Annals of **F**uzzy **M**athematics and **I**nformatics
Volume 4, No. 2, (October 2012), pp. 267–279
ISSN 2093–9310
http://www.afmi.or.kr

Attribute reduction in tolerance information systems based on evidence theory

Weihua Xu, Wenxin Sun, Qiaorong Wang, Yufeng Liu

Received 9 November 2011; Revised 28 December 2011; Accepted 10 January 2012

Abstract. Attribute reduction is one of the most important problems in rough set theory. This paper deals with attribute reduction in tolerance information systems based on Dempster-shafter theory of evidence. The concepts of plausibility consistent set and belief consistent set are introduced in tolerance information systems. Furthermore, relative plausibility reduction and belief reduction are discussed in tolerance systems and it is proved that a plausibility consistent set must be a consistent set. Moreover, it is shown that an attribute set is a belief reduction if and only if it is a classical reduction in tolerance information systems.

2010 AMS Classification: 94D05, 68T37

Keywords: Rough set, Attribute reduction, Tolerance relation, Evidence theory

Corresponding Author: Weihua Xu (chxuwh@gmail.com.)

1. Introduction

Rough set theory, proposed by Pawlak [11, 12, 13], has become a well-established mechanism for uncertainty management in a wide variety of applications related to artificial intelligence [5, 6, 7, 10]. The theory has been found its successful applications in the fields of pattern recognition, medical diagnosis, data mining, conflict analysis, algebra [1, 3, 14, 15, 16], which need to deal with an amount of imprecise, vague and uncertain information. In recent years, the rough set theory has generated a great deal of interest among more and more researchers.

However, in practice, due to the existence of uncertainty and complexity of particular problems, the problem would not be settled perfectly by means of classical rough sets. Therefore, it is vital to generalize the classical rough set model. To overcome this limitation, classical rough sets have been extended to several interesting and meaningful general models in recent years by proposing other binary relations, such as tolerance relations [27].

Another important method is 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 probability[4], and extended by Shafer as a theory[19]. 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 pair of lower and upper approximation operators induced by rough approximation spaces may be used to interpret the corresponding pairs of belief and plausibility functions induced by belief structures[20, 22].

It is well known that not all condition attributes in an information system are necessary. Knowledge reduction in the sense of reducing attributes is thus an outstanding contribution made by rough set research to data analysis[2, 9, 17, 18, 21, 23, 26]. In recent years, more attention has been paid to attribute reduction in tolerance information systems in rough set research[8].

In the next section, we give some basic notions related to tolerance information systems. We also review rough set approximations in tolerance information systems. Some basic notions of evidence theory are introduced in Section 3. The concepts of belief reduction and plausibility reduction in tolerance information systems are proposed and the relationships between the new concepts of reductions and the classical reduction are examined in Section 4. We then conclude the paper with a summary and out-look for further research in Section 5.

2. Rough set and tolerance information systems

In this section, we recall some necessary notions and preliminaries required in the sequel of our work. Detailed description of these theories can be found in the literature [28].

The concept 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 $I = (U, A, F)$, where U is a non-empty finite universe and A is a finite and non-empty set of attributes, such that there exists a map $f_l : U \mapsto V_{a_l}$ for any $a_l \in A$, where V_{a_l} is called the domain of the attribute a_l, and denoted $F = \{f_l | a_l \in A\}$.

If a binary relation R on the universe U is reflexive and symmetric, it is called a tolerance relation on U. The set of all tolerance relations on U is denoted by $\mathcal{R}$. Obviously, tolerance relation $R \in \mathcal{R}$ can construct a covering of the universe U[8].

Definition 2.1 ([24])**.** A tolerance information system is a triple $S = (U, A, \tau)$, where U is the non-empty finite object set known as universe; A is the non-empty finite set of attributes; the mapping τ is the mapping from 2^A into the family set $\mathcal{R}$ of tolerance relations on U.

