

5. Applications of rough concepts in decision tables

As one type of information system, a decision table plays an important role in decision applications. The majority of decision problems can be represented by decision tables.

Let $S = \{U, AT, V, f\}$ be an information system, where $AT = M \cup N$ and $M \cap N = \emptyset$. If M is called the set of conditional attributes and N is called the set of decision attributes, we say that $S = \{U, AT, V, f\}$ is a decision table. For example, Table 7 is a decision table, where $U = \{u_1, u_2, \ldots, u_8\}$, $M = \{a_1, a_2, \ldots, a_8\}$ and $N = \{d_1, d_2, \ldots, d_5\}$. We denote $U = \{u_1, u_2, \ldots, u_8\}$ as $U = \{1, 2, \ldots, 8\}$ in the simplified form.

A decision table consists of two information subsystems, a condition information subsystem $S_M = (U, M, V_M, f_M)$ and a decision information subsystem $S_N = (U, N, V_N, f_N)$, where $V_M = \bigcup_{a \in M} V_a$, f_M is a mapping of $U \times M$ to V_M with $V_N = \bigcup_{d \in N} V_d$; f_N is a mapping of $U \times N$ to V_N. Accordingly, we say that $(\widetilde{A}, B) \in \mathcal{B}(S_M)$ is a condition rough concept and $(\widetilde{C}, D) \in \mathcal{B}(S_N)$ is a decision rough concept, where $B \subseteq M$ and $D \subseteq N$.

5.1. Applications of rough concepts in decision rules

Suppose $(\widetilde{A}, B) \in \mathcal{B}(S_M)$ and $(\widetilde{C}, D) \in \mathcal{B}(S_N)$, then we say $(\widetilde{A}, B) \overset{\text{rough}}{\rightarrow} (\widetilde{C}, D)$ is a decision package composed of a set of decision rules, where the decision rule is defined as follows. let $X \in \widetilde{A}$ and $Y \in \widetilde{C}$, $des_B(X)$ denotes the description of equivalence class X, and $des_D(Y)$ denotes the description of equivalence class Y. A decision rule r is defined as

Table 7
The intent set $\mathcal{U}_S = \{L_1, \ldots, L_{12}\}$ of K_S.

L_1	ø	L_5	a_1a_2	L_9	$a_1a_2a_4$
L_2	a_1	L_6	a_1a_4	L_{10}	$a_1a_3a_4a_5$
L_3	a_2	L_7	$a_1a_3a_5$	L_{11}	$a_1a_2a_3a_5$
L_4	a_3a_5	L_8	$a_2a_3a_5$	L_{12}	$a_1a_2a_3a_4a_5$

Table 8
Cores of some attribute-subsets on the basis of $\mathcal{U}_S$.

T	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	$Core_S(T)$
$T = B$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$				
$B - \{a_3\}$				$\subseteq$			$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	$a_3 \notin$
$B - \{a_5\}$				$\subseteq$			$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	$a_5 \notin$
$T = C$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		
$C - \{a_2\}$										$\subseteq$		$\subseteq$	$a_2 \in$
$C - \{a_4\}$								$\subseteq$			$\subseteq$	$\subseteq$	$a_4 \in$
$C - \{a_5\}$									$\subseteq$			$\subseteq$	$a_5 \in$
$T = D$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		
$D - \{a_1\}$										$\subseteq$		$\subseteq$	$a_1 \notin$
$D - \{a_3\}$										$\subseteq$		$\subseteq$	$a_3 \notin$
$D - \{a_4\}$							$\subseteq$			$\subseteq$	$\subseteq$	$\subseteq$	$a_4 \in$
$D - \{a_5\}$										$\subseteq$		$\subseteq$	$a_5 \notin$

Table 9
Some computing results on IND_S on the basis of $\mathcal{U}_S$.

T	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	IND_S
$B - \{a_3\}$				$\subseteq$			$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	
$B - \{a_5\}$				$\subseteq$			$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	
$T = B$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$				$\notin$
$C - \{a_2\}$										$\subseteq$		$\subseteq$	
$C - \{a_4\}$								$\subseteq$			$\subseteq$	$\subseteq$	
$C - \{a_5\}$									$\subseteq$			$\subseteq$	
$T = C$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\in$
$D - \{a_1\}$										$\subseteq$		$\subseteq$	
$D - \{a_3\}$										$\subseteq$		$\subseteq$	
$D - \{a_4\}$							$\subseteq$			$\subseteq$	$\subseteq$	$\subseteq$	
$D - \{a_5\}$										$\subseteq$		$\subseteq$	
$T = D$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		$\notin$

Table 10
Reducts of some attribute-subsets on the basis of $\mathcal{U}_S$.

