Knowledge-Based Systems 23 (2010) 427–433

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Approximation reduction in inconsistent incomplete decision tables

Yuhua Qian, Jiye Liang *, Deyu Li, Feng Wang, Nannan Ma

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

ARTICLE INFO

Article history:
Received 30 June 2009
Received in revised form 8 January 2010
Accepted 16 February 2010
Available online 24 February 2010

Keywords:
Rough set theory
Inconsistent incomplete decision table
Maximal consistent block
Discernibility function
Approximation reduction

ABSTRACT

This article deals with approaches to attribute reductions in inconsistent incomplete decision table. The main objective of this study is to extend a kind of attribute reductions called a lower approximation reduct and an upper approximation reduct, which preserve the lower/upper approximation distribution of a target decision. Several judgement theorems of a lower/upper approximation consistent set in inconsistent incomplete decision table are educed. Then, the discernibility matrices associated with the two approximation reductions are examined as well, from which we can obtain approaches to attribute reduction of an incomplete decision table in rough set theory.

1. Introduction

Since the original exposition of the rough set theory (RST) by Pawlak [20–22] as a method of set approximation, it has continued to flourish as a tool for data mining and data analysis [2,3,24,25,31,34]. One fundamental aspect of RST involves a search for particular subsets of condition attributes that provide the same information for classification purposes as the full set of available attributes. Such subsets are called attribute reductions [5,6].

Attribute reduction is performed in information systems (ISs) by means of the notion of a reduct based on a specialization of the general notion of independence because of Marczewski [15]. Many types of knowledge reductions have been proposed in the area of rough sets [1–4,9,10,14,16–19,26–30,32,33,35], each of the reductions aimed at some basic requirements. It is required to provide their consistent classification. In the real world, most decision information systems (also called decision tables) are inconsistent because of various factors such as noise in data, compact representation, prediction capability, etc. To acquire brief decision rules from an inconsistent decision table, knowledge reductions are needed. The inconsistence of a system makes it infeasible to induce a set of certain definite rules covering all system objects with confidence 1. Hence, one has to resort to obtain rules from an inconsistent decision information system with confidence less than 1 (usually called possible rules or probabilistic rules). At present, rough set theory offers an unique approach for generating such possible rules without a priori knowledge.

In recent years, more attention has been paid to knowledge reduction in inconsistent systems in rough set research [9,10,14,16,17,26,29,33,36]. β-reduct was studied in the variable precision rough set model proposed by Ziarko [32]. The notions of α-reduct and α-relative reduct for decision tables were defined. The α-reduct allows the occurrence of additional inconsistency that is controlled by means of α-parameter [19]. In [29], Slezak presented an attribute reduction that preserves the class membership distribution for all objects in the ISs. It was shown by Slezak that the attribute reduction preserving the membership distribution is equivalent to the attribute reduction preserving the value of generalized inference measure function [29]. A generalized attribute reduction also was introduced in [28] that allows the value of generalized inference measure function after the attribute reduction to be different from the original one by user-specified threshold. Five kinds of attribute reductions and the relationships among them in inconsistent systems were investigated by Kryszkiewicz [9], Li [10] and Mi [16]. By eliminating the rigorous conditions required by distribution reduct, maximum distribution reduct was introduced by Zhang et al. in [33]. Unlike possible reduct [18], maximum distribution reduct can derive decision rules that are compatible with the original systems.

According to whether or not there are missing data (null values), information systems can be classified into two categories: complete and incomplete. By an incomplete information system we mean a system with missing data (null values). We do not

* Corresponding author. Address: Key Laboratory of Computational Intelligence and Chinese, Information Processing of Ministry of Education, Taiyuan 030006, China.

E-mail addresses: jinchengqyh@126.com (Y. Qian), ljy@sxu.edu.cn (J. Liang), ljdy@sxu.edu.cn (D. Li), sxuwangfeng@126.com (F. Wang), sxmanannan@126.com (N. Ma).

0950-7051/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2010.02.004

consider the case of null value meaning inapplicable value. This problem may be solved by adding a special symbol denoting inapplicable value to the attribute domains. In the paper we deal with the problem of unknown values. In other words, a null value may be some value in the domain of the corresponding attribute [7,11–14,23,28]. For an incomplete information system, if we distinguish condition attributes and decision attributes, then we call it an incomplete decision table. In the context of incomplete decision table, Leung and Li [11] proposed a so-called maximal consistent block technique, which had been used to rule acquisition from a consistent incomplete decision table. The maximal consistent block technique can describe the minimal units for information in incomplete information systems. In this paper, we concern on how to acquire attribute reductions in inconsistent incomplete decision tables. In recent years, some related researches have presented on attribute reductions in inconsistent incomplete decision tables. Zhou and Huang [37] developed several reduction methods in incomplete inconsistent decision tables, which are distribution reduction, maximum distribution reduction and assignment reduction. Miao et al. [18] concluded three kinds of relative reducts, such as region preservation reduct, decision preservation reduct and relationship preservation reduct in the context of decision tables. Unlike these existing works, in this study, we will employ the consistent block technique for attribute reduction in inconsistent incomplete information systems.

The main objective of this article is to introduce two concepts of approximation reduction named as a lower approximation reduct and an upper approximation reduct in inconsistent incomplete decision tables, which preserves the lower approximation and upper approximation of the decision classification in the context of maximal consistent blocks. The rest is organized as follows. Some preliminary concepts such as incomplete decision tables, inconsistency and maximal consistent blocks are briefly reviewed in Section 2. In Section 3, the notions of a lower approximation reduct and an upper approximation reduct are introduced to an inconsistent incomplete decision table in the context of maximal consistent blocks, and their some important properties are also obtained. In Section 4, the approaches to the two approximation reducts are provided and an illustrate example is employed to examine their validity. We then conclude the paper with a summary in Section 5.

2. Preliminaries

In this section, we briefly review several basic concepts, which are incomplete decision tables, inconsistency and maximal consistent blocks.

