Knowl Inf Syst (2010) 25:169–184
DOI 10.1007/s10115-009-0248-5

REGULAR PAPER

Attribute reduction in ordered information systems based on evidence theory

Wei-hua Xu · Xiao-yan Zhang · Jian-min Zhong · Wen-xiu Zhang

Received: 22 June 2007 / Revised: 21 May 2009 / Accepted: 9 August 2009 /
Published online: 3 September 2009

Abstract Attribute reduction is one of the most important problems in rough set theory. However, in real-world lots of information systems are based on dominance relation in stead of the classical equivalence relation because of various factors. The ordering properties of attributes play a crucial role in those systems. To acquire brief decision rules from the systems, attribute reductions are needed. This paper deals with attribute reduction in ordered information systems based on evidence theory. The concepts of plausibility and belief consistent sets as well as plausibility and belief reducts in ordered information systems are introduced. It is proved that a plausibility consistent set must be a consistent set and an attribute set is a belief reduct if and only if it is a classical reduction in ordered information system.

Keywords Attribute reduction · Consistent set · Evidence theory · Rough set · Ordered information system

1 Introduction

The rough set theory, proposed by Pawlak in the early 1980s [16], is an extension of sets theory for the study of intelligent systems characterized by inexact, uncertain or vague information and can serve as a new mathematical tool to soft computing. This theory has been applied successfully in machine learning, pattern recognition, decision support systems,

W.-h. Xu (✉) · X.-y. Zhang · J.-m. Zhong
School of Mathematics and Physics, Chongqing University of Technology, 400054 Chongqing, China
e-mail: datongxuweihua@126.com

X.-y. Zhang
e-mail: datongzhangxiaoyan@126.com

J.-m. Zhong
e-mail: ccqzjm@sina.com

W.-x. Zhang
School of Science, Xi'an Jiaotong University, Xi'an, China
e-mail: wxzhang@mail.xjtu.edu.cn

expert
systems, data analysis, data mining, and so on [11,17,18,26]. Recently, the theory has generated a great deal of interest among more and more researchers.

Another important method used to deal with uncertainty in information systems is the Dempster–Shafer theory of evidence. It was originated by Dempster's concept of lower and upper probabilities [4], and extended by Shafer as a theory [20]. The basic representational structure in this theory is a belief structure which consists of a family of subsets, called focal elements, with associated individual positive weights summing to one. The primitive numeric measures derived from the belief structure are a dual pair of belief and plausibility functions.

There are strong connections between rough set theory and Dempster–Shafer theory of evidence. It has been demonstrated that various belief structures are associated with various rough approximation spaces such that the different dual pairs of lower and upper approximation operators induced by rough approximation spaces may be used to interpret the corresponding dual pairs of belief and plausibility functions induced by belief structures [22,28,29,32].

It is well known that knowledge reduction is one of the hot research topics of rough set theory. Much study on this area had been reported and many useful results were obtained until now [1,13–15,23,30,33–35]. However, the original rough set theory approaches do not consider attributes with preference-ordered domains, that is, criteria. In many real situations, we are often face to the problems in which the ordering of properties of the considered attributes plays a crucial role. One such type of problem is the ordering of objects. For this reason, Greco, Matarazzo, and Slowinski proposed an extension of rough set theory, called the dominance-based rough set approach(DRSA) to take into account the ordering properties of criteria [5–10]. This innovation is mainly based on substitution of the indiscernibility relation by a dominance relation. Moreover, Greco, Matarazzo, and Slowinski characterize the DRSA as well as decision rules induced from rough approximations, while the usefulness of the DRSA and its advantages over the CRSA (classical rough set approach) are presented [5–10]. In DRSA, condition attributes are criteria and classes are preference ordered. Several studies have been made about properties and algorithmic implementations of DRSA [2,3,12,19,21,24,25,27]. In particular, in [24], an algorithm for reduction of attributes in ordered information systems has been given, based on generalization of the concept of discernibility matrix. In this paper, we attempt to investigate attribute reduction in ordered information systems basing on evidence theory and its strong relations with rough set theory.

The organization of the paper is as follows. In Sect. 2, we give some basic notions of rough set and ordered information systems. In Sect. 3, evidence theory in ordered information systems have been constructed. And we provide an approach for transforming from approximations to belief structures of ordered information systems in this section. In Sect. 4, the notions of belief and plausibility reductions in ordered information systems are proposed and the relationships between the new concepts of reductions and the classical reduction are examined. We then conclude the paper with a summary and out-look for further research in Sect. 5.

2 Rough sets and ordered information systems

The following recalls necessary concepts and preliminaries required in the sequel of our work. Detailed description of the theory can be found in the source papers [6,8–10]. A description has also been made in [35].

The notion of information system (sometimes called data tables, attribute-value systems, knowledge representation systems, etc.) provides a convenient tool for the representation of objects in terms of their attribute values.

An information system is an ordered triple $\mathcal{I} = (U, A, F)$, where $U = \{x_1, x_2, \ldots, x_n\}$ is a non-empty finite set of objects called the universe, and $A = \{a_1, a_2, \ldots, a_p\}$ is a non-empty finite set of attributes, such that there exists a map $f_l : U \longrightarrow V_{a_l}$ for any $a_l \in A$, where V_{a_l} is called the domain of the attribute a_l, and denoted $F = \{f_l | a_l \in A\}$.

In an information systems, if the domain of an attribute is ordered according to a decreasing or increasing preference, then the attribute is a criterion.

Definition 2.1 (See [6,8,9]) An information system is called an ordered information system(OIS) if all condition attributes are criteria.

Assumed that the domain of a criterion $a \in A$ is complete pre-ordered by an outranking relation $\succeq_a$, then $x \succeq_a y$ means that x is at least as good as y with respect to criterion a. And we can say that x dominates y. In the following, without any loss of generality, we consider criterions having a numerical domain, that is, $V_a \subseteq \mathcal{R}$($\mathcal{R}$ denotes the set of real numbers).