T	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	$Red_S(T)$
$T = C$									$\subseteq$			$\subseteq$	
$a_1 \cap T$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$				$\not\subseteq$					
$a_2 \cap T$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$			
$a_4 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$		
$a_1a_2 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$			
$a_1a_4 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$		
$\underline{a_2a_4} \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$		$\in$
$\underline{C} \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$		$\notin$
$T = D$							$\subseteq$			$\subseteq$	$\subseteq$	$\subseteq$	
$a_1 \cap T$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$				$\not\subseteq$					
$a_3 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$				
$a_5 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$				
$\underline{a_1a_3} \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$				$\in$
$\underline{a_1a_5} \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$				$\in$
$a_3a_5 \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$				
$\underline{D} \cap T$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$				$\notin$

Table 11
Some functional dependencies on the basis of $\mathcal{U}_S$.

	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	
a_4	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			$\not\subseteq$		
a_1		$\subseteq$			$\subseteq$	$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	$\subseteq$	$a_4 \to a_1$
a_2a_5	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$	$\not\subseteq$			
a_3				$\subseteq$			$\subseteq$	$\subseteq$		$\subseteq$	$\subseteq$	$\subseteq$	$a_2a_5 \to a_3$
a_4a_5	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		
a_1a_2					$\subseteq$				$\subseteq$		$\subseteq$	$\subseteq$	$a_4a_5 \nrightarrow a_1a_2$
a_3a_4	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		
a_1a_5							$\subseteq$			$\subseteq$	$\subseteq$	$\subseteq$	$a_3a_4 \to a_1a_5$
a_1a_4	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$	$\not\subseteq$		$\not\subseteq$		
a_2a_3					$\subseteq$				$\subseteq$		$\subseteq$	$\subseteq$	$a_1a_4 \nrightarrow a_2a_3$

$$r : des_B(X) \Rightarrow des_D(Y)$$

The corresponding certainty factor is defined as

$$\mu(X, Y) = |Y \cap X|/|X|, \; 0 < \mu(X, Y) \leqslant 1$$

- when $\mu(X,Y) = 1$, r is a certainty decision rule,
- when $0 < \mu(X,Y) < 1$, r is a uncertainty decision rule.
  For example, in Table 12 the decision package $(\{1234, \; 56, \; 78\}, a_1a_3a_7) \in \mathcal{B}(S_M) \overset{\text{rough}}{\rightarrow} (\{12, \; 3456, \; 78\}, \; d_3d_5) \in \mathcal{B}(S_N)$ contains following decision rules:
- Certainty decision rules:

$$(a_1, 0) \; and \; (a_3, 0) \; and \; (a_7, 1) \Rightarrow (d_3, 0) \; and \; (d_5, 0)$$
$$(a_1, 1) \; and \; (a_3, 0) \; and \; (a_7, 1) \Rightarrow (d_3, 1) \; and \; (d_5, 1)$$

- Uncertainty decision rules:

$$(a_1, 1) \; and \; (a_3, 1) \; and \; (a_7, 0) \Rightarrow (d_3, 0) \; and \; (d_5, 1),$$

the certainty factor of the decision rule is 0.5.

$$(a_1, 1) \; and \; (a_3, 1) \; and \; (a_7, 0) \Rightarrow (d_3, 0) \; and \; (d_5, 0),$$

the certainty factor of the decision rule is 0.5.

In a decision package, the significance of different condition attributes (or sets of attributes) relative to the set of decision attributes may differ. To measure the significance of the condition attribute (or sets of attributes) $B_1 \subseteq B$ relative to D in the decision package $(\widetilde{A}, B) \overset{\text{rough}}{\rightarrow} (\widetilde{C}, D)$, we define the importance factor $\rho_{BD}(B_1)$ based on rough concepts as follows:

Table 12
A decision table.