An information system is a pair $S=(U,A)$, where,

(1) U is a non-empty finite set of objects;
(2) A is a non-empty finite set of attributes;
(3) for every $a\in A$, there is a mapping $a,\ a: U\rightarrow V_a$, where V_a is called the value set of a.

Each subset of attributes $P\subseteq A$ determines a binary indistinguishable relation $IND(P)$ as follows:

$$IND(P)=\{(u,v)\in U\times U|\forall a\in P, a(u)=a(v)\}.$$

It can be easily shown that $IND(P)$is an equivalence relation on the set U. For $P\subseteq A$, the relation $IND(P)$ constitutes a partition of U, which is denoted by $U/IND(P)$, just U/P.

It may happen that some of the attribute values for an object are missing. For example, in medical information systems there may exist a group of patients for which it is impossible to perform all the required tests. These missing values can be represented by the set of all possible values for the attribute or equivalence by the domain of the attribute. To indicate such a situation, a distinguished value, a so-called null value is usually assigned to those attributes.

If V_a contains a null value for at least one attribute $a\in A$, then S is called an incomplete information system, otherwise it is complete [7,8,11–14]. Further on, we will denote the null value by *. Thus, if the value of an attribute a is missing, then the real value must be from the set $V_a-\{*\}$. Any domain value different from "*" will be called regular.

Let $S=(U,A)$ be an information system, $P\subseteq A$ an attribute set. We define a binary relation on U as follows:

$$SIM(P)=\{(u,v)\in U\times U|\forall a\in P, a(u)=a(v) \text{ or } a(u)=* \text{ or } a(v)=*\}.$$

In fact, $SIM(P)$ is a tolerance relation on U, the concept of a tolerance relation has a wide variety of applications in classification [11–14,23]. It can be easily shown that $SIM(P)=\bigcap_{a\in P}SIM(\{a\})$.

Let $S_P(u)$ denote the set $\{v\in U|(u,v)\in SIM(P)\}$. $S_P(u)$ is the maximal set of objects which are possibly indistinguishable by P with u. Let $U/SIM(P)$ denote the family sets $\{S_P(u)|u\in U\}$, the classification or the knowledge induced by P. A member $S_P(u)$ from $U/SIM(P)$ will be called a tolerance class or a granule of information. It should be noticed that the tolerance classes in $U/SIM(P)$ do not constitute a partition of U in general. They constitute a cover of U, i.e., $S_P(u)\neq O$ for every $u\in U$, and $\bigcup_{u\in U}S_P(u)=U$.

An incomplete information system $S=(U,C\cup D)$ is called an incomplete decision table if condition attributes and decision attributes are distinguished, where C is the condition attribute set, and D is the decision attribute set. For an incomplete decision table $S=(U,C\cup D)$, if $SIM(C)\subseteq IND(D)$, then we say the decision table S is consistent, otherwise we say it is inconsistent.

Example 1. Consider descriptions of several cars in Table 1 [8].

This is an incomplete decision table, where $U=\{u_1,u_2,u_3,u_4,u_5,u_6\}$, $C=\{a_1,a_2,a_3,a_4\}$ with a_1-Price, a_2-Mileage, a_3-Size, a_4-Max-Speed, and $D=\{d\}$. By computing, it follows that:

$$U/SIM(C)=\{S_C(u_1),S_C(u_2),S_C(u_3),S_C(u_4),S_C(u_5),S_C(u_6)\},$$

where $S_C(u_1)=\{u_1\},\ S_C(u_2)=\{u_2,u_6\},\ S_C(u_3)=\{u_3\},\ S_C(u_4)=\{u_4,u_5\},\ S_C(u_5)=\{u_4,u_5,u_6\},\ S_C(u_6)=\{u_2,u_5,u_6\}$.

Note that the order pair $(u_5,u_4)\in SIM(C)$ do not belong to the indistinguishable relation $IND(D)$. Hence, it is obvious that S is an inconsistent incomplete decision table. It is trivial to observe that the value of the generalized decision ∂_d for an object in an incomplete decision table is a superset of its generalized decision's value (see ∂_d in Table 1).

However, tolerance classes are not the minimal units for describing knowledge or information in an incomplete information system or an incomplete decision table [2,11].

Let $S=(U,A)$ be an information system, $P\subseteq A$ an attribute set and $X\subseteq U$ a subset of objects. We say X is consistent with respect to P if $(u,v)\in SIM(P)$ for any $u,\ v\in X$. If there does not exist a subset $Y\subseteq U$ such that $X\subset Y$, and Y is consistent with respect to P, then X is called a maximal consistent block of P. Obviously, in a

Table 1
The incomplete decision table about car [16].

Car	Price	Mileage	Size	Max-Speed	d	∂_d
u_1	High	Low	Full	Low	Good	{Good}
u_2	Low	*	Full	Low	Good	{Good}
u_3	*	*	Compact	Low	poor	{Poor}
u_4	High	*	Full	High	Good	{Good, excellent}
u_5	*	*	Full	High	Excellent	{Good, excellent}
u_6	Low	High	Full	*	Good	{Good, excellent}

maximal consistent block, all objects are not indiscernible with available information provided by a similarity relation [11].

Henceforth, we denote the set of all maximal consistent blocks determined by $P \subseteq A$ as MC_P, and the set of all maximal consistent blocks of P which includes some object $u \in U$ is denoted as $MC_P(u)$. It is obvious that $X \in MC_P$ if and only if $X = \bigcap_{u \in X} S_P(u)$ [11]. In fact, the set of all maximal consistent blocks MC_P will degenerate into the partition U/P induced by attribute set P in a complete information system, i.e., $MC_P = U/P$.

Example 2. Computing all maximal consistent blocks of C in Table 1.

