Attribute reduction based on maximal rules in decision formal context

Leijun Li[a,b], Jusheng Mi[a*], Bin Xie[a]

[a] *College of Mathematics and Information Science,*
Hebei Normal University, Shijiazhuang, 050016, P. R. China
[b] *Harbin Institute of Technology, Harbin, 150001, P. R. China*
E-mail: lileijun1985@163.com, mijsh@263.net, happyxiebin@126.com

Abstract

One of the key issues in the theory of concept lattices is to extract the useful rules from the decision formal context. The maximal rules implicate the others, thus people are interested in them. This paper proposes two new kinds of attribute reduction in the decision formal context based on maximal rules. The reducts preserve all the condition extensions and the decision extensions related to the original maximal rules. The internal relationship between the original maximal rules and the maximal rules in the reduced decision formal context is derived. The reducts can make the maximal rules more concise and accurate. The mathematical property of the proposed attribute reduction is investigated and we construct the discernibility matrix and function to compute all the reducts. Finally, all the attributes are classified into three types based on the maximal rules. The characteristics of these types of attributes are also analyzed.

Keywords: concept lattices; maximal rule; condition reduct; decision reduct

1. Introduction

The theory of concept lattices (also called formal concept analysis), proposed by Wille in 1982, is a relatively new model to represent the formal concepts associated with a context [30]. It can unravel hierarchical concepts organized as a lattice from relational information systems. A formal context consists of a universe of discourse U, an attributes set A, and a binary relation I from U to A. All the formal concepts are constructed from the formal context and form a concept lattice which reflects the relationship of generalization and the specialization among concepts. It has been studied from various perspectives, such as its mathematical property [4,8,13,17,18,25] and the applications in knowledge discovery [3,5,6,26,33,35].

In this paper, we investigate attribute reduction based on maximal rules in the decision formal context. Attribute reduction is performed in information systems by means of the notion of a reduct based on a specialization of the general notion of independence [15]. After eliminating the superfluous attributes, the formal context becomes simpler and the discovery of the initial knowledge hidden in the formal context becomes easier, thus it is a key problem in knowledge discovery. In recent years, many scholars have studied attribute reduction extensively from various perspectives and each of them aims at some basic requirements [14,16,19,27,32,36]. In 2005, a method to attribute reduction was proposed in concept lattices via the discernibility matrix and Boolean function [36]. It can get all the reducts and each of them preserves all the original concepts and

their hierarchy in the reduced context. Then inspired by rough set theory [22], two kinds of attribute reduction methods and object reduction methods for concept lattices were developed, respectively [14]. After that, based on irreducible elements in the object and property oriented concept lattices, a new approach to attribute reduction was presented [27]. Recently, the relation among attribute reductions in formal, object-oriented and property-oriented concept lattices is studied and the equivalence among them has also been proved [16]. Besides, much attention has been paid to attribute reduction of the decision formal context [11,12,28,29,34]. A decision formal context consists of a universe of discourse U, a conditional attributes set A, a decision attribute set C, and binary relations I from U to A and J from U to C. Based on the discernibility matrix and Boolean function, attribute reduction of the strongly consistent decision formal context and the weakly consistent decision formal context was discussed [29]. Then a knowledge-reduction heuristic algorithm is developed to search for a minimal reduct of a consistent decision formal context and its time complexity is polynomial [11]. Recently, an order-preserving mapping is proposed to classify the decision formal contexts into consistent and inconsistent formal decision contexts. Then their corresponding knowledge reduction methods are given by constructing the suitable discernibility matrices and Boolean functions [12].

On the other hand, rules extraction is an important problem and it has been studied extensively from different perspectives [1,2,9,10,20,21,23]. In the decision formal context, a rule is derived based on the condition formal concept (generated from the condition attribute set) and the decision formal concept (generated from the decision attribute set). Each formal concept is constituted with the extension and the intension. For a rule, there often exist some redundant attributes, thus the reduction based on rules extraction is necessary. As we know, the maximal rules implicate the other rules, thus the formal concepts related to the maximal rules are important. Attribute reduction based on maximal rules is proposed in this paper. The proposed reducts preserve all the condition extensions and decision extensions related to the original maximal rules. Furthermore, it is proven

that all the maximal rules in the reduced decision formal are derived from these condition extensions and decision extensions. In other words, every original maximal rule corresponds to a maximal rule in the reduced decision formal context and they are derived based on the same condition extension and decision extension. However, the depiction of the maximal rule becomes more concise and accurate after eliminating the superfluous attributes. Using the similar idea of attribute reduction in the classical formal context [36], the corresponding discernibility matrix is presented. All attributes are classified into three types based on the maximal rules. The characteristics of these types of attributes are then analyzed.

The remainder of the paper is organized as follows: in the next section, some basic concepts related to this paper are reviewed. In Section 3, two kinds of reduction based on maximal rules are proposed and their properties are investigated. In Section 4, we study attribute reduction based on maximal rules in the decision formal context. Some judgment theorems of a reduct are developed, and the discernibility matrix is constructed to compute all the reducts via Boolean approach. In Section 5, all the attributes are classified into three types based on maximal rules. The characteristics of these attributes are also studied. Finally, in Section 6, we conclude the paper with a summary and an outlook for further research.

2. Preliminaries

To facilitate our discussion, some basic notions and results related to this paper are reviewed in this section [7,29,31].

A formal context is a triplet (U, A, I), where U is a nonempty finite set of objects, A is a nonempty finite set of attributes, and $I \subseteq U \times A$ is a binary relation from U to A. $(x, a) \in I$ means that x has the attribute a.