We define $x \succeq y$ by $f(x, a) \geq f(y, a)$ according to increasing preference, where $a \in A$ and $x, y \in U$. For a subset of attributes $B \subseteq A$, $x \succeq_B y$ means that $x \succeq_a y$ for any $a \in B$, and that is to say x dominates y with respect to all attributes in B. Furthermore, we denote $x \succeq_B y$ by $x R_B^{\succeq} y$. In general, we denote a ordered information systems by $\mathcal{I}^{\succeq} = (U, A, F)$. Thus the following definition can be obtained.

Definition 2.2 (See [6,8,9]) Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information, for $B \subseteq A$, denote

$$R_B^{\succeq} = \{(x, y) \in U \times U | f_l(x) \geq f_l(y), \forall a_l \in B\};$$

$R_B^{\succeq}$ are called dominance relations of ordered information system $\mathcal{I}^{\succeq}$.

Let denote

$$\begin{aligned} [x_i]_B^{\succeq} &= \{x_j \in U | (x_j, x_i) \in R_B^{\succeq}\} \\ &= \{x_j \in U | f_l(x_j) \geq f_l(x_i), \forall a_l \in B\}; \\ U/R_B^{\succeq} &= \{[x_i]_B^{\succeq} | x_i \in U\}, \end{aligned}$$

where $i \in \{1, 2, \ldots, |U|\}$, then $[x_i]_B^{\succeq}$ will be called a dominance class or the granularity of information, and $U/R_B^{\succeq}$ be called a classification of U about attribute set B.

The following properties of a dominance relation are trivial by the above definition.

Proposition 2.1 (See [6,8,9]) *Let $R_A^{\succeq}$ be a dominance relation.*

(1) *$R_A^{\succeq}$ is reflexive,transitive, but not symmetric, so it is not an equivalence relation.*
(2) *If $B \subseteq A$, then $R_A^{\succeq} \subseteq R_B^{\succeq}$.*
(3) *If $B \subseteq A$, then $[x_i]_A^{\succeq} \subseteq [x_i]_B^{\succeq}$*
(4) *If $x_j \in [x_i]_A^{\succeq}$, then $[x_j]_A^{\succeq} \subseteq [x_i]_A^{\succeq}$ and $[x_i]_A^{\succeq} = \cup\{[x_j]_A^{\succeq} | x_j \in [x_i]_A^{\succeq}\}$.*
(5) *$[x_j]_A^{\succeq} = [x_i]_A^{\succeq}$ iff $f(x_i, a) = f(x_j, a)$ for all $a \in A$.*
(6) *$|[x_i]_B^{\succeq}| \geq 1$ for any $x_i \in U$.*
(7) *$U/R_B^{\succeq}$ constitute a covering of U, i.e., for every $x \in U$ we have that $[x]_B^{\succeq} \neq \phi$ and $\bigcup_{x \in U} [x]_B^{\succeq} = U$.*

where $|\cdot|$ denotes cardinality of the set.

For any subset X of U and A of $\mathcal{I}^{\succeq}$, the lower and upper approximation of X with respect to a dominance relation $R_A^{\succeq}$ could be defined as following [8]:

$$\underline{R_A^{\succeq}}(X) = \{x \in U | [x]_A^{\succeq} \subseteq X\};$$
$$\overline{R_A^{\succeq}}(X) = \{x \in U | [x]_A^{\preceq} \cap X \neq \phi\}.$$

where $[x_i]_B^{\preceq} = \{x_j \in U | f_l(x_j) \leq f_l(x_i), \forall a_l \in B\}$.

From the above definition of rough approximation, the following important properties in ordered information systems have been proved, which are similar to those of Pawlak approximation spaces.

Proposition 2.2 (See [10]) *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information systems and $X \subseteq U$. The rough approximation can be expressed as union of elementary sets. That is to say the following holds.*

$$\underline{R_A^{\succeq}}(X) = \bigcup_{x \in U} \{[x]_A^{\succeq} | [x]_A^{\succeq} \subseteq X\};$$
$$\overline{R_A^{\succeq}}(X) = \bigcup_{x \in U} \{[x]_A^{\preceq} | [x]_A^{\preceq} \cap X \neq \phi\}.$$

Proposition 2.3 (See [10]) *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information systems and $X, Y \subseteq U$, then its lower and upper approximations satisfy the following properties.*

(1) $\underline{R_A^{\succeq}}(X) \subseteq X \subseteq \overline{R_A^{\succeq}}(X)$.

(2) $\overline{R_A^{\succeq}}(X \cup Y) = \overline{R_A^{\succeq}}(X) \cup \overline{R_A^{\succeq}}(Y)$;
$\underline{R_A^{\succeq}}(X \cap Y) = \underline{R_A^{\succeq}}(X) \cap \underline{R_A^{\succeq}}(Y)$.

(3) $\underline{R_A^{\succeq}}(X) \cup \underline{R_A^{\succeq}}(Y) \subseteq \underline{R_A^{\succeq}}(X \cup Y)$;
$\overline{R_A^{\succeq}}(X \cap Y) \subseteq \overline{R_A^{\succeq}}(X) \cap \overline{R_A^{\succeq}}(Y)$.

(4) $\underline{R_X^{\succeq}}(\sim X) = \sim \overline{R_A^{\succeq}}(X)$;
$\overline{R_A^{\succeq}}(\sim X) = \sim \underline{R_A^{\succeq}}(X)$.

(5) $\underline{R_A^{\succeq}}(U) = U$; $\overline{R_A^{\succeq}}(\phi) = \phi$.

(6) $\underline{R_A^{\succeq}}(X) = \underline{R_A^{\succeq}}(\underline{R_A^{\succeq}}(X)) = \overline{R_A^{\succeq}}(\underline{R_A^{\succeq}}(X))$;
$\overline{R_A^{\succeq}}(X) = \underline{R_A^{\succeq}}(\overline{R_A^{\succeq}}(X)) = \overline{R_A^{\succeq}}(\overline{R_A^{\succeq}}(X))$.