By computing, from Example 1, we have that

$$MC_C = \{\{u_1\}, \{u_2, u_6\}, \{u_3\}, \{u_4, u_5\}, \{u_5, u_6\}\},$$

where MC_C is the set of all maximal consistent blocks determined by C on U.

3. Approximation reduction in inconsistent incomplete decision tables

Attribute reduction is needed to simplify a decision table. The approximation reduction proposed by Mi et al. is an important kind of attribute reduction, which can be used to simplify an inconsistent complete decision table [17,34]. To date, however, there is not any practical approach to attribute reduction in inconsistent incomplete decision tables. In this section, we present the notions of a lower approximation reduct and an upper approximation reduct in an inconsistent incomplete decision table and then deduce their some important properties.

In order to introduce the notion of approximation reduction, firstly, we redefine the consistency of an incomplete decision table by using maximal consistent blocks.

Definition 1. Let $S = (U, C \cup D)$ be an incomplete decision table. If for any maximal consistent block $X \in MC_C$, there exists a decision class $Y \in U/IND(D)$ such that $X \subseteq Y$, then the decision table S is said to be consistent, otherwise it is said to be inconsistent.

From the consistency of incomplete decision tables and Definition 1, one can prove the following Theorem 1.

Theorem 1. *Let $S = (U, C \cup D)$ be an incomplete decision table. If $SIM(C) \subseteq IND(D)$, then, for any maximal consistent block $X \in MC_C$, there exists a decision class $Y \in U/IND(D)$ such that $X \subseteq Y$.*

Proof. From the definition of maximal consistent blocks, we know that X is consistent with respect to C if $(u, v) \in SIM(C)$ for any $u, v \in X$. If $SIM(C) \subseteq IND(D)$, then we have that $(u, v) \in IND(D)$ for any $(u, v) \in SIM(C)$. In other words, there exists an equivalence class $[u]_D$ or $[v]_D$ in $U/IND(D)$ such that $u, v \in [u]_D$ and $u, v \in [v]_D$. Hence, from $X = \bigcap_{u \in X} S_C(u)$ [11], one can obtain that $u, v \in [u]_D$ for any $(u, v) \in X \times X$, i.e., $X \subseteq [u]_D$. Therefore, for any maximal consistent block $X \in MC_C$, there exists a decision class $Y \in U/IND(D)$ such that $X \subseteq Y$. This completes the proof. □

Let $S = (U, A)$ be an incomplete information system, $B \subseteq A$ and $X \subseteq U$. In [11], the approximation operators $\underline{apr}_B$ and $\overline{apr}_B$ are defined by

$$\underline{apr}_B(X) = \bigcup \{Y \in MC_B | Y \subseteq X\},$$
$$\overline{apr}_B(X) = \bigcup \{Y \in MC_B | Y \cap X \neq \emptyset\}.$$

Let $S = (U, C \cup D)$ be an incomplete decision table and $B \subseteq C$. Denoted by $U/IND(D) = \{D_1, D_2, \ldots, D_r\}$. The lower and upper approximation distribution functions with respect to B are defined as

$$\underline{apr}_B(D) = (\underline{apr}_B(D_1), \underline{apr}_B(D_2), \ldots, \underline{apr}_B(D_r)),$$
$$\overline{apr}_B(D) = (\overline{apr}_B(D_1), \overline{apr}_B(D_2), \ldots, \overline{apr}_B(D_r)).$$

By the lower and upper approximation distribution functions, we then introduce the notions of a lower/upper approximation consistent set and a lower/upper approximation reduct in incomplete decision tables.

Definition 2. Let $S = (U, C \cup D)$ be an incomplete decision table and $B \subseteq C$.

(1) If $\underline{apr}_B(D) = \underline{apr}_C(D)$, we say that B is a lower approximation consistent set of A. If B is a lower approximation consistent set, and no proper subset of B is lower approximation consistent, then B is called a lower approximation reduct of C.
(2) If $\overline{apr}_B(D) = \overline{apr}_C(D)$, we say that B is an upper approximation consistent set of A. If B is a upper approximation consistent set, and no proper subset of B is upper approximation consistent, then B is called an upper approximation reduct of C.

From the definition of the set approximation in the context of maximal consistent blocks, it follows that the lower approximation of a set must be contained by its upper approximation. Hence, it is easy to prove that an upper approximation consistent set must be a lower approximation consistent set. However, the converse relation is not true for inconsistent incomplete decision tables. In fact, for a consistent incomplete decision table, we can obtain the following theorem.

Theorem 2. *Let $S = (U, C \cup D)$ be a consistent incomplete decision table and $B \subseteq C$. Then, B is a lower approximation consistent set iff B is an upper approximation consistent set.*

Proof. It is straightforward. □

Let $S = (U, C \cup D)$ be an incomplete decision table, $B \subseteq C, U/IND(D) = \{D_1, D_2, \ldots, D_r\}$ and $X \in MC_B$. For convenient representation, denoted by

$$\varphi_B(u) = \begin{cases} D_j, & \text{if } \exists j \leqslant r \quad \text{such that } u \in X, X \subseteq \underline{apr}_B(D_j), \\ \emptyset, & \text{otherwise.} \end{cases}$$
$$\eta_B(u) = \{D_j : u \in X, X \subseteq \overline{apr}_B(D_j)\}.$$

Theorem 3. *φ and η have the following properties:*

(1) $\underline{apr}_B(D_j) = \bigcup\{u : \varphi_B(u) = \{D_j\}, u \in U\}$.
(2) $\overline{apr}_B(D_j) = \bigcup\{u : D_j \in \eta_B(u), u \in U\}$.

Proof. It is straightforward. □

In the following, we investigate some judgement methods of a lower/upper approximation consistent set in an inconsistent incomplete decision table.