Definition 1. Let (U, A, I) be a formal context. For all $X \subseteq U$ and $B \subseteq A$, two operations are defined respectively

$$X^* := \{a \in A \mid \forall x \in X, (x, a) \in I\}, \quad (1)$$

$$B^* := \{x \in U \mid \forall a \in B, (x, a) \in I\}. \quad (2)$$

Evidently, X^* consists of exactly those attributes which the objects in X have in common, and B^* contains exactly those objects which share all attributes in B.

Definition 2. Let (U,A,I) be a formal context. A pair (X,B) with $X \subseteq U$ and $B \subseteq A$ is called a formal concept iff $X^* = B$ and $B^* = X$. Here X and B are called the extension and the intension of (X,B), respectively.

Theorem 1. *Let (U,A,I) be a formal context. For all $X,X_1,X_2 \subseteq U$ and $B,B_1,B_2 \subseteq A$, the following statements hold:*

(1) $X_1 \subseteq X_2 \Longrightarrow X_1^* \supseteq X_2^*$; $B_1 \subseteq B_2 \Longrightarrow B_1^* \supseteq B_2^*$.
(2) $X \subseteq X^{**}$, $B \subseteq B^{**}$.
(3) $X^* = X^{***}$, $B^* = B^{***}$.
(4) $(X_1 \cup X_2)^* = X_1^* \cap X_2^*$, $(B_1 \cup B_2)^* = B_1^* \cap B_2^*$.
(5) (X^{**}, X^*) and (B^*, B^{**}) are both concepts.

Here $(X_1, B_1) \leqslant (X_2, B_2)$ if and only if $X_1 \subseteq X_2$ (or equivalently $B_2 \subseteq B_1$). The set of all the concepts in (U,A,I) together with the partial order is denoted by $L(U,A,I)$. Furthermore, the set of all the extensions is defined as $L_U(U,A,I) = \{X | (X,B) \in L(U,A,I)\}$. It can be seen that $L_U(U,A,I) = \{X | (X,B) \in L(U,A,I)\} = \{X^{**} | X \subseteq U\}$.

Definition 3. A five-tuple (U,A,I,C,J) is called a decision formal context, if (U,A,I) and (U,C,J) are both the formal contexts, and A is called a condition attribute set, C is a decision attribute set.

Let (U,A,I,C,J) be a decision formal context, with each $B \subseteq A$, we associate a binary relation $I_B = I \cap (U \times B)$. Then (U,B,I_B) is a sub-context of (U,A,I). It is easy to see that $I_A = I, X^{*B} = X^* \cap B, X^{*B} \subseteq X^*$.

The following theorem follows immediately from the definition of $L_U(U,A,I)$.

Theorem 2. *Let (U,A,I) be a formal context. For all $B \subseteq A (B \neq \emptyset)$, $L_U(U,B,I_B) \subseteq L_U(U,A,I)$.*

Definition 4. Let $(X,K) \in L(U,A,I)$ and $(Y,H) \in L(U,C,J)$. If $X \subseteq Y(Y \neq U,\emptyset)$, then "$K \to H$" is called a rule. Furthermore, if there exist $(X',K') \in L(U,A,I)$ and $(Y',H') \in L(U,C,J)$ satisfying $K \subseteq K'$ and $H \supseteq H'$ (or equivalently $X' \subseteq X$ and $Y' \supseteq Y$), then it means that "$K \to H$" implies "$K' \to H'$"

and the rule "$K' \to H'$" is redundant. If there is no rule "$K' \to H'$" satisfying $K' \subseteq K, H \subseteq H'$, then the rule "$K \to H$" is called a maximal rule.

It should be noted that the rule proposed above is in fact a decision implication [24]. However, a decision implication may not be a rule because its premise or conclusion may not be the intension of any formal concept. Besides, the premise or the conclusion of the implication rules given in [7] may also not be the intension of any formal concept. Thus, compared with the decision implication [24] and implication rules [7], the rule proposed above has more semantic explanation than a decision implication.

3. Attribute reduction based on maximal rules

In this section, we introduce the definition of attribute reduction based on maximal rules in the decision formal context, and then present some properties related to the definition.

For any $X_i \in L_U(U,A,I)$, denote by $S_C(X_i) = \{Y_j | X_i \subseteq Y_j \in L_U(U,C,J)$ and for any $Y_j' \in L_U(U,C,J)$, if $X_i \subseteq Y_j'$, then $Y_j' \not\subseteq Y_j\}$.

Clearly, for any $Y_j \in S_C(X_i), Y_j$ is the minimal extension which contains X_i in $L_U(U,C,J)$ and denote by

$$\Phi_C = \bigcup_{X_i \in L_U(U,A,I)} S_C(X_i). \qquad (3)$$

For any $Y_j \in \Phi_C$, $T_A(Y_j) = \{X_i | Y_j \supseteq X_i \in L_U(U,A,I)$ and for any $X_i' \in L_U(U,A,I)$, if $X_i' \subseteq Y_j$, then $X_i' \not\supseteq X_i\}$ and denote by

$$\Psi_A = \bigcup_{Y_j \in \Phi_C} T_A(Y_j). \qquad (4)$$

Evidently, for any $X_i \in T_A(Y_j), X_i$ is the maximal extension which is contained in Y_j in $L_U(U,A,I)$.

From the definition of Φ_C and $T_A(Y)$, the following theorem can be obtained:

Theorem 3. *Let (U,A,I,C,J) be a decision formal context. For any $(X,K) \in L(U,A,I)$ and $(Y,H) \in L(U,C,J)$, "$K \to H$" is a maximal rule if and only if $Y \in S_C(X)(Y \neq U,\emptyset)$ and $X \in T_A(Y)$.*

Definition 5. Let (U,A,I,C,J) be a decision formal context. $D \subseteq C$ is called a decision consistent set

if for all $X_i \in L_U(U,A,I), S_D(X_i) = S_C(X_i)$. Furthermore, if for any $d \in D$, there exists $X_i \in L_U(U,A,I)$ such that $S_{D\setminus\{d\}}(X_i) \neq S_C(X_i)$, then D is called a decision reduct of (U,A,I,C,J). The intersection of all the decision reducts are called the decision core of (U,A,I,C,J).