(7) *If $X \subseteq Y$, then $\underline{R_A^{\succeq}}(X) \subseteq \underline{R_A^{\succeq}}(Y)$ and $\overline{R_A^{\succeq}}(X) \subseteq \overline{R_A^{\succeq}}(Y)$.*

where $\sim X$ is the complement of X.

Example 2.1 Given an ordered information system in Table 1.

From Table 1 we have

$$[x_1]_A^{\succeq} = \{x_1, x_2, x_5, x_6\},$$
$$[x_2]_A^{\succeq} = \{x_2, x_5, x_6\},$$
$$[x_3]_A^{\succeq} = \{x_2, x_3, x_4, x_5, x_6\},$$
$$[x_4]_A^{\succeq} = \{x_4, x_6\},$$
$$[x_5]_A^{\succeq} = \{x_5\},$$
$$[x_6]_A^{\succeq} = \{x_6\},$$

Table 1 An ordered information system

U	a_1	a_2	a_3
x_1	1	2	1
x_2	3	2	2
x_3	1	1	2
x_4	2	1	3
x_5	3	3	2
x_6	3	2	3

and

$$\begin{aligned}
[x_1]_A^{\preceq} &= \{x_1\},\\
[x_2]_A^{\preceq} &= \{x_1, x_2, x_3\},\\
[x_3]_A^{\preceq} &= \{x_3\},\\
[x_4]_A^{\preceq} &= \{x_3, x_4\};\\
[x_5]_A^{\preceq} &= \{x_1, x_2, x_3, x_5\},\\
[x_6]_A^{\preceq} &= \{x_1, x_2, x_3, x_4, x_6\}.
\end{aligned}$$

Thus, it is obviously that $U/R_A^{\succeq} = \{[x_1]_A^{\succeq}, [x_2]_A^{\succeq}, [x_3]_A^{\succeq}, [x_4]_A^{\succeq}, [x_5]_A^{\succeq}, [x_6]_A^{\succeq}\}$.
And if we let $X = \{x_3, x_5, x_6\}$, then

$$\overline{R_A^{\succeq}}(X) = \{x_2, x_3, x_4, x_5, x_6, \}$$

$$\underline{R_A^{\succeq}}(X) = \{x_5, x_6\}.$$

It is clear that

$$\underline{R_A^{\succeq}}(X) \subseteq X \subseteq \overline{R_A^{\succeq}}(X).$$

For simple description, in the following context information systems are based on dominance relations generally, i.e., ordered information systems.

3 Evidence theory in ordered information systems

In evidence theory [4,20], for a universe U a mass function can be defined by a map $m : 2^U \longrightarrow [0, 1]$, which is called a basic probability assignment and satisfies two axioms:

$$\begin{aligned}
&(M1) \qquad m(\phi) = 0\\
&(M2) \qquad \sum_{X \subseteq U} m(X) = 1.
\end{aligned}$$

The value $m(X)$ represents the degree of belief that a specific element of U belongs to set X, but not to any particular subset of X. A subset $X \subseteq U$ with $m(X) > 0$ is called a focal element.

We denote by $\mathcal{M}$ the family of all focal elements of m. The pair $(\mathcal{M}, m)$ is called a belief structure. Associated with each belief structure in information systems based classical equivalence relation, a pair of belief and plausibility functions can be derived.

Definition 3.1 (See [4,20]) Let $(\mathcal{M}, m)$ be a belief structure. A set function $Bel : 2^U \longrightarrow [0, 1]$ is referred to as a belief function on U, if

$$Bel(X) = \sum_{Y \subseteq X} m(Y), \quad \forall X \in 2^U.$$

A set function $Pl : 2^U \longrightarrow [0, 1]$ is referred to as a plausibility function on U, if

$$Pl(X) = \sum_{Y \cap X \neq \phi} m(Y), \quad \forall X \in 2^U.$$

From above, we can shown that a mass function of classical information systems is a basic probability assignment. However, relations which are induced by attributes sets are not equivalence relations in ordered information systems. Based on the observation, a mass function is defined in ordered information systems as follows.

Definition 3.2 Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system. If we denote

$$h(X) = \left\{x \in U | [x]_A^{\succeq} = X\right\}$$

for any $X \in U/R_A^{\succeq}$, then a mass function of $\mathcal{I}^{\succeq}$ can be defined by a map $m : U/R_A^{\succeq} \rightarrow [0, 1]$, where

$$m(X) = \frac{|h(X)|}{|U|}.$$

By above definition, we can easily find that a mass function of ordered information systems still satisfies two basic axioms. In other words, for any $X \in U/R_A^{\succeq}$ in ordered information systems, the following hold directly.

$$(M1) \quad m(\phi) = 0$$
$$(M2) \quad \sum_{X \in U/R_A^{\succeq}} m(X) = 1.$$

As same as classical information systems we denote by $\mathcal{M}$ the family of all focal elements of m in ordered information systems. The pair $(\mathcal{M}, m)$ is called a belief structure of an ordered information system, and a pair of belief and plausibility functions in ordered information systems can be constructed immediately.

Definition 3.3 Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system, and $(\mathcal{M}, m)$ be a belief structure. A set function $Bel : 2^U \longrightarrow [0, 1]$ is referred to as a belief function on U, if

$$Bel(X) = \sum_{Y \subseteq X,\ Y \in U/R_A^{\succeq}} m(Y), \quad \forall X \in 2^U.$$

A set function $Pl : 2^U \longrightarrow [0, 1]$ is referred to as a plausibility function on U, if

$$Pl(X) = \sum_{Y \cap X \neq \phi,\ Y \in U/R_A^{\succeq}} m(Y), \quad \forall X \in 2^U.$$

Belief and plausibility functions based on the same belief structure are connected by the dual property:

$$Pl(X) = 1 - Bel(\sim X)$$

and furthermore, $Bel(X) \leq Pl(X)$ for all $X \in 2^U$.