Here, we define a partial relation in incomplete information systems. Let $S = (U, A)$ be an incomplete information system, $P, Q \subseteq A$, $MC_P = \{X_1, X_2, \ldots, X_m\}$ and $MC_Q = \{Y_1, Y_2, \ldots, Y_n\}$. A partial relation $\sqsubseteq$ is defined as follows:

$$P \sqsubseteq Q \Longleftrightarrow \quad \text{for every } X_i \in MC_P, \quad \text{there exists } Y_j \in MC_Q \quad \text{such that } X_i \subseteq Y_j.$$

From the definition of partial relation $\sqsubseteq$, it is easy to obtain the following lemma.

Lemma 1. *Let $S = (U, A)$ be an incomplete information system and $B \subseteq A$, then $MC_A \sqsubseteq MC_B$.*

Theorem 4 (Judgement theorem of consistent set I). *Let $S = (U, C \cup D)$ be an incomplete decision table and $B \subseteq C$. Then,*

(1) *B is a lower approximation consistent set iff* $\varphi_B(u) = \varphi_C(u), \ \forall u \in U$;

(2) *B is an upper approximation consistent set iff* $\eta_B(u) = \eta_C(u),\ \forall u \in U$.

Proof. Suppose $U/IND(D) = \{D_1, D_2, \ldots, D_r\}$. Since $B \subseteq C$, from Lemma 1, it follows that $MC_C \sqsubseteq MC_B$.

(1) "⇒" If B is a lower approximation consistent set, then $\underline{apr}_B(D) = \underline{apr}_C(D)$, i.e., one has that $\underline{apr}_B(D_j) = \underline{apr}_C(D_j)$, $j \leqslant r$. Hence, for any $u \in D_j$, if $u \in \underline{apr}_B(D_j)$, then u must belong to $\underline{apr}_C(D_j)$, i.e., $\varphi_B(u) = \varphi_C(u) = \{D_j\}$; if $u \notin \underline{apr}_B(D_j)$, then u does not belong to $\underline{apr}_C(D_j)$, i.e., $\varphi_B(u) = \varphi_C(u) = \emptyset$. From $U/IND(D)$ being a partition on the universe U, one therefore has that $\varphi_B(u) = \varphi_C(u),\ \forall u \in U$."⇐" Suppose $\varphi_B(u) = \varphi_C(u),\ \forall u \in U$. When $\varphi_B(u) = \{D_j\}$, there must exist some $X \in MC_B(u)$ such that $X \subseteq D_j$, i.e., $X \subseteq \underline{apr}_B(D_j)$. And since $MC_C \sqsubseteq MC_B$, we know that there exist $Y \in MC_C(u)$ such that $Y \subseteq X$. Hence, $Y \subseteq D_j$ and $Y \subseteq \underline{apr}_C(D_j)$. Thus, $u \in \underline{apr}_C(D_j)$. When $\varphi_B(u) = \emptyset$, we suppose $u \in \underline{apr}_C(D_j)$ and $u \notin \underline{apr}_B(D_j)$. In this situation, $\varphi_B(u) = \emptyset,\ \varphi_C(u) = \{D_j\}$, i.e., $\varphi_B(u) \neq \varphi_C(u)$. This yields a contradiction. Therefore, if $u \in \underline{apr}_C(D_j)$, one must have $u \in \underline{apr}_B(D_j)$ when $\varphi_B(u) = \varphi_C(u)$. Hence, B is a lower approximation consistent set if $\varphi_B(u) = \varphi_C(u),\ \forall u \in U$.

(2) "⇒" Denoted by $T(u) = \{D_k : u \in \overline{apr}_B(D_k),\ u \in U,\ D_k \in U/IND(D)\}$ and $T'(u) = \{D_k : u \in \overline{apr}_C(D_k),\ u \in U,\ D_k \in U/IND(D)\}$. If B is an upper approximation consistent set, then $\overline{apr}_B(D) = \overline{apr}_C(D)$, i.e., one has that $\overline{apr}_B(D_j) = \overline{apr}_C(D_j),\ j \leqslant r$. Hence, $T(u) = T'(u),\ \forall u \in U$. Therefore, we have that $\eta_B(u) = \eta_C(u)$."⇐" Let $\eta_B(u) = \eta_C(u),\ \forall u \in U$. We suppose that there exists $D_0 \in U/IND(D)$ such that $\overline{apr}_B(D_0) \neq \overline{apr}_C(D_0)$. Since $MC_C \sqsubseteq MC_B$, it follows that $\overline{apr}_C(D_0) \subseteq \overline{apr}_B(D_0)$. Let $u_0 \in \overline{apr}_B(D_0)$ and $u_0 \notin \overline{apr}_C(D_0)$, one has that $D_0 \in \eta_B(u_0)$ and $D_0 \notin \eta_C(u_0)$. Obviously, $\eta_B(u_0) \neq \eta_C(u_0)$. This yields a contradiction. Hence, B is an upper approximation consistent set if $\eta_B(u) = \eta_C(u),\ \forall u \in U$.

This completes the proof. □

Theorem 4 provides an approach to judge whether a subset of condition attributes is a lower/upper approximation consistent set or not.

Denoted by $r(X_C(u)) = \bigcup\{v : f_a(v) \text{ is regular},\ v \in X(u), a \in C\}$, where $X_C(u)$ is a maximal consistent block under C containing u and $f_a(v)$ is the value of v under the attribute a. Here, $r(X_C(u))$ is called a regular set of $X_C(u)$.

Based these discussions, in the following we give another judgement theorem of approximation consistent sets.

Theorem 5 (Judgement theorem of consistent set II). *Let* $S = (U, C \cup D)$ *be an incomplete decision table and* $B \subseteq C$. *Then,*

(1) *B is a lower approximation consistent set iff if* $u,\ v \in U$ *such that* $\varphi_C(u) \neq \varphi_C(v)$, *then* $r(X_B(u)) \cap r(X_B(v)) = \emptyset$ *when* $r(X_B(u)) \neq r(X_B(v))$;

(2) *B is an upper approximation consistent set iff if* $u,\ v \in U$ *such that* $\eta_C(u) \neq \eta_C(u)$, *then* $r(X_B(u)) \cap r(X_B(v)) = \emptyset$ *when* $r(X_B(u)) \neq r(X_B(v))$.