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	d_1	d_2	d_3	d_4	d_5
u_1	1	1	1	2	1	0	0	0	1	2	0	1	1
u_2	1	1	1	1	0	2	0	0	1	2	0	1	1
u_3	1	2	1	2	1	2	0	1	0	2	0	2	0
u_4	1	2	1	2	1	2	0	1	0	2	0	2	0
u_5	0	1	0	1	0	1	1	1	1	0	0	0	0
u_6	0	1	0	0	1	1	1	1	1	0	0	0	0
u_7	1	0	0	0	1	0	1	0	1	1	1	2	1
u_8	1	0	0	1	0	0	1	0	1	1	1	2	1

$$\rho_{BD}(B_1) = \frac{1}{|U|}\cdot\left(\left|\bigcup_{Y\in\widetilde{C}}\underline{R_1}Y\right| - \left|\bigcup_{Y\in\widetilde{C}}\underline{R_2}Y\right|\right)$$
$$\underline{R_1}Y = \bigcup_{P\subseteq Y \text{ and } P\in\widetilde{A}} P$$
$$\underline{R_2}(Y) = \bigcup_{P\subseteq Y \text{ and } P\in\mathcal{O}_{B-B_1}} P$$

where $(\mathcal{O}_{B-B_1}, (B-B_1)^+) \in \mathcal{B}(S), R_1 = Ind(B)$ and $R_2 = Ind(B - B_1)$. Greater $\rho_{BD}(B_1)$ indicates that B_1 is more important relative to D. As an example, in Table 7 we can obtain a decision package

$$(\{1234, 56, 78\}, B) \in \mathcal{B}(S_M) \overset{\text{rough}}{\rightarrow} (\{12, 3456, 78\}, D) \in \mathcal{B}(S_N)$$

the importance factor $\rho_{BD}(B_1)$ of $B_1 \subseteq B$ relative to D is shown in Table 13, where $B = a_1a_3a_7$ and $D = d_3d_5$.

From Table 8 we can see that a_1a_3 and a_1a_7 are more important than a_3a_7, and a_1 is more important than a_3 and a_7 relative to D.

5.2. Applications of rough concepts in decision dependencies

Since decision dependency has become a common form of knowledge representation because of its properties of expressiveness and ease of understanding, it has been widely used in practice. Therefore, in this section we focus on decision dependencies in a decision table.

In $S = \{U, M \cup N, V, f\}$, if $B \subseteq M$ and $D \subseteq N$, then a function dependency $B \to D$ is called a decision dependency; If $(B', B) \in \mathcal{B}(S_M)$ and $(D', D) \in \mathcal{B}(S_N)$, we say that $B \to D$ is a concept decision dependency.

For convenience, some formal symbols are defined as follows. $\widetilde{A}'$ in S_M is denoted as $\widetilde{A}^{I_1}$ and $\widetilde{A}'$ in S_N is denoted as $\widetilde{A}^{I_2}$. Correspondingly, B' in S_M is denoted as B^{I_1} and B' in S_N is denoted as B^{I_2}. B^+ is denoted as B^+_M with respect to $B \subseteq M$, and B^+ is denoted as B^+_N with respect to $B \subseteq N$.

Theorem 9. *Let B_1, $B_2 \subseteq M$, D_1, $D_2 \subseteq N$. If $B \to D$ is a concept decision dependency, then*

(1) *If $B \subseteq B_1$ and $D_1 \subseteq D$, then $B_1 \to D_1$.*
(2) *If $(B_2)^+_M = B$ and $(D_2)^+_N = D$, then $B_2 \to D_2$.*

Proof 5.

(1) Because $B \to D$ is a concept decision dependency, there exists $B^{I_1} \tilde{\subseteq} D^{I_2}$. In addition, $B_1^{I_1} \tilde{\subseteq} B^{I_1}$ and $D^{I_2} \tilde{\subseteq} D_1^{I_2}$ can be deduced from $B_1 \supseteq B$ and $D_1 \subseteq D$. Hence $B_1^{I_1} \tilde{\subseteq} D_1^{I_2}$, that is, $B_1 \to D_1$ holds.
(2) we can obtain $B_2^{I_1 I_1} = B$ and $D_2^{I_2 I_2} = D$ on the basis of $(B_2)^+_M = B$ and $(D_2)^+_N = D$, then $B_2^{I_1} = B^{I_1}$ and $D_2^{I_2} = D^{I_2}$ hold. In addition, since $B \to D$ is a concept decision dependency, there exists $B^{I_1} \tilde{\subseteq} D^{I_2}$ such that $B_2^{I_1} \tilde{\subseteq} D_2^{I_2}$. That is, $B_2 \to D_2$ holds. □

In general, the number of decision dependencies in a decision table is quite large, and in a given set there are always many redundant decision dependencies that can be deduced from others by means of so-called $\alpha\beta$-decision inference.