A decision reduct D of (U,A,I,C,J) is a minimal attribute subset satisfying for all $X_i \in L_U(U,A,I)$, $S_D(X_i) = S_C(X_i)$.

Definition 6. Let (U,A,I,C,J) be a decision formal context, $E \subseteq A$ is called a condition consistent set if for all $Y_j \in \Phi_C$, $T_E(Y_j) = T_A(Y_j)$. Furthermore, if for any $e \in E$ there exists $Y_j \in \Phi_C$ such that $T_{E\setminus\{e\}}(Y_j) \neq T_A(Y_j)$, then E is called a condition reduct of (U,A,I,C,J). The intersection of all the condition reducts are called the condition core of (U,A,I,C,J).

A condition reduct E of (U,A,I,C,J) is a minimal attribute subset satisfying for all $Y_j \in \Phi_C$, $T_E(Y_j) = T_A(Y_j)$.

It follows from the Definitions 5 and 6 that there are at least one condition reduct and one decision reduct in the decision formal context.

We can see that the reducts preserve all the condition extensions and the decision extensions related to the original maximal rules. And further, the following theorem is given:

Theorem 4. *Let (U,A,I,C,J) be a decision formal context, $D \subseteq C$ is a decision consistent set and $E \subseteq A$ is a condition consistent set. Then for all $(X,K) \in L(U,E,I_E), (Y,H) \in L(U,D,J_D)$, "$K \to H$" is a maximal rule in (U,E,I_E,D,J_D) if and only if $Y \in S_C(X), X \in T_A(Y)$.*

Proof. $\Rightarrow$) "$K \to H$" is a maximal rule in (U,E,I_E,D,J_D), thus $Y \in S_D(X)$ and $X \in T_E(Y)$. From Theorem 2 and the definition of decision consistent set, $X \in L_U(U,E,I_E) \subseteq L_U(U,A,I), S_D(X) = S_C(X)$, thus $Y \in S_C(X)$. From the definition of the condition consistent set and $Y \in S_C(X)$, we have that $X \in T_E(Y) = T_A(Y)$.

$\Leftarrow$) From the definition of $S_C(X)$ and $T_A(Y)$, there is no $X' \in L_U(U,A,I)$ and $Y' \in L_U(U,C,J)$ satisfying $X \subset X' \subseteq Y$ and $X \subset Y' \subseteq Y$. By Theorem 2, $L_U(U,E,I_E) \subseteq L_U(U,A,I)$ and $L_U(U,D,J_D) \subseteq L_U(U,C,J)$ hold. Thus, there is no $X' \in L_U(U,E,I_E)$

and $Y' \in L_U(U,D,J_D)$ satisfying $X \subset X' \subseteq Y$ and $X \subset Y' \subseteq Y$. Therefore, "$K \to H$" is a maximal rule of (U,E,I_E,D,J_D). □

From Theorem 3, we know that, if $Y \in \Phi_C, X \in T_A(Y)$, then "$X^{*A} \to Y^{*C}$" is a maximal rule in (U,A,I,C,J). Thus, it can be seen that "$X^{*A} \to Y^{*C}$" and "$K \to H$" are derived based on the same condition extensions and decision extensions. However, compared with "$X^{*A} \to Y^{*C}$", the rule "$K \to H$" is more concise and accurate.

Furthermore, if $(X,K) \in L(U,E,I_E), (Y,H) \in L(U,D,J_D)$, then $K = X^{*A} \cap E, H = Y^{*C} \cap D$. We can see that the maximal rule "$K \to H$" in the reduced decision formal context (U,E,I_E,D,J_D) can be obtained based on the original maximal rule "$X^{*A} \to Y^{*C}$", the condition consistent set E and the decision consistent set D. What is more, every original maximal rule corresponds to a maximal rule in the reduced decision formal context. For example, if "$X^{*A} \to Y^{*C}$" is a maximal rule in (U,A,I,C,J), then from Theorem 3, $Y \in S_C(X), X \in T_A(Y)$. $E \subseteq A$ and $D \subseteq C$ are the condition consistent set and the decision consistent set, thus $X \in T_A(Y) = T_E(Y)$ and $Y \in S_C(X) = S_D(X)$. Therefore, "$X^{*A} \cap E \to Y^{*C} \cap D$" is a maximal rule in (U,E,I_E,D,J_D).

In what follows, we only consider the reduction method and attribute characteristics of decision attributes in detail. As to condition attributes, we only give its discernibility matrix and the other related problems can be discussed as the decision attributes.

4. Approaches to attribute reduction based on maximal rules

In this section, inspired by the approaches to attribute reduction in classical concept lattices, we construct the discernibility matrix in the decision formal context based on maximal rules, and then obtain all the reducts via Boolean approach.

Definition 7. Let (U,A,I,C,J) be a decision formal context. For all $(Y_i,B_i),(Y_j,B_j) \in L(U,C,J)$,

$$D((Y_i,B_i),(Y_j,B_j)) = \begin{cases} B_i \triangle B_j, & \text{if } Y_i \in \Phi_C, Y_i \subset Y_j \text{ or } Y_j \in \Phi_C, Y_j \subset Y_i \\ \emptyset, & \text{else} \end{cases} \quad (5)$$

is called the discriminable attribute set between (Y_i, B_i) and (Y_j, B_j). Here $B_i \triangle B_j = B_i \cup B_j - B_i \cap B_j$ and $\Lambda_{FC} = (D((Y_i, B_i), (Y_j, B_j)) | (Y_i, B_i), (Y_j, B_j) \in L(U, C, J))$ is called the discernibility matrix of the decision attributes. Denote by

$$D_0 = \{D((Y_i, B_i), (Y_j, B_j)) | D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset\} \quad (6)$$

In what follows, some judgment theorems of the decision consistent sets are proposed.