Proof

(1) "⇒" If $r(X_B(u)) \cap r(X_B(v)) \neq \emptyset$, one has that $r(X_B(u)) = r(X_B(v))$. So, $X_B(u) = X_B(v)$. From the definition of φ function, it follows that $\varphi_B(u) \neq \varphi_B(v)$. The assumption that B is a lower approximation consistent set implies that $\varphi_B(u) = \varphi_C(u)$ and $\varphi_B(v) = \varphi_C(v)$. Therefore, $\varphi_C(u) = \varphi_C(v)$. This yields a contradiction. Hence, $r(X_B(u)) \cap r(X_B(v)) = \emptyset$ when $r(X_B(u)) \neq r(X_B(v))$."⇐" From Lemma 1, we know that $MC_C \sqsubseteq MC_B$.If $\varphi_C(u) = \emptyset$, then for any maximal consistent block $X_C(u) \in MC_C(u)$ containing u, there exists some $X_B(u) \in MC_B(u)$ such that $X_C(u) \subseteq X_B(u)$, thus, we have that $\varphi_B(u) = \emptyset$.If $\varphi_C(u) = \{D_j\}$, for some $j \leqslant r$, there exists a maximal consistent block $X_C(u) \in MC_C(u)$ such that $X_C(u) \subseteq D_j$. For any $v \in X_B(u)$, since $r(X_B(u)) \cap r(X_B(v)) \neq \emptyset$, we have that $r(X_B(u)) = r(X_B(v))$, i.e., $X_B(u) = X_B(v)$. Hence, it follows from the assumption that $\varphi_C(u) = \varphi_C(v)$. Thus, $X_C(v) \subseteq D_j$. Therefore, $v \in D_j$. So, one has that $X_B(u) \subseteq D_j$. That is to say $\varphi_B(u) = \{D_j\}$.Therefore, $\varphi_C(u) = \varphi_B(u),\ \forall u \in U$. By Theorem 4 we conclude that B is a lower approximation consistent set.

(2) It is similar to the proof of (1).

This completes the proof. □

Theorem 5 provides another approach to judge whether a subset of condition attributes is lower/upper approximation consistent set or not.

Remark. As we know, if S is a complete decision table, then the maximal consistent blocks induced by the condition attribute set C will generate into the equivalence classes induced by C. In this case, the regular set $r(X_C(u))$ of a maximal consistent block $X_C(u)$ containing u is equivalent to the equivalence class $[u]_C$. Therefore, the judgement approaches to lower/upper approximation consistent set in Theorem 5 are also be used in a complete decision table.

4. Approaches to approximation reduction in inconsistent incomplete decision tables

In this section, the approaches to the lower approximation reduct and the upper approximation reduct are provided and an illustrate example is also employed to show their mechanisms.

Let $S = (U, C \cup D)$ be an incomplete decision table and $X, Y \in MC_C$. In order to introduce the notions of a lower/upper approximation discernibility attribute set, we denote

$$\varphi_C(X) = \begin{cases} D_j, & \text{if } \exists j \leqslant r \quad \text{such that } X \subseteq D_j, D_j \in U/IND(D), \\ \emptyset, & \text{otherwise}. \end{cases}$$

$$\eta_C(X) = \{D_j : X \cap D_j \neq \emptyset\}.$$

and

$$D_1^* = \{(X, Y) : \varphi_C(X) \neq \varphi_C(Y)\},$$
$$D_2^* = \{(X, Y) : \eta_C(X) \neq \eta_C(Y)\}.$$

where (X, Y) identifies with (Y, X) in $D_l^*,\ l \in \{1, 2\}$.

From the above denotations, it is easy to see that the maximal consistent blocks $X,\ Y$ in D_1^* cannot be included in the same decision class $D_j \in U/IND(D)$, and the maximal consistent blocks $X,\ Y$ in D_2^* cannot be included in the upper approximation of same decision class $D_j \in U/IND(D)$.

Denoted by $f_a(X) = f_a(u),\ a \in C, X \in U$ and $u \in r(X)$, where $r(X)$ is the regular set of a maximal consistent block X. For $a \in C$, we denote $f_a(X) = f_a(Y)$ iff $f_a(u) = f_a(v),\ u \in r(X),\ v \in r(Y)$ and $a \in C$; otherwise $f_a(X) \neq f_a(Y)$.

Based on these representations, we give the definitions of a lower approximation discernibility attribute set and an upper approximation discernibility attribute set in the following.

Definition 3. Let $S = (U, C \cup D)$ be an incomplete decision table and $MC_C = \{X_1, X_2, \ldots, X_m\}$. We denote

$$D_l(X_i, X_j) = \begin{cases} \{a \in C : f_a(X_i) \neq f_a(X_j)\}, & (X_i, X_j) \in D_l^*, \\ C, & (X_i, X_j) \notin D_l^*. \end{cases} \quad l \in \{1, 2\},$$

then $D_l(X_i, X_j),\ l \in \{1, 2\}$, are called a lower approximation discernibility attribute set and an upper approximation discernibility attribute set, respectively.

Based on the denotation in Definition 3, the below judgement theorem of a consistent set can be also obtained.