The following theorems shows that the classical belief and plausibility functions can be interpreted in terms of the Pawlak's lower and upper approximations of sets [32].

Theorem 3.1 (See [32]) *Let (U, A, F) an information system, for any $X \subseteq U$, $B \subseteq A$, denoted*

$$Bel_B(X) = \frac{|\underline{R_B}(X)|}{|U|};$$

$$Pl_B(X) = \frac{|\overline{R_B}(X)|}{|U|},$$

Then $Bel_B(X)$is the belief function and $Pl_B(X)$ is the plausibility function of U, where the corresponding mass distribution is

$$m_B(Y) = \begin{cases} P(Y), & \text{if } Y \in U/R_B; \\ 0, & \text{otherwise.} \end{cases}$$

There are strong connections between rough set theory and the Dempster–Shafer theory of evidence, the detailed description can be found in [22,28,31].

Hence, we can acquire the following results which show that the pair of lower and upper approximation operators in ordered information systems generates a pair of belief and plausibility functions respectively.

Theorem 3.2 *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system, for any $X \subseteq U$, $B \subseteq A$, denoted*

$$Bel_B^{\succeq}(X) = \frac{|\underline{R_B^{\succeq}}(X)|}{|U|};$$

$$Pl_B^{\succeq}(X) = \frac{|\overline{R_B^{\succeq}}(X)|}{|U|},$$

Then $Bel_B^{\succeq}(X)$is the belief function and $Pl_B^{\succeq}(X)$ is the plausibility function of U, where the corresponding mass distribution is

$$m_B(Y) = \begin{cases} \frac{|h(Y)|}{|U|}, & \text{if } Y \in U/R_B; \\ 0, & \text{otherwise.} \end{cases}$$

.

Proof From [32] we can see that we only prove the following facts.

(1) $\frac{|\underline{R_B^{\succeq}}(\phi)|}{|U|} = 0;$

(2) $\frac{|\underline{R_B^{\succeq}}(U)|}{|U|} = 1;$

(3) For every positive integer n and every collection $X_1, X_2, \ldots, X_n \subseteq U$,

$$\frac{|\underline{R_B^{\succeq}}(X_1 \cup X_2 \cup \cdots \cup X_n)|}{|U|} \geq \sum_i \frac{|\underline{R_B^{\succeq}}(X_i)|}{|U|} - \sum_{i<j} \frac{|\underline{R_B^{\succeq}}(X_i \cap X_j)|}{|U|}$$
$$\pm \cdots \pm (-1)^{n+1} \frac{|\underline{R_B^{\succeq}}(X_1 \cap X_2 \cap \cdots \cap X_n)|}{|U|}.$$

Next, we will give proof of the above.

In facts, one can find that $\frac{|\underline{R_B^{\succeq}}(X)|}{|U|}$ satisfies (1) and (2) from Proposition 2.3(5). Consider a collection $X_1, X_2, \ldots, X_n \subseteq U$, according to (2) and (3) of Proposition 2.3, we have

$$
\begin{aligned}
&\frac{|\underline{R_B^{\succeq}}(X_1 \cup X_2 \cup \cdots \cup X_n)|}{|U|} \\
&\geq \frac{|\underline{R_B^{\succeq}}(X_1) \cup \cdots \cup \underline{R_B^{\succeq}}(X_n)|}{|U|} \\
&= \sum_i \frac{|\underline{R_B^{\succeq}}(X_i)|}{|U|} - \sum_{i<j} \frac{|\underline{R_B^{\succeq}}(X_i) \cap \underline{R_B^{\succeq}}(X_j)|}{|U|} \\
&\quad \pm \cdots \pm (-1)^{n+1} \frac{|\underline{R_B^{\succeq}}(X_1) \cap \underline{R_B^{\succeq}}(X_2) \cap \cdots \cap \underline{R_B^{\succeq}}(X_n)|}{|U|} \\
&= \sum_i \frac{|\underline{R_B^{\succeq}}(X_i)|}{|U|} - \sum_{i<j} \frac{|\underline{R_B^{\succeq}}(X_i \cap X_j)|}{|U|} \\
&\quad \pm \cdots \pm (-1)^{n+1} \frac{|\underline{R_B^{\succeq}}(X_1 \cap X_2 \cap \cdots \cap X_n)|}{|U|}.
\end{aligned}
$$

Thus, we proved that $Bel_B^{\succeq}(X)$ is the belief function of U. And one can obtain directly that $Pl_B^{\succeq}(X)$ is the plausibility function since the duality between $Bel_B^{\succeq}(X)$ and $Pl_B^{\succeq}(X)$.

So, the theorem was proved. □

Combining Theorem 3.2 and Proposition 2.3, we have the following corollary.

Corollary 3.1 *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system and $C \subseteq B \subseteq A$, then for any $X \subseteq U$,,*

$$Bel_C^{\succeq}(X) \leq Bel_B^{\succeq}(X) \leq \frac{|X|}{|U|} \leq Pl_B^{\succeq}(X) \leq Pl_C^{\succeq}(X).$$

Example 3.1 From Example 2.1, for $X = \{x_3, x_5, x_6\}$, we have got

$$
\begin{aligned}
&\overline{R_A^{\succeq}}(X) = \{x_2, x_3, x_4, x_5, x_6\}, \\
&\underline{R_A^{\succeq}}(X) = \{x_5, x_6\}.
\end{aligned}
$$

So, we can calculate

$$
\begin{aligned}
Bel_A^{\succeq}(X) &= \frac{|\underline{R_A^{\succeq}}(X)|}{|U|} = \frac{1}{3}, \\
Pl_A^{\succeq}(X) &= \frac{|\overline{R_A^{\succeq}}(X)|}{|U|} = \frac{5}{6}.
\end{aligned}
$$