Theorem 5. *Let (U, A, I, C, J) be a decision formal context. For all $D \subseteq C$ ($D \neq \emptyset$), the following statements are equivalent.*

(1) D is a decision consistent set.

(2) $\Phi_C \subseteq L_U(U, D, J_D)$.

(3) For $(Y_i, B_i), (Y_j, B_j) \in L(U, C, J)$, if $Y_i \in \Phi_C, Y_i \subset Y_j$ (or $Y_j \in \Phi_C, Y_j \subset Y_i$), then $B_i \cap D \neq B_j \cap D$.

(4) $D \cap D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$ ($\forall D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$).

(5) For all $B \subseteq C$, if $B \cap D = \emptyset$, then $B \notin D_0$.

Proof.

$(1) \Leftrightarrow (2)$ It follows immediately from the definition of decision consistent set.

$(2) \Rightarrow (3)$ For $(Y_i, B_i), (Y_j, B_j) \in L(U, C, J)$, without loss of generality, we assume that $Y_i \in \Phi_C$ and $Y_i \subset Y_j$. Since $\Phi_C \subseteq L_U(U, D, J_D)$, then there exists $(Y_i, B_i') \in L(U, D, J_D)$ with $B_i' = B_i \cap D$. Assume that $B_i \cap D = B_j \cap D$, then $Y_j \subseteq (B_j \cap D)^* = (B_i \cap D)^* = Y_i \subset Y_j$, which contradicts to the fact that $Y_j = Y_j$. Therefore, $B_i \cap D \neq B_j \cap D$.

$(3) \Rightarrow (2)$ For $(Y, B) \in \Phi_C$, we need to prove $(Y, B \cap D) \in L(U, D, J_D)$.

On the one hand, it is easy to see that $Y^{*D} = B \cap D$.

On the other hand, assume that $(B \cap D)^* \neq Y$, then $(B \cap D)^* \supset Y$. Since $((B \cap D)^*, (B \cap D)^{**}) \in L(U, C, J)$, $(B \cap D)^{**} \cap D \neq B \cap D$ holds by (3). Since $B \cap D \subseteq B \Rightarrow (B \cap D)^* \supseteq B^* = Y \Rightarrow (B \cap D)^{**} \subseteq Y^* = B \Rightarrow (B \cap D)^{**} \cap D \subseteq B \cap D$, and $B \cap D \subseteq (B \cap D)^{**} \Rightarrow B \cap D \subseteq (B \cap D)^{**} \cap D$, we conclude that $B \cap D = (B \cap D)^{**} \cap D$, which contradicts to the fact that $B \cap D \neq (B \cap D)^{**} \cap D$. Therefore, $(B \cap D)^* = Y$.

$(3) \Rightarrow (4)$ For $(Y_i, B_i), (Y_j, B_j) \in L(U, C, J)$, if $D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$, which implies that $Y_i \in \Phi_C, Y_i \subset Y_j$ or $Y_j \in \Phi_C, Y_j \subset Y_i$, thus $B_i \cap D \neq B_j \cap D$

holds by (3). And further, $B_i \cap D - B_j \cap D = D \cap B_i \cap \overline{B_j} \neq \emptyset$, or $B_j \cap D - B_i \cap D = D \cap B_j \cap \overline{B_i} \neq \emptyset$. Therefore, we have that

$$
\begin{aligned}
D \cap D((Y_i, B_i), (Y_j, B_j)) &= D \cap (B_i \cup B_j - B_i \cap B_j) \\
&= D \cap (B_i \cup B_j) \cap (\overline{B_i} \cup \overline{B_j}) \\
&= (D \cap B_i \cap \overline{B_j}) \cup (D \cap B_i \cap \overline{B_j}) \\
&\neq \emptyset.
\end{aligned}
$$

$(4) \Rightarrow (3)$ For $(Y_i, B_i), (Y_j, B_j) \in L(U, C, J)$, if $Y_i \in \Phi_C, Y_i \subset Y_j$ or $Y_j \in \Phi_C, Y_j \subset Y_i$, then $D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$ holds by Definition 7, thus $D \cap D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$. Therefore, there exists $d \in D \cap D((Y_i, B_i), (Y_j, B_j))$, which implies $d \in B_i$ or $d \in B_j$, and $d \notin B_i \cap B_j$. If $d \in B_i$ and $d \notin B_j$, then $d \in D \cap B_i$ and $d \notin D \cap B_j$; If $d \in B_j$ and $d \notin B_i$, then $d \in D \cap B_j$ and $d \notin D \cap B_i$. Therefore, $B_i \cap D \neq B_j \cap D$.

$(4) \Leftrightarrow (5)$ It follows immediately from the definition of D_0. $\square$

From Theorem 5, we know that the decision reduct of the decision formal context (U, A, I, C, J) is a minimal subset D of C such that $D \cap D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$ ($\forall D((Y_i, B_i), (Y_j, B_j)) \neq \emptyset$).

Furthermore, the discernibility function is given based on its discernibility matrix.

$$f(D_0) = \bigwedge_{H \in D_0} (\bigvee_{h \in H} h). \quad (7)$$

By the absorption law and the distributive law, $f(D_0)$ can be translated to the minimal disjunctive normal form. From Theorem 5 and the definition of the minimal disjunctive normal form, it follows immediately that the components of conjunctive normal forms are all the decision reducts of the decision formal context (U, A, I, C, J).

The discernibility matrix of the condition attributes is given as follows and its reducts can be computed as the procedure of the decision reducts.