Theorem 6 (Judgement theorem of consistent set III). *Let $S=(U,C\cup D)$ be an incomplete decision table and $B\subseteq C$. Then,*

(1) *B is a lower approximation consistent set iff $B\cap D_1(X_i,X_j)\neq\emptyset$ for all $(X_i,X_j)\in D_1^*$;*
(2) *B is an upper approximation consistent set iff $B\cap D_2(X_i,X_j)\neq\emptyset$ for all $(X_i,X_j)\in D_2^*$.*

Proof

(1) "$\Rightarrow$" Suppose B is a lower approximation consistent set. For $\forall(X_i,X_j)\in D_1^*$, one can find $u,\ v\in U$ such that there exist two maximal consistent blocks $X_C(u),\ X_C(v)$ with $X_i=X_C(u)$ and $X_j=X_C(v)$, then $\varphi_C(u)\neq\varphi_C(v)$. We obtain by (1) in Theorem 5 that $r(X_B(u))\cap r(X_B(v))=\emptyset$ when $r(X_B(u))\neq r(X_B(v))$. Thus, there exists $a_0\in B$ such that $f_{a_0}(u)\neq f_{a_0}(v)$, i.e., $f_{a_0}(X_i)\neq f_{a_0}(X_j)$. It implies $a_0\in D_1(X_i,X_j)$. Then, $B\cap D_1(X_i,X_j)\neq\emptyset$."$\Leftarrow$" Conversely, if there exists $(X_i,X_j)\in D_1^*$ such that $B\cap D_1(X_i,X_j)=\emptyset$, one can select $u,\ v\in U$ and $X_C(u),\ X_C(v)$ satisfying $X_i=X_C(u),\ X_j=X_C(v)$. It should be noted that $\varphi_C(u)\neq\varphi_C(v)$. Then, for any $a\in B$, one has $a\notin D_1(X_i,X_j)$. Therefore, $f_a(X_i)=f_a(X_j)$. Consequently, $f_a(u)=f_a(v)$ for all $a\in B$, which implies $X_C(u)=X_C(v)$. Thus, by (1) in Theorem 5 we conclude that B is not a lower approximation consistent set. Hence, if $B\cap D_1(X_i,X_j)\neq\emptyset$ for all $(X_i,X_j)\in D_1^*$, then B must be a lower approximation consistent set.
(2) It is similar to the proof of (1).

This completes the proof. □

Theorem 6 provides an approach to approximation reduction in inconsistent incomplete decision tables. It is also available for a consistent incomplete decision table.

Definition 4. Let $S=(U,C\cup D)$ be an incomplete decision table, $D_l=(D_l(X_i,X_j),i,j\leqslant m),\ l\in\{1,2\},\ X_i,\ X_j\in MC_C$, are called lower and upper approximation discernibility matrices, respectively. Denoted by

$$M_l=\bigwedge\left\{\bigvee\{a:a\in D_l(X_i,X_j)\}:i,j\leqslant m\right\}$$
$$=\bigwedge\left\{\bigvee\{a:a\in D_l(X_i,X_j)\}:(X_i,X_j)\in D_l^*,\ (l\in\{1,2\})\right\}.$$

Then, $M_l(l\in\{1,2\})$ are, respectively, referred to as the lower approximation and upper approximation discernibility functions.

Through the lower approximation and upper approximation discernibility functions, we can design the approaches to the lower/upper approximation reduct in an inconsistent incomplete decision table as follows.

Theorem 7. *Let $S=(U,C\cup D)$ be an incomplete decision table. The minimal disjunctive normal form of each discernibility function $M_l\ (l\in\{1,2\})$ is*

$$M_l=\bigvee_{k=1}^{t}\left(\bigwedge_{s=1}^{q_k}a_{i_s}\right),\quad(l=1,2).$$

Denoted by $B_{lk}=\{a_{i_s}:s=1,2,\ldots,q_k\}$, and then $\{B_{lk}:k=1,2,\ldots,t\}\ (l=1,2)$ are, respectively, the set of all lower and upper approximation reducts.

Proof. The proof is same as the idea of all minimal disjunctive normal form of each discernibility function in rough set theory. We omit this proof in this paper. □

Theorem 7 provides practical approaches to some attribute reductions in inconsistent incomplete decision tables.

In the following, an illustrative example is employed to analyze the mechanisms of the approaches established in the present research.

Example 3 (*Continued from Example 1*). Compute the lower approximation reduct and upper approximation reduct of this decision table (see Table 1).

By computing, it follows that:

$$MC_C=\{\{u_1\},\{u_2,u_6\},\{u_3\},\{u_4,u_5\},\{u_5,u_6\}\},$$

We denote the maximal consistent blocks of objects by

$$X_1=\{u_1\},\quad X_2=\{u_2,u_6\},\quad X_3=\{u_3\},\quad X_4=\{u_4,u_5\},$$
$$X_5=\{u_5,u_6\}.$$

And, we denote the decision classes of objects by

$$D_1=\{u_1,u_2,u_4,u_6\},\quad D_2=\{u_3\},\quad D_3=\{u_5\}.$$

At first, we compute the lower approximation reducts of this incomplete decision table.

It can easily be calculated that

$$\varphi_C(X_1)=\{D_1\},\quad\varphi_C(X_2)=\{D_1\},\quad\varphi_C(X_3)=\{D_2\},$$
$$\varphi_C(X_4)=\{\emptyset\},\quad\varphi_C(X_5)=\{\emptyset\}.$$

Hence,

$$D_1^*=\{(X_1,X_3),(X_1,X_4),(X_1,X_5),(X_2,X_3),(X_2,X_4),(X_2,X_5),(X_3,X_4),(X_3,X_5)\}.$$

From Definitions 3 and 4, one can obtain the lower approximation discernibility matrix. Because of its symmetry, it is only necessary to write half of the discernibility matrix (see Table 2).

Therefore, it follows from Theorem 7 that:

$$M_1=(a_1\vee a_2\vee a_3\vee a_4)\wedge a_3\wedge a_4\wedge(a_1\vee a_2\vee a_4)\wedge(a_1\vee a_4)\wedge(a_3\vee a_4)\wedge\{a_1,a_2,a_3,a_4\}=a_3\wedge a_4.$$

By Theorem 7, we conclude that $\{a_3,a_4\}$ is the unique lower approximation reduct of this incomplete decision table.

Then, we compute the upper approximation reducts of this incomplete decision table.