Table 13
The importance factor of $B_1 \subseteq B$ relative to D in $(\widetilde{A}, B) \overset{\text{rough}}{\rightarrow} (\widetilde{C}, D)$.

B_1	a_1	a_3	a_7	a_1a_3	a_1a_7	a_3a_7
$\rho_{BD}(B_1)$	0.5	0	0	0.5	0.5	0.25

($\alpha\beta$-DECISION INFERENCE) Let $B \to D$ and $B_1 \to D_1$ be concept decision dependencies, and $B_2 \to D_2$ is a decision dependency, then

- (α-INFERENCE RULE) If $B \subseteq B_1$ and $D_1 \subseteq D$, then $B_1 \to D_1$ can be inferred from $B \to D$.
- (β-INFERENCE RULE) If $(B_2)_M^+ = B$ and $(D_2)_N^+ = D$, then $B_2 \to D_2$ can be inferred from $B \to D$.

The α-inference rule can be characterized by the following form:

$$\frac{B \subseteq B_1,\ D_1 \subseteq D,\ \text{concept decision dependency } B \to D}{\text{concept decision dependency } B_1 \to D_1}$$

which means that, if $B \subseteq B_1$, $D_1 \subseteq D$ and $B \to D$ is a concept decision dependency, then a concept decision dependency $B_1 \to D_1$ can be inferred. In a similar way, the β-inference rule can be characterized in the following form:

$$\frac{(B_2)_M^+ = B, (D_2)_N^+ = D,\ \text{concept decision dependency } B \to D}{B_2 \to D_2}$$

which means that, if $(B_2)_M^+ = B, (D_2)_N^+ = D$, and $B \to D$ is a concept decision dependency, then a decision dependency $B_2 \to D_2$ can be inferred.

Let Σ be a set of decision dependencies and $B \to C$ be a decision dependency. If $B \to C$ can be inferred from Σ by some decision inference τ, we say $B \to C$ can be τ-inferred from Σ. In this case, we call $B \to C$ is redundant relative to Σ. Furthermore, if all decision dependencies of S can be τ-inferred from $\sum$, we say Σ is τ-complete relative to S. In addition, a concept decision dependency $B \to D$ is maximal, if

(1) there is no concept decision dependency $B_1 \to D$ satisfying $B_1 \subset B$.
(2) there is no concept decision dependency $B \to D_1$ satisfying $D \subset D_1$.

In this case, we also say that $B \to D$ is a maximal concept decision dependency in S.

Theorem 10. *In a decision table S, the set Σ of all maximal concept decision dependencies is $\alpha\beta$-complete and $\alpha\beta$-non-redundant.*

Proof 6. Firstly, we prove that Σ is $\alpha\beta$-non-redundant. Assume Σ is $\alpha\beta$-redundant, then there must exist a maximal concept decision dependency $B_1 \to D_1$ which can be inferred from $B_2 \to D_2$ in $\Sigma \backslash (B_1 \to D_1)$ by the α-inference rule. Obviously we have $B_2 \subseteq B_1$ and $D_1 \subseteq D_2$. In addition, since there exists $B_1 \to D_1 \notin \Sigma \backslash (B_1 \to D_1)$, there must exist $B_1 \neq B_2$ or $D_1 \neq D_2$.

If $D_1 \neq D_2$, then $D_1 \subset D_2$. Since there exists $B_2 \subseteq B_1$, then $B_1^{I_1} \tilde{\subseteq} B_2^{I_1}$ holds. In addition, we can obtain $B_2^{I_1} \tilde{\subseteq} D_2^{I_2}$ by $B_2 \to D_2$, then $B_1^{I_1} \tilde{\subseteq} D_2^{I_2}$ holds, that is, $B_1 \to D_2$ holds. Obviously, $B_1 \to D_2$ contradicts with the condition of $B_1 \to D_1$ being the maximum concept decision dependency.