Definition 8. Let (U, A, I, C, J) be a decision formal context. For all $(X_i, K_i), (X_j, K_j) \in L(U, A, I)$,

$$D((X_i, K_i), (X_j, K_j)) = \begin{cases} K_i \triangle K_j, & \text{if } X_i \in \Psi_A, X_i \subset X_j \text{ or } X_j \in \Psi_A, X_j \subset X_i \\ \emptyset, & \text{else}. \end{cases} \quad (8)$$

is called the discriminable attribute set between (X_i, K_i) and (X_j, K_j). Here $K_i \triangle K_j = K_i \cup K_j - K_i \cap$

K_j and $\Lambda_{FC} = (D((X_i,K_i),(X_j,K_j))|(X_i,K_i),(X_j,K_j) \in L(U,A,I))$ is called the discernibility matrix of the condition attributes.

The following example illustrates our idea of attribute reduction based on maximal rules in further.

Example 1. A decision formal context $S = (U,A,I,C,J)$ is shown in Table 1, where the condition attributes set $A = \{a,b,c,d,e,f\}$ and the decision attributes set $C = \{g,h,k,l,m,n\}$).

Table 1. A decision formal context $S = (U,A,I,C,J)$

	a	b	c	d	e	f	g	h	k	l	m	n
1	1	1	1	0	0	1	1	1	1	1	0	1	1
2	1	0	1	1	1	1	0	0	1	0	0	0
3	0	0	0	1	0	0	0	1	0	1	0	1
4	0	1	1	0	1	1	1	1	0	0	0	1	0
5	1	0	1	1	0	1	0	0	0	0	1	1
6	0	0	0	0	0	1	0	1	0	1	0	1

In what follows, the condition concept lattice (Figure 1) and the decision concept lattice (Figure 2) are presented .
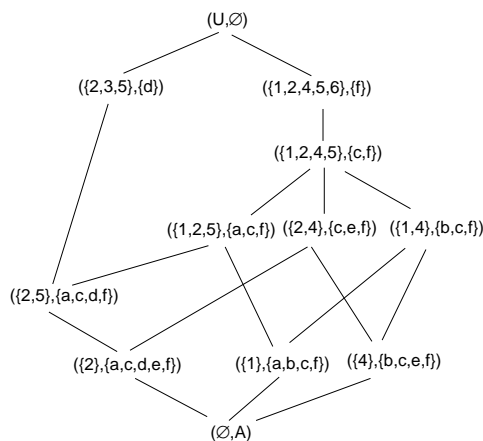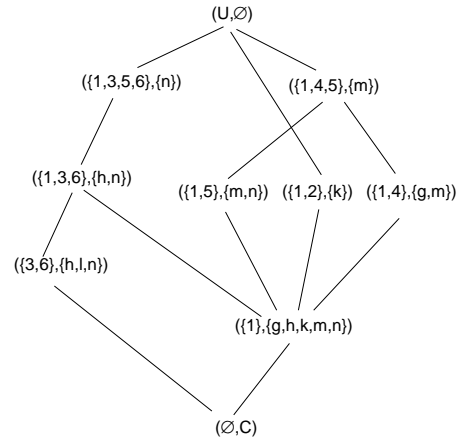

Fig.1. condition concept lattice

Fig.2. decision concept lattice

Clearly:

$L_U(U,A,I) = \{X_1 = U, X_2 = \{1,2,4,5,6\}, X_3 = \{2,3,5\}, X_4 = \{1,2,4,5\}, X_5 = \{1,2,5\}, X_6 = \{2,5\}, X_7 = \{2,4\}, X_8 = \{1,4\}, X_9 = \{4\}, X_{10} = \{2\}, X_{11} = \{1\}, X_{12} = \emptyset\}$;

$L_U(U,C,J) = \{Y_1 = U, Y_2 = \{1,3,5,6\}, Y_3 = \{1,4,5\}, Y_4 = \{1,3,6\}, Y_5 = \{1,5\}, Y_6 = \{1,2\}, Y_7 = \{1,4\}, Y_8 = \{3,6\}, Y_9 = \{1\}, Y_{10} = \emptyset\}$.

Thus $S_C(X_1) = \{Y_1\}, S_C(X_2) = \{Y_1\}, S_C(X_3) = \{Y_1\}, S_C(X_4) = \{Y_1\}, S_C(X_5) = \{Y_1\}, S_C(X_6) = \{Y_1\}, S_C(X_7) = \{Y_1\}, S_C(X_8) = \{Y_7\}, S_C(X_9) = \{Y_7\}, S_C(X_{10}) = \{Y_6\}, S_C(X_{11}) = \{Y_9\}, S_C(X_{12}) = \{Y_{10}\}$.

Consequently, $\Phi_C = \{Y_1,Y_6,Y_7,Y_9,Y_{10}\}$.

Furthermore, $T_A(Y_1) = \{X_1\}, T_A(Y_6) = \{X_{10}\}, T_A(Y_7) = \{X_8\}, T_A(Y_9) = \{X_{11}\}, T_A(Y_{10}) = \{X_{12}\}$.

Therefore, $\Psi_A = \{X_1,X_8,X_{10},X_{11},X_{12}\}$.

Then the maximal rules in the decision formal context $S = (U,A,I,C,J)$ can be given as follows:

(1) if $b \wedge c \wedge f$, then $g \wedge m$;

(2) if $a \wedge c \wedge d \wedge e \wedge f$, then k;

(3) if $a \wedge b \wedge c \wedge f$, then $g \wedge h \wedge k \wedge m \wedge n$.

Based on Ψ_A and Φ_C, the discernibility matrixs of the condition attributes and the decision attributes are presented as Tables 2 and 3, respectively.