Another, if let $B = \{a_1\} \subseteq A$, we have

$$[x_1]_B^{\succeq} = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$[x_2]_B^{\succeq} = \{x_2, x_5, x_6\},$$
$$[x_3]_B^{\succeq} = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$[x_4]_B^{\succeq} = \{x_2, x_4, x_5, x_6\},$$
$$[x_5]_B^{\succeq} = \{x_2, x_5, x_6\},$$
$$[x_6]_B^{\succeq} = \{x_2, x_5, x_6\},$$

and

$$[x_1]_B^{\preceq} = \{x_1, x_3\},$$
$$[x_2]_B^{\preceq} = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$[x_3]_B^{\preceq} = \{x_1, x_3\},$$
$$[x_4]_B^{\preceq} = \{x_2, x_3, x_4\},$$
$$[x_5]_B^{\preceq} = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$[x_6]_B^{\preceq} = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$

So

$$\overline{R_B^{\succeq}}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$\underline{R_B^{\succeq}}(X) = \phi.$$

Thus, we get

$$Bel_B^{\succeq}(X) = \frac{|\underline{R_B^{\succeq}}(X)|}{|U|} = 0,$$
$$Pl_B^{\succeq}(X) = \frac{|\overline{R_B^{\succeq}}(X)|}{|U|} = 1.$$

Hence, the following is obvious.

$$Bel_B^{\succeq}(X) \leq Bel_A^{\succeq}(X) \leq \frac{|X|}{|U|} \leq Pl_A^{\succeq}(X) \leq Pl_B^{\succeq}(X).$$

4 Attributes reduction in ordered information systems

In this section, we discuss the attribute reduction in ordered information systems by proposing the concepts of belief and plausibility reductions in ordered information systems, and compare them with the existing classical reduction.

Definition 4.1 Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system, then

(1) an attribute subset $B \subseteq A$ is referred to as a classical consistent set of $\mathcal{I}^{\succeq}$ if $R_B^{\succeq} = R_A^{\succeq}$. Moreover, if B is a classical consistent set of $\mathcal{I}^{\succeq}$ and no proper subset of B is a classical consistent set of $\mathcal{I}^{\succeq}$, then B is referred to as a classical reduct of $\mathcal{I}^{\succeq}$.

(2) an attribute subset $B \subseteq A$ is referred to as a belief consistent set of $\mathcal{I}^{\succeq}$ if $Bel_B^{\succeq}(X) = Bel_A^{\succeq}(X)$ for any $X \in U/R_A^{\succeq}$. Moreover, if B is a belief consistent set of $\mathcal{I}^{\succeq}$ and no proper subset of B is a belief consistent set of $\mathcal{I}^{\succeq}$, then B is referred to as a belief reduct of $\mathcal{I}^{\succeq}$.

(3) an attribute subset $B \subseteq A$ is referred to as a plausibility consistent set of $\mathcal{I}^{\succeq}$ if $Pl_B^{\succeq}(X) = Pl_A^{\succeq}(X)$ for any $X \in U/R_A^{\succeq}$. Moreover, if B is a plausibility consistent set of $\mathcal{I}^{\succeq}$ and no proper subset of B is a plausibility consistent set of $\mathcal{I}^{\succeq}$, then B is referred to as a plausibility reduct of $\mathcal{I}^{\succeq}$.

Theorem 4.1 *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system and $B \subseteq A$. Then the following holds.*

(1) *B is a classical consistent set of $\mathcal{I}^{\succeq}$ if and only if B is a belief consistent set of $\mathcal{I}^{\succeq}$.*
(2) *B is a classical reduction of $\mathcal{I}^{\succeq}$ if and only if B is a belief reduction of $\mathcal{I}^{\succeq}$.*

Proof (1) Assume that B is a classical consistent set of $\mathcal{I}^{\succeq}$. For any $X \in U/R_A^{\succeq}$, since $[x]_B^{\succeq} = [x]_A^{\succeq}$ for all $x \in U$, we can have

$$[x]_B^{\succeq} \subseteq X \Longleftrightarrow [x]_A^{\succeq} \subseteq X.$$

Then by the definition of lower approximation we have

$$x \in \underline{R_B^{\succeq}}(X) \Longleftrightarrow x \in \underline{R_A^{\succeq}}(X), \quad x \in U.$$

Hence $\underline{R_B^{\succeq}}(X) = \underline{R_A^{\succeq}}(X)$ for any $X \in U/R_A^{\succeq}$. By Theorem 3.2, it follows that $Bel_B^{\succeq}(X) = Bel_A^{\succeq}(X)$ for any $X \in U/R_A^{\succeq}$.

Thus B is a belief consistent of $\mathcal{I}^{\succeq}$

Conversely, if B is a belief consistent set of $\mathcal{I}^{\succeq}$, that is,

$$Bel_B^{\succeq}(X) = Bel_A^{\succeq}(X), \quad \text{for any } X \in U/R_A^{\succeq};$$

i.e.,

$$Bel_B^{\succeq}([x]_A^{\succeq}) = Bel_A^{\succeq}([x]_A^{\succeq}), \quad \text{for any } x \in U.$$

Then for any $x \in U$ we have

$$\frac{|\underline{R_A^{\succeq}}([x]_A^{\succeq})|}{|U|} = \frac{|\underline{R_B^{\succeq}}([x]_A^{\succeq})|}{|U|}.$$

By Corollary 3.1 and Proposition 2.3, we obtain

$$\underline{R_A^{\succeq}}([x]_A^{\succeq}) = \underline{R_B^{\succeq}}([x]_A^{\succeq}), \qquad \text{for all } x \in U.$$

So by the definition of lower approximation we have

$$\left\{y|[y]_A^{\succeq} \subseteq [x]_A^{\succeq}\right\} = \left\{y|[y]_B^{\succeq} \subseteq [x]_A^{\succeq}\right\}, \quad \text{for all } x \in U.$$

That is to say

$$[y]_A^{\succeq} \subseteq [x]_A^{\succeq} \Longleftrightarrow [y]_B^{\succeq} \subseteq [x]_A^{\succeq}, \qquad \text{for all } x, y \in U. \qquad (*)$$

In Eq. (*), we let $y = x$, then obtain $[x]_A^{\succeq} \subseteq [x]_A^{\succeq} \Longleftrightarrow [x]_B^{\succeq} \subseteq [x]_A^{\succeq}$. Hence, we have $[x]_B^{\succeq} \subseteq [x]_A^{\succeq}$ for all $x \in U$. Therefore, by Proposition 2.1 we conclude that $[x]_B^{\succeq} = [x]_A^{\succeq}$ for any $x \in U$.