It can easily be calculated that

$$\eta_C(X_1)=\{D_1\},\quad\eta_C(X_2)=\{D_1\},\quad\eta_C(X_3)=\{D_2\},$$
$$\eta_C(X_4)=\{D_1,D_3\},\quad\eta_C(X_5)=\{D_1,D_3\}.$$

Thus,

$$D_2^*=\{(X_1,X_3),(X_1,X_4),(X_1,X_5),(X_2,X_3),(X_2,X_4),(X_2,X_5),(X_3,X_4),(X_3,X_5)\}.$$

From Definitions 3 and 4, one can obtain the upper approximation discernibility matrix. Similar to Table 2, because of its symmetry we only write half of the discernibility matrix (see Table 3).

Table 2
The lower approximation discernibility matrix of the incomplete decision table about cars.

X_i/X_j	X_1	X_2	X_3	X_4	X_5
X_1	$\emptyset$	$\{a_1,a_2,a_3,a_4\}$	$\{a_3\}$	$\{a_4\}$	$\{a_1,a_2,a_4\}$
X_2		$\emptyset$	$\{a_3\}$	$\{a_1,a_4\}$	$\{a_4\}$
X_3			$\emptyset$	$\{a_3,a_4\}$	$\{a_3,a_4\}$
X_4				$\emptyset$	$\{a_1,a_2,a_3,a_4\}$
X_5					$\emptyset$

Therefore, it follows from Theorem 7 that:

$$M_2 = (a_1 \vee a_2 \vee a_3 \vee a_4) \wedge a_3 \wedge a_4 \wedge (a_1 \vee a_2 \vee a_4) \wedge (a_1 \vee a_4) \wedge (a_3 \vee a_4) \wedge \{a_1, a_2, a_3, a_4\} = a_3 \wedge a_4.$$

By Theorem 7, we conclude that $\{a_3, a_4\}$ is the unique upper approximation reduct of the incomplete decision table.

In what follows, we analyze the relationship between the attribute reduction method and other approaches of discernibility matrix to reduce attributes within rough set theory. As we know, discernibility matrix method can be used to find all attribute reducts from an information systems or a decision table, which has been widely applied in feature selection in rough set theory. Using the framework, many discernibility matrix versions have been developed, and each version is based on a special definition of discernibility attribute set for its corresponding definition of attribute reduction. In this study, unlike each of existing approaches of discernibility matrix, we define two more different discernibility matrix approaches through employing maximal consistent block technique for acquiring lower/upper approximation discernibility attribute set and obtaining lower/upper approximation reduction from an inconsistent incomplete decision table.

Five data sets from the University of California at Irvine (UCI) Machine Learning Repository are used in the empirical study. The information about these five data sets is shown in Table 4. The objective of these experiments is to show the power of the proposed method to attribute reduction.

Table 5 displays the lower approximation reducts on five public data sets. It can be seen from Table 5 that the voting-records has 1 reduct only, the soybean-large has 682 reducts, the spect has 267 reducts, the zoo has 101 reducts and the tic-tac-toe has 9 reducts. These results show that the proposed discernibility matrix approach can obtain all lower/upper approximation attribute reducts, but not a single reduct. It is deserved to point out that the time consumption of discernibility matrix approach is very disappointing, and a heuristic strategy is always employed when dealing with large-scale data sets.

By adopting the technique of maximal consistent blocks, a variety of discernibility functions in inconsistent incomplete decision tables becomes simpler in the proposed approach. In a maximal consistent block, all objects are not indiscernible with available information provided by a similarity relation. Therefore, by using maximal consistent blocks as units to construct discernibility matrices, their size can be reduced. This means that the algorithms for finding various approximation reducts requires smaller memory. Therefore, it makes the absorption law (used to simplify a discernibility function for obtaining all prime implicants) more efficient.

Table 3
The upper approximation discernibility matrix of the incomplete decision table about cars.

X_i/X_j	X_1	X_2	X_3	X_4	X_5
X_1	Ø	$\{a_1, a_2, a_3, a_4\}$	$\{a_3\}$	$\{a_4\}$	$\{a_1, a_2, a_4\}$
X_2		Ø	$\{a_3\}$	$\{a_1, a_4\}$	$\{a_4\}$
X_3			Ø	$\{a_3, a_4\}$	$\{a_3, a_4\}$
X_4				Ø	$\{a_1, a_2, a_3, a_4\}$
X_5					Ø

Table 4
Data description.

	Data sets	Samples	Features	Data type	Classes
1	Voting-records	435	16	Incomplete	2
2	Soybean-large	307	35	Incomplete	19
3	Spect	267	22	Complete	2
4	Zoo	101	16	Complete	7
5	Tic-tac-toe	958	9	Complete	2

Table 5
Lower approximation reducts.

	Data sets	Reduct numbers	Features of the shortest reduct	Features of the longest reduct
1	Voting-records	1	14	14
2	Soybean-large	682	11	16
3	Spect	267	14	15
4	Zoo	101	8	9
5	Tic-tac-toe	9	8	8

Remark. In this study, we develop an approach to find all the set of all lower and upper approximation reducts based on discernibility functions. But the process to calculate the disjunctive normal form is an NP-hard problem. However, it cannot be easily calculated by computer if there are hundreds and thousands objects in a given data set. In fact, it cannot be applied in practical applications. Therefore, more applicable approaches such as heuristic algorithm is desirable. We will investigate this work in the future.

5. Conclusions

To acquire brief decision rules from inconsistent incomplete decision tables, attribute reductions in the condition part are needed. This paper has introduced a kind of attribute reductions called a lower/upper approximation reduct, which preserve the lower/upper approximation distribution of the decision classes. The judgement theorems and discernibility matrices associated with the two reductions have been obtained. Then, the practical approaches to lower/upper approximation reduct in inconsistent incomplete decision tables have been provided as well. Finally, an illustrative example has been employed to explain the mechanism of this kind of attribution reduction methods. The proposed method renders a set of simpler discernibility functions for finding all approximation reducts of an inconsistent incomplete decision table. Though the procedure is developed in the context of inconsistent incomplete decision tables, it is apparent that the methodology is also applicable to consistent decision tables and complete decision tables.