Then the discernibility function of the condition attributes is

Table 2: Discernibility matrix of condition attributes

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
X_1	∅											
X_2	∅	∅										
X_3	∅	∅	∅									
X_4	∅	∅	∅	∅								
X_5	∅	∅	∅	∅	∅							
X_6	∅	∅	∅	∅	∅	∅						
X_7	∅	∅	∅	∅	∅	∅	∅					
X_8	{bcf}	{bc}	∅	{b}	∅	∅	∅	∅				
X_9	∅	∅	∅	∅	∅	∅	∅	∅	∅			
X_{10}	{acdef}	{acde}	{acef}	{ade}	{de}	{e}	{ad}	∅	∅	∅		
X_{11}	{abcf}	{abc}	∅	{ab}	{b}	∅	∅	{a}	∅	∅	∅	
X_{12}	A	{abcde}	{abcef}	{abde}	{bde}	{be}	{abd}	{ade}	{ad}	{b}	{de}	∅

Table 3: Discernibility matrix of decision attributes

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}
Y_1	∅									
Y_2	∅	∅								
Y_3	∅	∅	∅							
Y_4	∅	∅	∅	∅						
Y_5	∅	∅	∅	∅	∅					
Y_6	{k}	∅	∅	∅	∅	∅				
Y_7	{gm}	∅	{g}	∅	∅	∅	∅			
Y_8	∅	∅	∅	∅	∅	∅	∅	∅		
Y_9	{ghkm}	{ghkm}	{ghkn}	{gkm}	{ghk}	{ghmn}	{hkn}	∅	∅	
Y_{10}	C	{ghklm}	{ghkln}	{gklm}	{ghkl}	{ghlmn}	{hkln}	{gkm}	{l}	∅

$$\triangle_1 = (b \vee c \vee f) \wedge (b \vee c) \wedge (b) \wedge (a \vee c \vee d \vee f) \vee \wedge (a \vee c \vee d \vee e) \wedge (a \vee c \vee e \vee f) \wedge (a \vee d \vee e) \wedge (d \vee e) \wedge (e) \wedge (a \vee d) \wedge (a \vee b \vee c \vee f) \wedge (a \vee b \vee c) \wedge (a \vee b) \wedge (a) \wedge (a \vee b \vee c \vee d \vee e \vee f) \wedge (a \vee b \vee c \vee d \vee e) \wedge (a \vee b \vee d \vee e) \wedge (b \vee d \vee e) \wedge (b \vee e) \wedge (a \vee b \vee d) \wedge (a \vee b \vee c \vee e \vee f)$$
$$= a \wedge b \wedge e.$$

Therefore, $E = \{a,b,e\}$ is the unique condition reduct of the decision formal context (U,A,I,C,J).

The discernibility function of the decision at-

tributes is
$$\triangle_2 = (k) \wedge (g \vee m) \wedge (g) \wedge (g \vee h \vee k \vee m \vee n) \wedge (g \vee h \vee k \vee m) \wedge (g \vee h \vee k \vee n) \wedge (g \vee k \vee m) \wedge (g \vee h \vee k) \wedge (g \vee h \vee m \vee n) \wedge (h \vee k \vee n) \wedge (g \vee h \vee k \vee l \vee m \vee n) \wedge (g \vee h \vee k \vee l \vee m) \wedge (g \vee h \vee k \vee l \vee n) \wedge (g \vee k \vee l \vee m) \wedge (g \vee h \vee k \vee l) \wedge (g \vee h \vee l \vee m \vee n) \wedge (h \vee k \vee l \vee n) \wedge (l)$$
$$= g \wedge k \wedge l.$$

Therefore, $D = \{g,k,l\}$ is the unique decision reduct of the decision formal context (U,A,I,C,J).

It is easy to see that $\Phi_C \subseteq L(U,D,J_D)$ and $\Psi_A \subseteq L(U,E,I_E)$ in (U,E,I_E,D,J_D).

The maximal rules in the reduced decision formal context (U,E,I_E,D,J_D) can be obtained based on the original maximal rules, the condition reduct and the decision reduct. They are given as follows:

($1'$) if b, then g;

($2^{'}$) if $a \wedge e$, then k;

($3^{'}$) if $a \wedge b$, then $g \wedge k$.

These maximal rules are derived from the condition extensions and the decision extensions related to the original maximal rules. However, it can be seen that they are more concise and accurate.

5. Attribute characteristics based on maximal rules

Based on maximal rules, the significance of the decision attributes in a given decision formal context may be different and they play different roles in the decision reduct. These decision attributes can be classified into three types: the absolutely necessary decision attributes, the relatively necessary decision attributes, and the absolutely unnecessary decision attributes. In this section, we provide the characteristics of these types of attributes.

Definition 9. Let $S = (U, A, I, C, J)$ be a decision formal context and $D_i \subseteq C$, with $i \in \Upsilon$, are the decision reducts of S, then the attributes in C are classified into three types:

(1) Absolutely necessary decision attribute set: $L = \bigcap_{i \in \Upsilon} D_i$.

(2) Relatively necessary decision attribute set: $M = \bigcup_{i \in \Upsilon} D_i - \bigcap_{i \in \Upsilon} D_i$.

(3) Absolutely unnecessary decision attribute set: $N = C - \bigcup_{i \in \Upsilon} D_i$.

It should be noted that the method of classifying the attributes has been proposed in [36] and it is generalized to classify the attributes based on the maximal rules in this paper.

Denote by $\Phi_C = \{B | (Y, B) \in L(U, C, J) \text{ and } Y \in \Phi_U\}$, we get the following theorem.

Theorem 6. *Let $S = (U, A, I, C, J)$ be a decision formal context. $D \subseteq C$ is a decision consistent set of S if and only if for all $B \in \Phi_C$, $B^* = (B^{**} \cap D)^*$ holds.*

Proof. $\Rightarrow$) If D is a decision consistent set, then $\Phi_U \subseteq L_U(U, D, J_D)$. Since for all $B \in \Phi_C$ we have $B^* \in \Phi_U$, then $B^* \in L_U(U, D, J_D)$, which means that $(B^*, B^{**} \cap D) \in L(U, D, J_D)$, thus, $B^* = (B^{**} \cap D)^*$.