Thus B is a classical consistent set of $\mathcal{I}^{\succeq}$.

(2) It follows immediately from (1).
Thus the proof is completed. □

Definition 4.2 Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system and $U/R_A^{\succeq} = \{C_1^{\succeq}, C_2^{\succeq}, \ldots, C_t^{\succeq}\}$, denote

$$M = \sum_{i=1}^{t} Bel_A^{\succeq}(C_i^{\succeq}).$$

Then M is referred to as belief sum of $\mathcal{I}^{\succeq}$.

By above definition, we can have the following theorem.

Theorem 4.2 *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system and $B \subseteq A$. Then the following holds.*

(1) *B is classical consistent set of $\mathcal{I}^{\succeq}$ if and only if $\sum_{i=1}^{t} Bel_B^{\succeq}(C_i^{\succeq}) = M$.*
(2) *B is a classical reduction of $\mathcal{I}^{\succeq}$ if and only if $\sum_{i=1}^{t} Bel_B^{\succeq}(C_i^{\succeq}) = M$, and for any nonempty proper subset $B' \subset B$, $\sum_{i=1}^{t} Bel_{B'}^{\succeq}(C_i^{\succeq}) < M$ is true.*

Proof (1) By the Theorem 4.1, we know that B is a classical consistent set of $\mathcal{I}^{\succeq}$ if and only if B is a belief consistent set of $\mathcal{I}^{\succeq}$. Thus B is classical consistent set of $\mathcal{I}^{\succeq}$ if and only if $\sum_{i=1}^{t} Bel_B^{\succeq}(C_i^{\succeq}) = \sum_{i=1}^{t} Bel_A^{\succeq}(C_i^{\succeq})$. That is to say B is classical consistent set of $\mathcal{I}^{\succeq}$ if and only if $\sum_{i=1}^{t} Bel_B^{\succeq}(C_i^{\succeq}) = M$.

(2) It can be obtained from (1) and Definition 4.1. □

Example 4.1 Let consider the system in Example 2.1. Denote

$$\begin{aligned}
C_1^{\succeq} &= [x_1]_A^{\succeq} = \{x_1, x_2, x_5, x_6\};\\
C_2^{\succeq} &= [x_2]_A^{\succeq} = \{x_2, x_5, x_6\};\\
C_3^{\succeq} &= [x_3]_A^{\succeq} = \{x_2, x_3, x_4, x_5, x_6\};\\
C_4^{\succeq} &= [x_4]_A^{\succeq} = \{x_4, x_6\};\\
C_5^{\succeq} &= [x_5]_A^{\succeq} = \{x_5\};\\
C_6^{\succeq} &= [x_6]_A^{\succeq} = \{x_6\}.
\end{aligned}$$

So it can be calculated that

$$M = \sum_{i=1}^{6} Bel_A^{\succeq}(C_i^{\succeq}) = \frac{16}{6}.$$

Let $B = \{a_2, a_3\}$, we have

$$\begin{aligned}
[x_1]_B^{\succeq} &= \{x_1, x_2, x_5, x_6\},\\
[x_2]_B^{\succeq} &= \{x_2, x_5, x_6\},\\
[x_3]_B^{\succeq} &= \{x_2, x_3, x_4, x_5, x_6\},\\
[x_4]_B^{\succeq} &= \{x_4, x_6\},\\
[x_5]_B^{\succeq} &= \{x_5\},\\
[x_6]_B^{\succeq} &= \{x_6\},
\end{aligned}$$

Table 2 Another ordered information system

U	a_1	a_2	a_3
x_1	1	3	1
x_2	2	1	2
x_3	3	2	3
x_4	1	2	2

So

$$\sum_{i=1}^{6} Bel_B^{\succeq}(C_i^{\succeq}) = \sum_{i=1}^{6} \frac{|\underline{R_B^{\succeq}}(C_i^{\succeq})|}{|U|} = \frac{16}{6}.$$

On the other hand, it can be computed that

$$\sum_{i=1}^{6} Bel_{\{a_2\}}^{\succeq}(C_i^{\succeq}) = \frac{7}{6}, \quad \sum_{i=1}^{6} Bel_{\{a_3\}}^{\succeq}(C_i^{\succeq}) = \frac{5}{6},$$
$$\sum_{i=1}^{6} Bel_{\{a_1,a_2\}}^{\succeq}(C_i^{\succeq}) = \frac{12}{6}, \quad \sum_{i=1}^{6} Bel_{\{a_1,a_3\}}^{\succeq}(C_i^{\succeq}) = \frac{12}{6}.$$

Hence, from the above and Theorem 4.2 we can see that $\{a_2, a_3\}$ is the unique belief reduction of $\mathcal{I}^{\succeq}$. It can also be calculated by the discernibility matrix method [27] that the system has the unique classical reduction $\{a_2, a_3\}$.