When $D_1 = D_2$ and $B_1 \neq B_2$, because there exists $B_2 \subseteq B_1$, then $B_2 \subset B_1$ holds. Since there exists $D_1 = D_2$, $B_2 \to D_1$ holds, which contradicts with the condition of $B_1 \to D_1$ being the maximal concept decision dependency. Hence one can see Σ is $\alpha\beta$-non-redundant from above discussions.

Next, we prove that Σ is $\alpha\beta$-complete. Let $B_2 \to D_2$ be a decision dependency, one can easily see that $B_2 \to D_2$ can be inferred from the concept decision dependency $(B_2)_M^+ \to (D_2)_N^+$ by the β-inference rule. And further we can see that there must exist a maximal concept decision dependency $B_1 \to D_1$ and $(B_2)_M^+ \to (D_2)_N^+$ can be inferred from $B_1 \to D_1$ by the α-inference rule, hence $B_2 \to D_2$ can be inferred from $B_1 \to D_1$. It indicates that Σ is $\alpha\beta$-complete. □

This section was inspired by the work of Qu et al. [32] and can be viewed as an extension of that work. Based on $\mathcal{B}(S_M)$ and $\mathcal{B}(S_N)$, steps for generating an $\alpha\beta$-complete and $\alpha\beta$-non-redundant set Σ of decision dependencies are listed below. However, the time complexity is very high, which is not desirable, especially for larger experiments. Thus, the following method only serves as a basis for opportunities for further development.

1. For every condition rough concept $(\widetilde{A}, B)$, find $(\widetilde{C}_i, D_i)$, which is the minimal-decision rough concept satisfying $B \to D_i$; add $(\widetilde{C}_i, D_i)$ to Δ.
2. Choose $(\widetilde{C}_j, D_j)$ from Δ randomly and find $(\widetilde{A}_k, B_k)$, which is the maximal-condition rough concept satisfying $B_k \to D_j$; delete $(\widetilde{C}_j, D_j)$ from Δ.
3. Add $B_k \to D_j$ to Σ; if $\Delta \neq \emptyset$, then switch to step 2.
4. Output Σ.

From Table 12, we can obtain an $\alpha\beta$-complete and an $\alpha\beta$-non-redundant set Σ of decision dependencies using the methods described above. The experimental results are shown in Tables 14–16.

In Table 12, according to the $\alpha\beta$-decision inference rule, it is clear that all the decision dependencies can be inferred from the decision dependencies in Table 16 based on Tables 14 and 15. For example, since $a_1a_3a_7 = (a_1a_3)_M^+$ and $d_3 = (d_3)_N^+$ for the decision dependency $a_1a_3 \to d_3$, then $a_1a_3 \to d_3$ can be deduced from $a_1a_3a_7 \to d_3$ by the β-inference rule. In addition, because there exist $a_1a_3a_7 \subseteq a_1a_3a_7$ and $d_3 \subseteq d_2d_3$, then $a_1a_3a_7 \to d_3$ can be inferred from the decision dependency $a_1a_3a_7 \to d_2d_3$ by the α-inference rule. It is clear that $a_1a_3 \to d_3$ can be inferred from $a_1a_3a_7 \to d_2d_3$ by the $\alpha\beta$-decision inference.

Table 14
Condition rough concepts in Table 12.

Rough intent	Rough extent	Rough intent	Rough extent
$a_1a_2a_3a_7a_8$	12, 34, 56, 78	a_1a_6	178, 234, 56
$a_1a_3a_4a_5a_7$	134, 2, 5, 6, 7, 8	$a_2a_4a_5$	1, 25, 34, 6, 7, 8
$a_1a_3a_7$	1234, 56, 78	a_4a_5	134, 258, 67
a_2	1256, 34, 78	a_8	1278, 3456
a_2a_5	16, 25, 34, 7, 8	$a_1a_4a_5a_8$	1, 28, 34, 5, 6, 7
a_5	13467, 258	$a_1a_4a_5$	134, 28, 5 6, 7
$a_1a_5a_6a_8$	17, 34, 2, 5, 6, 8	$a_1a_5a_8$	17, 28, 34, 5, 6
a_1a_8	1278, 34, 56	$a_1a_2a_3a_4a_5a_6a_7a_8$	1, 2, 34, 5, 6, 7, 8
a_1a_5	1347, 28, 5, 6	a_5a_8	17, 28, 346, 5
a_1	123478, 56	$a_1a_2a_3a_6a_7a_8$	1, 2, 34, 56, 78
$a_1a_6a_8$	178, 2, 34, 56	a_3a_7	1234, 5678
$a_1a_3a_6a_7$	1, 234, 56, 78	$a_3a_4a_5a_7$	134, 2, 58, 67
ø	12345678		

Table 15
Decision rough concepts in Table 12.