$\Leftarrow$) For all $Y \in \Phi_U$, we have $B \in \Phi_C$ such that $B^* = Y$. On the other hand, $B^* = (B^{**} \cap D)^* \in$

$L_U(U, D, J_D)$, thus $\Phi_U \subseteq L_U(U, D, J_D)$. Therefore, D is a decision consistent set. $\square$

Theorem 7. *Let $S = (U, A, I, C, J)$ be a decision formal context, $d \in C$. Then*

(1) d is an absolutely necessary decision attribute iff there exists $B \in \Phi_C$ satisfying $(B^{**} - \{d\})^* \neq B^*$.

(2) d is a relatively necessary decision attribute iff for all $B \in \Phi_C$, satisfying $(B^{**} - \{d\})^* = B^*$ and there exists D_{i_0}, with $i_0 \in \Upsilon$, and $K \in \Phi_C$ satisfying $(K^{**} \cap (D_{i_0} - \{d\}))^* \neq K^*$.

(3) d is an absolutely unnecessary decision attribute iff for all $B \in \Phi_C$, satisfying $(B^{**} - \{d\})^* = B^*$ and for all D_i, with $i \in \Upsilon$, satisfying $(B^{**} \cap (D_i - \{d\}))^* = B^*$.

Proof. (1) $d \in C$ is an absolutely necessary decision attribute of S if and only if $C - \{d\}$ is not consistent set. It is equivalent to the fact that there exists $B \in \Phi_C$ satisfying $(B^{**} - \{d\})^* \neq B^*$ by Theorem 6.

(2) $d \in C$ is a relatively necessary decision attribute of S if and only if d is not a core attribute and there exists a reduct D_{i_0}, with $i_0 \in \Upsilon$, satisfying $d \in D_{i_0}$. Therefore the conclusion can be obtained from (1) and Theorem 6.

(3) It follows immediately from (1) and (2). $\square$

Example 2. In the decision formal context $S = (U, A, I, C, J)$ shown in Table 1, by Theorem 7, we can verify that $\{a, b, e\}$ are the absolutely necessary condition attributes set and $\{c, d, f\}$ are the absolutely unnecessary condition attributes set; $\{g, k, l\}$ are the absolutely necessary decision attributes set, and $\{h, m, n\}$ are the absolutely unnecessary decision attributes set.

6. Conclusions

Attribute reduction in the decision formal context is one of the important problems in knowledge discovery. In this paper, we propose the condition reduct and the decision reduct based on maximal rules. The reducts preserve all the condition extensions and decision extensions related to the original maximal rules. After eliminating the superfluous attributes, the depiction of the maximal rule becomes

more concise and accurate. It is proven that the maximal rules in the reduced decision formal context can be obtained based on the original maximal rules, the condition consistent set and the decision consistent set. The related discernibility matrix is constructed to compute all the reducts via Boolean approach. Based on the maximal rules, all the attributes are classified into three types. The characteristics of these types of attributes are also analyzed. Our further work is to study attribute reduction method based maximal rules in the fuzzy decision formal concept lattices and explore its mathematical property.

Acknowledgments

References

1. R. Agrawal, T. Imielinski and A. Swami, "Mining association rules between sets of items in large databases," *Proceedings of the ACM SIGMOD Conference on Management of data, Washington, DC, USA: ACM Press*, 207–216 (1993). `doi:10.1145/170035.170072`

2. P. Y. Bastide, R. Taouil and L. Lakhal, "Discovering frequent closed item sets for association rules," *Proceedings of 7th International Conference Database Theory (ICDT99), Jerusalem, Israel: Springer–Verlag*, 398–416 (1999). `doi:10.1007/3-540-49257-7_25`

3. R. Belohlavek, B. D. Baets, J. Outrata and V. Vychodil, "Inducing decision trees via concept lattices," *International Journal of General Systems*, 38, 1563–5104 (2009). `doi:10.1080/03081070902857563`

4. A. Berry and A. Sigayret, "Representing a concept lattice by a graph," *Discrete Applied Mathematics*, 144, 27–42 (2004). `doi:10.1016/j.dam.2004.02.016`

5. C. Carpineto and G. Romano, "A lattice conceptual clustering system and its application to browsing retrieval," *Maching Learning*, 10, 95–122 (1996). `doi:10.1007/BF00058654`

6. R. Cole, P. Eklund and G. Stumme, "Document retrieval for e–mail search and discovery using formal concept analysis," *Applied Artificial Intelligence*, 17, 257–280 (2003). `doi:10.1080/713827120`

7. B. Ganter and R. Wille, "*Formal concept analysis: Mathematical Foundations*," Springer, Berlin, 1999.

8. L. K. Guo, F. P. Huang, Q. G. Li and G. Q. Zhang, "Power contexts and their concept lattices," *Discrete Mathematics*, 311, 2049–2063 (2011). `doi:10.1016/j.disc.2011.04.033`

9. S. K. Harms and J. S. Deogun, "Sequential accociation rule with time lags," *Journal of Intelligent Information Systems*, 22, 7–22 (2004). `doi:10.1023/A:1025824629047`

10. Y. Leung and D. Y. Li, "Maximal consistent block technique for rule acquisition in incomplete information systems," *Information Sciences*, 153, 85–106 (2003). `doi:10.1016/S0020-0255(03)00061-6`

11. J. H. Li, C. L. Mei and Y. J. Lv, "A heuristic knowledge-reduction method for decision formal contexts," *Computers and Mathematics with Applications*, 61, 1096–1106 (2011). `doi:10.1016/j.camwa.2010.12.060`

12. J. H. Li, C. L. Mei and Y. J. Lv, "Knowledge reduction in formal decision contexts based on an order-preserving mapping," *International Journal of General Systems*, 41, 143–161 (2012). `doi:10.1080/03081079.2011.634410.`

13. M. Z. Li and J. S. Mi, "The strong direct product of formal contexts," *Information Sciences*, 226, 47–67 (2013). `doi:10.1016/j.ins.2012.10.032`

14. M. Liu, M. W. Shao, W. X. Zhang and C. Wu, "Reduction method for concept lattices based on rough set theory and its application," *Computers and Mathematics with Applications*, 53, 1390–1410 (2007). `doi:10.1016/j.camwa.2006.03.040`

15. E. Marczewski, "A general scheme of independence in mathematics," *Bulletin de L Academie of Polonaise des Sciences–Serie des Sciences Mathematiques Astronomiques et Physiques*, 6, 731–736 (1958).