Acknowledgements

The authors wish to thank the anonymous reviewers for their constructive comments on this study.

This work was supported by the National Natural Science Foundation of China (Nos. 60903110, 60773133, 70971080, 60970014), the National High Technology Research and Development Program of China (No. 2007AA01Z165), the Natural Science Foundation of Shanxi Province, China (Nos. 2008011038, 2009021017-1).

References

[1] M. Beynon, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research 134 (3) (2001) 592–605.
[2] Y.Y. Guan, H.K. Wang, Set-valued information systems, Information Science 176 (17) (2006) 2507–2525.
[3] Q. Hu, J. Liu, D. Yu, Mixed feature selection based on granulation and approximation, Knowledge based systems 21 (2008) 294–304.
[4] Q. Hu, D. Yu, Z. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, Pattern Recognition Letters 27 (2006) 414–423.
[5] Q. Hu, Z. Xie, D. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition 40 (2007) 3509–3521.

[6] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: a tutorial, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hybridization: A New Trend in Decision Making, Springer, Singapore, 1999, pp. 3–98.
[7] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1998) 39–49.
[8] M. Kryszkiewicz, Rule in incomplete information systems, Information Sciences 113 (1999) 271–292.
[9] M. Kryszkiewicz, Comparative study of alternative type of knowledge reduction in inconsistent systems, International Journal of Intelligent Systems 16 (2001) 105–120.
[10] D.Y. Li, B. Zhang, Y. Leung, On knowledge reduction in inconsistent decision information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (5) (2004) 651–672.
[11] Y. Leung, D.Y. Li, Maximal consistent block technique for rule acquisition in incomplete information systems, Information Sciences 153 (2003) 85–106.
[12] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Science in China F: Information Sciences 51 (10) (2008) 1427–1444.
[13] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information system, International Journal of General Systems 35 (6) (2006) 641–654.
[14] J.Y. Liang, Z.B. Xu, The algorithm on knowledge reduction in incomplete information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 24 (1) (2002) 95–103.
[15] E. Marczewski, A general scheme of independence in mathematics, Bulletin de l'Acad'emie Polonaise des Sciences, S'erie des sci. math, Astronom. Phys. 6 (1958) 331–362.
[16] J.S. Mi, W.Z. Wu, W.X. Zhang, Comparative studies of knowledge reductions in inconsistent systems, Fuzzy Systems and Mathematics 17 (3) (2003) 54–60.
[17] J.S. Mi, W.Z. Wu, W. X Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (2004) 255–272.
[18] D.Q. Miao, Y. Zhao, Y.Y. Yao, H.X. Li, F.F. Xu, Relative reducts in consistent and inconsistent decision tables of the Pawlak rough set model, Information Sciences 179 (24) (2009) 4140–4150.
[19] H.S. Nguyen, D. Slezak, Approximate reducts and association rules correspondence and complexity results, in: N. Zhong, A. Skowron, S. Oshuga (Eds.), Proceeding of RSFDGrc'99, Yamaguchi, Japan, LNAI 1711, 1999, pp. 137–145.
[20] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[21] Z. Pawlak, Rough set theory and its applications in data analysis, Cybernetics and Systems 29 (1998) 661–688.
[22] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
[23] Y.H. Qian, J.Y. Liang, C.Y. Dang, H.Y. Zhang, J.M. Ma, On the evaluation of the decision performance of an incomplete decision table, Data & Knowledge Engineering 65 (3) (2008) 373–400.
[24] Y.H. Qian, J.Y. Liang, C.Y. Dang, Converse approximation and rule extracting from decision tables in rough set theory, Computers and Mathematics with Applications 55 (2008) 1754–1765.
[25] Y.H. Qian, J.Y. Liang, D.Y. Li, H.Y. Zhang, C.Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences 178 (2008) 181–202.
[26] M. Quafatou, α-RST: a generalization of rough set theory, Information Sciences 124 (2000) 301–316.
[27] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R. Słowiński (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer, Academic, Dordrecht, 1992, pp. 331–362.
[28] D. Slezak, Approximate reducts in decision tables, in: Proceeding of IPMU'96, Granada, Spain, vol. 3, 1996, pp. 1159–1164.
[29] D. Slezak, Searching for dynamic reducts in inconsistent decision tables, in: Proceeding of IPMU'98, France, vol. 2, 1998, pp. 1362–1369.
[30] W.Z. Wu, M. Zhang, H.Z. Li, J.S. Mi, Knowledge reduction in random information systems via Dempster–Shafer theory of evidence, Information Sciences 174 (2005) 143–164.
[31] D. Yu, Q. Hu, C. Wu, Uncertainty measures for fuzzy relations and their applications, Applied Soft Computing 7 (2007) 1135–1143.
[32] W. Ziarko, Variable precision rough set model, Journal of Computer System Science 46 (1993) 39–59.
[33] W.X. Zhang, J.S. Mi, W.Z. Wu, Approaches to knowledge reductions in inconsistent systems, International Journal of Intelligent Systems 18 (2003) 989–1000.
[34] W.X. Zhang, G.F. Qiu, Uncertain Decision Making Based on Rough Sets, Tsinghua Press, Beijing, China, 2005.
[35] W.X. Zhang, G.F. Qiu, W.Z. Wu, A general approach to attribute reduction in rough set theory, Science in China F: Information Sciences 50 (2) (2007) 188–197.
[36] W. Zhu, F.Y. Wang, Reduction and axiomization of covering generalized rough sets, Information Sciences 152 (2003) 217–230.
[37] X.Z. Zhou, B. Huang, Rough set-based attribute reduction under incomplete information systems, Journal of Nanjing University of Science and Technology 27 (5) (2003) 630–635.