Rough intent	Rough extent	Rough intent	Rough extent
d_2d_3	1234, 56, 78	$d_1d_2d_3d_4d_5$	12, 34, 56, 78
d_1	125678, 34	d_1d_3	1256, 34, 78
d_5	1278, 3456	d_1d_5	1278, 34, 56
d_4	12, 3478, 56	d_3d_5	12, 3456, 78
d_3	123456, 78	ø	12345678

Table 16
A $\alpha\beta$-complete and $\alpha\beta$-not-redundant set Σ in Table 12.

$\emptyset \to \emptyset$	$a_8 \to d_5$	$a_1a_8 \to d_1d_5$
$a_2 \to d_1d_3$	$a_1a_3a_7 \to d_2d_3$	$a_1a_2a_3a_7a_8 \to d_1d_2d_3d_4d_5$

6. Conclusions

This paper introduces FCA into rough set theory naturally and proposes a rough set model based on FCA that can be viewed as expansion and application of the theories of Ganter and Wille [6] to rough set theory. In this model, we provide a solution to the problem of algebraic structure in an information system; that is, a rough concept lattice is inferred from the information system. We also investigated applications of rough concepts in rough set theory. In addition, since the number of decision dependencies in a decision table increases exponentially with the scale of the decision table, we presented some inference rules to eliminate superfluous decision dependencies. Thus, we can obtain a complete and non-redundant set of decision dependencies from a decision table. In future research we will investigate whether this theory can be widely applied to some special rough set models, such as the variable-precision rough set model, the probability rough set model, the fuzzy rough set model, and rough set models based on random sets. Exploration of wider combinations of FCA and rough set theory will also be a focus of our future work.

Acknowledgements

The work was supported by the National Natural Science Foundation of China (Nos. 60970014, 61070100, 61175067, 61100058, 61170059 and 60875040), the Natural Science Foundation of Shanxi, China (Nos. 2010011021-1 and 2011021013-2), the Foundation of Doctoral Program Research of Ministry of Education of China (No. 200801080006), Shanxi Foundation of Tackling Key Problem in Science and Technology (No. 20110321027-02), the Natural Science Foundation of Anhui Province of China (KJ2011A086) and the Graduate Innovation Project of Shanxi Province, China (No. 20103004).

References

[1] Y.H. Chen, Y.Y. Yao, A multiview approach for intelligent data analysis based on data operators, Information Sciences 178 (1) (2008) 1–20.
[2] Y. Cheng, D.Q. Miao, Q.R. Feng, Positive approximation and converse approximation in interval-valued fuzzy rough sets, Information Sciences 181 (11) (2011) 2086–2110.
[3] I. Düntsch, G. Gediga, Approximation operators in qualitative data analysis, in: H. de Swart, E. Orlowska, G. Schmidt, M. Roubens (Eds.), Theory and Application of Relational Structures as Knowledge Instruments, Springer, Heidelberg, 2003, pp. 216–233.
[4] A.A. Estaji, S. Khodaii, S. Bahrami, On rough set and fuzzy sublattice, Information Sciences 181 (18) (2011) 3981–3994.