Theorem 4.3 *Let $\mathcal{I}^{\succeq} = (U, A, F)$ be an ordered information system and $B \subseteq A$. If B is a classical consistent set of $\mathcal{I}^{\succeq}$, then B is a plausibility consistent set of $\mathcal{I}^{\succeq}$.*

Proof Assume that B is a classical consistent set of $\mathcal{I}^{\succeq}$. For any $X \in U/R_A^{\succeq}$, since $[x]_B^{\succeq} = [x]_A^{\succeq}$ for all $x \in U$, we can have $[x]_B^{\preceq} = [x]_A^{\preceq}$, Thus, we know

$$[x]_B^{\preceq} \cap X \neq \phi \Longleftrightarrow [x]_A^{\preceq} \cap X \neq \phi.$$

Then by the definition of upper approximation we have

$$x \in \overline{R_B^{\succeq}}(X) \Longleftrightarrow x \in \overline{R_A^{\succeq}}(X), \quad x \in U.$$

Hence $\overline{R_B^{\succeq}}(X) = \overline{R_A^{\succeq}}(X)$ for any $X \in U/R_A^{\succeq}$. By Theorem 3.2, it follows that $Pl_B^{\succeq}(X) = Pl_A^{\succeq}(X)$ for any $X \in U/R_A^{\succeq}$.

Thus B is a plausibility consistent of $\mathcal{I}^{\succeq}$.

The proof is completed. □

This theorem shown the classical consistent set is the plausibility consistent set in the ordered information systems. However, the reversion of Theorem 4.3 does not hold. And we can show this fact by the following example.

Example 4.2 Table 2 gives another ordered information system, let consider the system.

If we denote $A = \{a_1, a_2, a_3\}$ and $B = \{a_2, a_3\}$, then from Table 2 we have

$$C_1^{\succeq} = [x_1]_A^{\succeq} = \{x_1\};$$
$$C_2^{\succeq} = [x_2]_A^{\succeq} = \{x_2, x_3\};$$
$$C_3^{\succeq} = [x_3]_A^{\succeq} = \{x_3\};$$
$$C_4^{\succeq} = [x_4]_A^{\succeq} = \{x_3, x_4\}.$$

Thus we can obtain

$$[x_1]_A^{\preceq} = \{x_1, x_2, x_3, x_4\};$$
$$[x_2]_A^{\preceq} = \{x_1, x_2, x_4\};$$
$$[x_3]_A^{\preceq} = \{x_1, x_2, x_3, x_4\};$$
$$[x_4]_A^{\preceq} = \{x_1, x_2, x_4\};$$

and

$$\overline{R_A^{\succeq}}(C_1^{\succeq}) = \{x_1, x_2, x_3, x_4\};$$
$$\overline{R_A^{\succeq}}(C_2^{\succeq}) = \{x_1, x_2, x_3, x_4\};$$
$$\overline{R_A^{\succeq}}(C_3^{\succeq}) = \{x_1, x_3\};$$
$$\overline{R_A^{\succeq}}(C_4^{\succeq}) = \{x_1, x_2, x_3, x_4\}.$$

So

$$Pl_A^{\succeq}(C_1^{\succeq}) = 1, \quad Pl_A^{\succeq}(C_2^{\succeq}) = 1,$$
$$Pl_A^{\succeq}(C_3^{\succeq}) = \frac{1}{2}, \quad Pl_A^{\succeq}(C_4^{\succeq}) = 1.$$

On the other hand, we have

$$[x_1]_B^{\succeq} = \{x_1\};$$
$$[x_2]_B^{\succeq} = \{x_2, x_3, x_4\};$$
$$[x_3]_B^{\succeq} = \{x_3\};$$
$$[x_4]_B^{\succeq} = \{x_3, x_4\}.$$

Thus we can calculate

$$Pl_B^{\succeq}(C_1^{\succeq}) = 1, \quad Pl_B^{\succeq}(C_2^{\succeq}) = 1,$$
$$Pl_B^{\succeq}(C_3^{\succeq}) = \frac{1}{2}, \quad Pl_B^{\succeq}(C_4^{\succeq}) = 1.$$

Hence, we can see that $B = \{a_2, a_3\}$ is a plausibility consistent set of the system. But it is not a classical consistent set of the system, because $R_B^{\succeq} \neq R_A^{\succeq}$.

5 Conclusion

It is well-known that rough set theory has been regarded as a generalization of the classical set theory in one way. Furthermore, this is an important mathematical tool to deal with uncertainty. In this article, we have introduced the notions of belief and plausibility reductions in ordered information system and examined the relationships between new concepts of reduct and the classical reduct. Moreover, we have proved that an attribute set in ordered information system is a classical reduct if and only if it is a belief reduct. Though an attribute set in classical information system is a belief reduction if and only if it is a plausibility reduct [25], we have shown that the belief reduct and plausibility reduct in ordered information system are different concepts. In this paper, we only discussed the issue of attribute reduction by the theory of evidence in ordered information system without decision. We will investigate their application for knowledge acquisition in the form of rule induction in our further study.

Acknowledgments The authors would like to thank the anonymous referees for their valuable comments and suggestions. The project is supported by Science and Technology Program of Jiulongpo District, Chongqing (No. 2008Q98), and Science and Technology Program of Board of Education of Chongqing(KJ090612).