Definition 2.2 ([8]). Let $I = (U, A, \tau)$ be a tolerance information system, for $B \subseteq A$, $X \subseteq U$,$R_B \in \mathcal{R}$. Denote

$$[x_i]_B = \{x_j \in U|(x_i, x_j) \in R_B\},$$
$$U/R_B = \{[x_i]_B|x_i \in U\},$$

where $i \in \{1, 2, ..., |U|\}$, then $[x_i]_B$ will be called a tolerance class. And U/R_B is called a classification of U about attribute set B.

X is said to be a maximal tolerance class with B if there does not exist another tolerance class Y with respect to B such that $X \subseteq Y$.

Proposition 2.3 ([8]). *Let R be a tolerance relation, $B \subseteq A$.*

(1) *If $B \subseteq A$, then $R_A \subseteq R_B$.*

(2) *If $B \subseteq A$, then $[x_i]_A \subseteq [x_i]_B$.*

(3) *If $x_j \in [x_i]_B$, then $x_i \in [x_j]_B$.*

(4) *$|[x_i]_B| \geq 1$ for any $x_i \in U$, $B \subseteq A$.*

(5) *U/R_B constitutes a covering of U, i.e., for every $x \in U$ we have that $[x]_B \neq \varnothing$ and $\bigcup\limits_{x\in U} [x]_B = U$.*

Where $|\,.\,|$ denotes cardinality of the set.

Definition 2.4. Let $I = (U, A, \tau)$ be a tolerance information system. The lower approximation and the upper approximation of a set $X \subseteq U$ are respectively defined by

$$\underline{R_A}(X) = \{x \in U : [x]_A \subseteq U\},$$
$$\overline{R_A}(X) = \{x \in U : [x]_A \cap X \neq \varnothing\}.$$

The set $bn_R(x) = \overline{R_A}(X) - \underline{R_A}(X)$ is called the boundary of X. The set $\underline{R_A}(X)$ consists of elements which surely belong to X in view of the knowledge provided by R, while $\overline{R_A}(X)$ consists of elements which possibly belong to X. The boundary is the actual area of uncertainty. It consists of elements whose membership in X can not be determined when R-related objects can not be distinguished from each other.

Some basic properties of approximations are shown in the following.

Proposition 2.5 ([26]). *Let $I = (U, A, \tau)$ be a tolerance information system and R be a tolerance relation on U. If $X, Y \subseteq U$, $\sim X$ is the complement of X. Then the following properties hold.*

(1) $\underline{R_A}(X) \subseteq X \subseteq \overline{R_A}(X)$;

(2) $\underline{R_A}(\varnothing) = \overline{R_A}(\varnothing) = \varnothing, \underline{R_A}(U) = \overline{R_A}(U) = U$;

(3) $\sim \overline{R_A}(X) = \underline{R_A}(\sim X), \overline{R_A}(\sim X) =\sim \underline{R_A}(X)$;

(4) $\underline{R_A}(X \cap Y) = \underline{R_A}(X) \cap \underline{R_A}(Y)$, $\overline{R_A}(X \cup Y) = \overline{R_A}(X) \cup \overline{R_A}(Y)$;

(5) $\underline{R_A}(X \cup Y) \supseteq \underline{R_A}(X) \cup \underline{R_A}(Y)$, $\overline{R_A}(X \cap Y) \subseteq \overline{R_A}(X) \cap \overline{R_A}(Y)$;

(6) *If $X \subseteq Y$, then $\underline{R_A}(X) \subseteq \underline{R_A}(Y)$ and $\overline{R_A}(X) \subseteq \overline{R_A}(Y)$;*

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

Example 2.6. Suppose $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, A consists of three attributes called a, b, c respectively. Their attribute values are given in Table 1. Let us define a mapping $\tau : 2^A \mapsto \mathcal{R}$(where $\mathcal{R}$ is a family of tolerance relations), $\forall B \in 2^A, R \in \mathcal{R}$

$$R = \{(x, y)||f_{l_i}(x) - f_{l_i}(y)| \leq 1\}, \forall l_i \in B.$$

Table 1.