16. J. Medina, "Relating attribute reduction in formal, object-oriented and property-oriented concept lattices," *Computers and Mathematics with Applications* 64, 1992–2002 (2012). `doi:10.1016/j.camwa.2012.03.087`

17. J. Medina, M. Ojeda–Aciego and J. Ruiz–Calviño, "Relating generalized concept lattices and concept lattices for non–commutative conjunctors," *Applied Mathematics Letters*, 21, 1296–1300 (2008). `doi:10.1016/j.aml.2007.12.026`

18. J. Medina, M. Ojeda-Aciego and J. Ruiz-Calvi, "For-

mal concept analysis via multi-adjoint concept lattices," *Fuzzy Sets and Systems*, 160, 130–144 (2009). `doi:10.1016/j.fss.2008.05.004`

19. J. S. Mi, Y. Leung and W. Z. Wu, "Approaches to attribute reduction in concept lattices induced by axialities," *Knowledge–Based Systems*, 23, 504–511 (2010). `doi:10.1016/j.knosys.2010.03.007`

20. T. T. Nguyen, S. C. Hui and K. Y. Chang, "A lattice–based approach for mathematical search using Formal Concept Analysis," *Expert Systems with Applications*, 39, 5820–5828 (2012). `doi:10.1016/j.eswa.2011.11.085`

21. T. Oyama, K. Kitano and T. K. Satou, "Extraction of knowledge on protein–protein interaction by association rule discovery," *Bioinformatics*, 18, 705–714 (2002). `doi:10.1093/bioinformatics/18.5.705`

22. Z. Pawlak, "Rough sets," *International Journal of Computer and Information Science*, 11, 341–356 (1982). `doi:10.1007/BF01001956`

23. J. R. Quinlan, "Simplifying decision trees," *International Journal of Man–Machine Studies*, 27, 221–234 (1987). `doi:10.1016/S0020-7373(87)80053-6`

24. K. S. Qu, Y. H. Zhai, J. Y. Liang and M. Chen, "Study of decision implications based on formal concept analysis," *International Journal of General Systems*, 36, 147–156 (2007). `doi:10.1080/03081070600913650`

25. Y. C. Shao and K. Y. Qin, "Fuzzy Soft Sets and Fuzzy Soft Lattices," *International Journal of Computational Intelligence Systems*, 5, 1135–1147 (2012). `doi:10.1080/18756891.2012.747667`

26. P. Tonella, "Using a concept lattice of decomposition slices for program understanding and impact analysis," *IEEE Transactions on Software Engineering*, 29, 495–509 (2003). `doi:10.1109/TSE.2003.1205178`

27. X. Wang and W. X. Zhang, "Relations of attribute reduction between object and property oriented concept lattices," *Knowledge Based Systems*, 21, 398–403 (2008). `doi:10.1016/j.knosys.2008.02.005`

28. H. Wang and W. X. Zhang, "Approaches to knowledge reduction in generalized consistent decision formal context," *Mathematical and Computer Modelling*, 48, 1677–1684 (2008). `doi:10.1016/j.mcm.2008.06.007`

29. L. Wei, J. J. Qi and W. X. Zhang, "Attribute reduction theory of concept lattice based on decision formal contexts," *Science in China Series F: Information Sciences*, 51, 910–923 (2008). `doi:10.1007/s11432-008-0067-4`

30. R. Wille, "Restructuring lattice theory: an approach based on hierarchies of concepts," *Formal Concept Analysis*, 314–339 (2009). `doi:10.1007/978-3-642-01815-2_23`

31. K. E. Wolff, "A conceptual view of knowledge bases in rough set theory," *Lecture Notes in Computer Science*, 220–228 (2005). `doi:10.1007/3-540-45554-X_26`

32. W. Z. Wu, Y. Leung and J. S. Mi, "Granular computing and knowledge reduction in formal context," *IEEE Transactions on Knowledge and Data Engineering*, 21, 1461–1474 (2009). `doi:10.1109/TKDE.2008.223`

33. Y. Y. Yao, "A comparative study of formal concept analysis and rough set theory in data analysis," *In: S. Tsumoto, R. Slowinski, H. Jan Komorowski, J.W. Grzymala–Busse (Eds.) Springer, Berlin*, 59–68 (2004). `doi:10.1007/978-3-540-25929-9_6`

34. H. Z. Yang, Y. Leung, and M. W. Shao, "Rule acquisition and attribute reduction in real decision formal contexts," *Soft Computing*, 15, 1115–1128 (2011). `doi:10.1007/s00500-010-0578-y`

35. L. Yang, Y. H. Wang and Y. Xu, "A Combination Algorithm of Multiple Lattice-Valued Concept Lattices," *International Journal of Computational Intelligence Systems*, 6, 881–892 (2013). `doi:10.1080/18756891.2013.807109`

36. W. X. Zhang, L. Wei and J. J. Qi, "Attribute reduction in concept lattice based on discernibility matrix," *Lecture Notes in Computer Science*, 3642, 157–165 (2005). `doi:10.1007/11548706_17`