References

1. Beynon M (2001) Reducts with in the variable precision rough set model: a further investigation. Eur J Oper Res 134:592–605
2. Dembczynski K, Pindur R, Susmaga R (2003) Generation of exhaustive set of rules within Dominance-based rough set approach. Electronic Notes Theory Computer Science 82(4):96–107
3. Dembczynski K, Pindur R, Susmaga R (2003) Dominance-based rough set classifier without induction of decision rules. Electronic Notes Theory Computer Science 82(4):84–95
4. Dempster AP (1967) Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 38:325–339
5. Greco S, Matarazzo B, Slowinski R (1999) Rough approximation of a preference relation by dominance relation. ICS Research Report 16/96, Warsaw University of Technology; 1996 and in Eur J Oper Res 117:63–83
6. Greco S, Matarazzo B, Slowinski R (1998) A new rough set approach to multicriteria and multiattribute classification. In: Polkowsik L, Skowron A (eds) Rough sets and current trends in computing (RSCTC'98). Lecture notes in Artificial Intelligence, vol 1424. Springer, Berlin, pp 60–67
7. Greco S, Matarazzo B, Slowinski R (1999) A new rough sets approach to evaluation of bankruptcy risk. In: Zopounidis X, (ed) Operational tools in the management of financial risks. Kluwer, Dordrecht, pp 121–136
8. Greco S, Matarazzo B, Slowinski R (2001) Rough set theory for multicriteria decision analysis. Eur J Oper Res 129:11–47
9. Greco S, Matarazzo B, Slowinski R (2002) Rough sets methodology for sorting problems in presence of multiple attributes and criteria. Eur J Oper Res 138:247–259
10. Greco S, Matarazzo B, Slowinski R (2007) Dominance-based Rough Set approach as a proper way of handling graduality in rough set theory. Transactions on Rough Sets VII. Lecture notes in computer science, 4400:36–52
11. Guo D, Xiong H, Atluri V, Adam NR (2009) Object discovery in high-resolution remote sensing images: a semantic perspective. Knowl Inform Syst 19(5):211–233
12. Intan R, Mukaidono M (2004) Fuzzy conditional probability relations and their applications in fuzzy information systems. Knowl Inform Syst 6:345–365
13. Kryszkiewicz M (1998) Rough set approach to incomplete information system. Inform Sci 112:39–49
14. Kryszkiewicz M (2001) Comparative studies of alternative type of knowledge reduction in inconsistent systems. Int J Intell Syst 16:105–120
15. Leuang Y, Wu WZ, Zhang WX (2006) Knowledge acquisition in incomplete information systems: a rough set approach. Eur J Oper Res 168(1):164–180
16. Pawlak Z (1982) Rough sets. Int J Comput Inform Sci 11(5):341–356
17. Qiu L, Li YJ, Wu XT (2008) Protecting business intelligence and customer privacy while outsourcing data mining tasks. Knowl Inform Syst 17(10):99–120
18. Resconi G, Kovalerchuk B (2009) Agents model of uncertainty. Knowl Inform Syst 18(2):213–229
19. Sai Y, Yao YY, Zhong N (2001) Data analysis and mining in ordered information tables. In: Proceeding of the 2001 IEEE international conference on data mining. IEEE Computer Society Press, pp 497–504
20. Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton

21. Shao MW, Zhang WX (2005) Dominance relation and relus in an incomplete ordered information system. Int J Intell Syst 20:13–27
22. Skowron A, Graymala-Busse J (1994) From rough set theory to evidence theory. In: Yager RR, Fedrizzi M, Kacprzyk J (eds) Advance in the Dempster–Shafer theory of evidence. Wiley, New York, pp 193–236
23. Slezak D (1998) Searching for dynamic reducts in inconsistent decision tables. In: Proceedings of IPMU'98, Paris, France 2:1362–1369
24. Susmaga R, Slowinski R, Greco S, Matarazzo B (2000) Generation of reducts and rules in multi-attribute and multi-criteria classification. Control Cybernet 4:969–988
25. Starzyk JA, Nelson DE, Sturtz K (2000) A mathematical foundation for improved reduct generation in information systems. Knowl Inform Syst 2:131–146
26. Wang KJ, Zhang JY, Shen FS, Shi LF (2008) Adaptive learning of dynamic Bayesian networks with changing structures by detecting geometric structures of time series. Knowl Inform Syst 17(10):121–133
27. Wu QX, Bell D, McGinnity M (2005) Multiknowledge for decisioin making. Knowl Inform Syst 7:246–266
28. Wu WZ, Leung Y, Zhang WX (2002) Connections between rough set theory and Dempster–Shafer theory of evidence. Int J Gen Syst 31:405–430
29. Wu WZ, Zhang M, Li HZ, Mi JS (2005) Knowledge reduction in random information systems via Dempster-Shafer theory of evidence. Inform Sci 174:143–164
30. Xu WH, Zhang WX (2007) Measuring roughness of generalized rough sets induced by a covering. Fuzzy Sets Syst 158:2443–2455
31. Xu WH, Shao MW, Zhang WX (2006) Knowledge reduction based on evidence reasoning theory in ordered information systems. Lect Notes Artif Intell 4092:535–547
32. Yao YY, Lingras P (1998) Interpretations of belief functions in the theory of rough sets. Inform Sci 104:81–106
33. Zhang M, Xu LD, Zhang WX, Li HZ (2003) A rough set approach to knowledge reduction based on inclusion degree and evidence reasoning theory. Expert Syst 20:298–304
34. Zhang WX, Mi JS, Wu WZ (2003) Approaches to knowledge reductions in inconsistent Systems. Int J Intell Syst 21(9):989–1000
35. Zhang WX, Wu WZ, Liang JY, Li DY (2001) Theory and method of rough sets. Science Press, Beijing

Author biographies

Wei-hua Xu is a Professor and Master Supervisor at Chongqing University of Technology, China. He received a PhD from Xi'an Jiaotong University, China in 2007. He has published over 20 journal papers. His main research fields are rough sets, fuzzy sets and the mathematical foundation of information science.

Xiao-yan Zhang is a teacher at Chongqing University of Technology, China. She received a Master Degree from Guangxi University, China in 2005. She has published over 10 journal papers. Her main research fields are rough sets, fuzzy sets and topological dynamical system.

Jian-min Zhong is a Professor at Chongqing University of Technology, China. He has published over 10 journal papers. His main research fields are information science and numerical analysis.

Wen-xiu Zhang is a Professor and Doctoral Supervisor at Xi'an Jiaotong University, China. He graduated from Nankai University, China in 1967. He is a member of the Standing Council of Mathematics Society of China and of IFSA. He is Vice-Chairman of the Editorial Board of the *Journal of Fuzzy Systems and Mathematics*, and Vice-Chairman of the Editorial Board of the *Journal of Engineering Mathematics*. He has published over 140 journal papers and 20 books. His research specialisation covers fuzzy sets, random sets, rough sets, genetic algorithms and the mathematical foundation of information science.