[5] F. Feng, Y.M. Li, V. Leoreanu-Fotea, Y.B. Jun, Soft sets and soft rough sets, Information Sciences 181 (6) (2011) 1125–1137.
[6] B. Ganter, R. Wille, Formal Concept Analysis:Mathematical Foundations, Springer-Verlag, Berlin, 1999.
[7] G. Gediga, I. Düntsch. Modal-style operators in qualitative data analysis, in: Proceedings of the 2002 IEEE International Conference on Data Mining, 2002, pp. 155–162.
[8] L.K. Guo, F.P. Huang, Q.G. Li, G. Q Zhang, Power contexts and their concept lattices, Discrete Mathematics 311 (18–19) (2011) 2049–2063.
[9] T.B. Ho, Acquiring concept approximations in the framework of rough concept analysis, in: Proceedings of the 7th European–Japanese Conference on Information Modelling and Knowledge Bases, 1997, pp. 186–195.
[10] K. Hu, Y. Sui, Y. Lu, J. Wang, C. Shi, Concept approximation in concept lattice, in: Proceedings of the 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD'01, 2001, pp. 167–173.
[11] X.G. Hu, Y.H. Zhang, X.Y. Wang, The Representation and Resolution of Rough Sets Based on the Extended Concept Lattice, in: Proceedings of FSKD(1), 2005, pp. 1309–1312.
[12] J. Järvinen. Representation of Information Systems and Dependence Spaces, and Some Basic Algorithms. Licentiate's thesis. Ph.D. Thesis, University of Turku, Department of Mathematics, Turku, Finland, 1997.
[13] M. Kaytoue, S. Duplessis, S.O. Kuznetsov, A. Napoli, Mining gene expression data with pattern structures in formal concept analysis, Information Sciences 181 (10) (2011) 1989–2001.
[14] R.E. Kent, Rough concept analysis, Fundamenta Informaticae 27 (1996) 169–181.
[15] H. L Lai, D.X. Zhang, Concept lattices of fuzzy contexts: formal concept analysis vs. rough set theory, International Journal of Approximate Reasoning 50 (5) (2009) 695–707.
[16] J.H. Li, C.L. Mei, Y.J. Lv, Knowledge reduct in decision formal contexts, Knowledge-Based Systems 24 (5) (2011) 709–715.
[17] H.R. Li, W.X. Zhang, H. Wang. Classification and reduct of attributes in concept lattices, in: Proceedings of the 2006 IEEE International Conference on Granular Computing, GrC'06, 2006, pp. 142–147.
[18] M. Liu, M.W. Shao, W.X. Zhang, C. Wu, Reduct method for concept lattices based on rough set theory and its application, Computers and Mathematics with Applications 53 (9) (2007) 1390–1410.
[19] J.M. Ma, W.X. Zhang, X. Wang, Dependence space of concept lattice based on rough set, in: Proceedings of the 2006 IEEE International Conference on Granular Computing, GrC'06, 2006, pp. 200–204.
[20] J. Medina, M. Ojeda-Aciego, Multi-adjoint t-concept lattices, Information Sciences 180 (5) (2010) 712–725.
[21] J.S. Mi, Y. Leung, W.Z. Wu, Approaches to attribute reduct in concept lattices induced by axialities, Knowledge-Based Systems 23 (6) (2010) 504–511.
[22] G.D. Oosthuizen, Rough sets and concept lattices, in: W.P. Ziarko (Ed.), Rough Sets, and Fuzzy Sets and Knowledge Discovery (RSKD'93), Springer-Verlag, London, 1994, pp. 24–31.
[23] J. Outrata, V. Vychodil, Fast algorithm for computing fixpoints of Galois connections induced by object–attribute relational data, Information Sciences 185 (1) (2012) 114–127.
[24] P. Pagliani, From concept lattices to approximation spaces: algebraic structures of some spaces of partial objects, Fundamenda Informaticae 18 (1993) 1–25.
[25] P. Pagliani, M.K. Chakraborty, Information quanta and approximation spaces I: Non-classical approximation operators, in: Proceedings of the 2005 IEEE International Conference on Granular Computing, GrC'05, 2005, pp. 605–610.
[26] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.
[27] D.W. Pei, Z.B. Xu, Rough set models on two universes, International Journal of General Systems 33 (2004) 569–581.
[28] J. Pócs, Note on generating fuzzy concept lattices via Galois connections, Information Sciences 185 (1) (2012) 128–136.
[29] J.J. Qi, L. Wei, Z.Z. Li. A partitional view of concept lattice, in: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, RSFDGrC'05, I, LNAI 3641, 2005, pp. 74–83.
[30] Y.H. Qian, J.Y. Liang, W. Pedrycz, C.Y. Dang, An efficient accelerator for attribute reduction from incomplete data in rough set framework, Pattern Recognition 44 (8) (2011) 1658–1670.
[31] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary GrC model, IEEE Transactions on Fuzzy Systems 19 (2) (2011) 253–264.
[32] K.S. Qu, Y.H. Zhai, J.Y. Liang, M. Chen, Study of decision implication based on formal concept analysis, International Journal of General Systems 36 (2) (2007) 147–156.
[33] J. Saquer, J. Deogun. Formal rough concept analysis, in: Proceedings of the 7th International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing, RSFDGrC'99, LNAI1711, 1999, pp. 91–99.
[34] J. Saquer, J. Deogun, Concept approximations based on rough sets and similarity measures, International Journal of Applied Mathematics and Computer Science 11 (2001) 655–674.
[35] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
[36] M.W. Shao, M. Liu, W.X. Zhang, Set approximations in fuzzy formal concept analysis, Fuzzy Sets and Systems 158 (23) (2007) 2627–2640.
[37] M.W. Shao, W.X. Zhang, Approximation in formal concept analysis, in: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'05, I, LNAI 3641, 2005, pp. 43–53.
[38] Z.H. Shi, Z.T. Gong, The further investigation of covering-based rough sets: uncertainty characterization, similarity measure and generalized models, Information Sciences 180 (19) (2010) 3745–3763.
[39] J.Y. Shyng, H.M. Shieh, G.H. Tzeng, An integration method combining Rough Set Theory with formal concept analysis for personal investment portfolios, Knowledge-Based Systems 23 (6) (2010) 586–597.
[40] L.D. Wang, X.D. Liu, Concept analysis via rough set and AFS algebra, Information Sciences 178 (21) (2008) 4125–4137.
[41] H. Wang, W.X. Zhang, Relationships between concept lattice and rough set, Lecture Notes in Computer Science, 4029, Springer-Verlag, 2006.
[42] P. Wasilewski, Concept lattices vs. approximation spaces, in: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'05, I, LNAI 3641, 2005, pp. 114–123.
[43] L. Wei, J.J. Qi, Relation between concept lattice reduct and rough set reduct, Knowledge-Based Systems 23 (8) (2010) 934–938.
[44] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: I. Rival (Ed.), Ordered Sets, Reidel, Dordrecht-Boston, 1982, pp. 445–470.
[45] M. Wolski, Galois connections and data analysis, Fundamenta Informaticae 60 (2004) 401–415.
[46] M. Wolski, Formal Concept Analysis and Rough Set Theory from the Perspective of Finite Topological Approximations, Transactions on Rough Sets, III, LNCS, Springer, 2005, pp. 230–243.
[47] Q. Wu, Z.T. Liu, Real formal concept analysis based on grey-rough set theory, Knowledge-Based Systems 22 (1) (2009) 38–45.
[48] Q. Wu, Z.T. Liu, Y. Li, Rough formal concepts and their accuracies, in: Proceedings of the 2004 International Conference on Services Computing, SCC'04, 2004, pp. 445–448.
[49] H.Z. Yang, Y. Leung, M.W. Shao, Rule acquisition and attribute reduct in real decision formal contexts, Soft Computing 15 (6) (2011) 1115–1128.
[50] Y.Y. Yao, A comparative study of formal concept analysis and rough set theory in data analysis, in: Proceedings of the 3rd International Conference on Rough Sets and Current Trends in Computing, RSCTC'04, LNCS, 2004, pp. 59–68 (3066).
[51] Y.Y. Yao, Concept lattices in rough set theory, in: Proceedings of the 23rd International Meeting of the North American Fuzzy Information Processing Society, 2004, pp. 796–801.
[52] Y.Y. Yao, On unifying formal concept analysis and rough set analysis, in: W.X. Zhang, Y.Y. Yao, Y. Leung (Eds.), Rough Set and Concept Lattice, Xi'an Jiaotong University Press, Xi'an, 2006, pp. 1–20.
[53] Y.Y. Yao, A partition model of granular computing, in: LNCS Transactions on Rough Sets I, LNCS 3100, 2004, pp. 232–253.

[54] Y.Y. Yao, Y.H. Chen, Rough set approximations in formal concept analysis, LNCS Transactions on Rough Sets V 4100 (2006) 285–305.
[55] W.X. Zhang, L. Wei, J.J. Qi, Attribute reduct in concept lattice based on discernibility matrix, in: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'05, Part II, LNAI 3642, 2005, pp. 157–165.
[56] X.Y. Zhang, Z.W. Mo, F. Xiong, W. Cheng, Comparative study of variable precision rough set model and graded rough set model, International Journal of Approximate Reasoning 53 (1) (2012) 104–116.
[57] J. Zhao, L. Liu, Construction of concept granule based on rough set and representation of knowledge-based complex system, Knowledge-Based Systems 24 (6) (2011) 809–815.