U	a	b	c
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

From Table 1 we have

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

Thus, it is obvious that $U/R_A = \{[x_1]_A, [x_2]_A, [x_3]_A, [x_4]_A, [x_5]_A, [x_6]_A\}$.

And if we let $X = \{x_1, x_3, x_4\}$, then

$$\begin{aligned}
\underline{R_A}(X) &= \{x_1, x_3\};\\
\overline{R_A}(X) &= \{x_1, x_2, x_3, x_4, x_6\}.
\end{aligned}$$

It is clear that

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

3. Evidence theory in tolerance information systems

In evidence theory[4, 19], let U be a non-empty finite universe of discourse, a set function $m : 2^U \mapsto [0,1]$(where 2^U is the classes of all subset of the U)is referred to as a basic probability assignment if it satisfies :

$$\begin{aligned}
&(M1) \quad m(\varnothing) = 0\\
&(M2) \quad \sum_{X \in U/R_A} m(X) = 1
\end{aligned}$$

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 set $X \subseteq U$ with nonzero basic probability assignment is referred to as a focal element.

We denoted the family of all focal elements of m by μ . The pair (μ, m) is called a belief structure or a body of evidence. Associated with each belief structure in information systems based on classical equivalence relation, a pair of belief and plausibility functions can be derived.

Definition 3.1 ([4, 19])**.** Let (μ, m) be a belief structure. A set function $Bel : 2^U \mapsto [0,1]$ is referred to as a belief function on U, if

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

A belief function $Bel : 2^U \mapsto [0,1]$ can be equivalently defined by axioms, if Bel is a belief function iff it satisfies the axioms:
(1) $Bel(\varnothing) = 0$,
(2) $Bel(U) = 1$,
(3) For every collection of subsets $X_1, X_2, ..., X_n \subseteq U$,

$$Bel(\bigcup_{i=1}^{n} X_i) \geq \sum_{\varnothing \neq J \subseteq \{1,2,...,n\}} (-1)^{|J|+1} Bel(\bigcap_{i \in J} X_i).$$

Where $|J|$ is the cardinality of the set J.

Definition 3.2 ([4, 19])**.** A set function $Pl : 2^U \mapsto [0,1]$ is referred to as a plausibility function on U, if

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

We can define a relational partition function or a basic set assignment.

Definition 3.3. Let $I = (U, A, \tau)$ be a tolerance information system. We denote

$$j(X) = \{x \in U | [x]_A = X\}.$$

for any $X \in U/R_A$.

It is easy to verify that j satisfies the properties
(1) $j(\varnothing) = \varnothing$;
(2) $\bigcup_{Y \subseteq U} j(Y) = U$;
(3) $A \neq B \Rightarrow j(A) \cap j(B) = \varnothing$;
(4) $\underline{R}(X) = \bigcup_{Y \subseteq X} j(Y), \overline{R}(X) = \bigcup_{Y \cap U \neq \varnothing} j(Y)$.
Then a mass function of I can be defined by a map $m : U/R_A \to [0,1]$, where

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

By above definition, we can easily find out that a mass function of a tolerance information system still satisfies two basic axioms.

Definition 3.4. Let $I = (U, A, \tau)$ be a tolerance information system. A set function $Bel : 2^U \mapsto [0,1]$ is referred to as a belief function on U, if

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

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

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

There are strong connections between rough set theory and the evidence theory. The following theorem shows that the classical belief and plausibility functions can be interpreted in terms of the Pawlak's lower and upper approximation of sets[21].

Theorem 3.5 ([25]). *Let $I = (U, A, F)$ be an information system. For any $X \subseteq U$, $B \subseteq A$, denote*

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

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

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

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

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

Theorem 3.6. *Let $I = (U, A, \tau)$ be a tolerance information system. For any $X \subseteq U$, $B \subseteq A$, denote*

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

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

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

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

Proof. From Proposition 2.5 we can see that

(1) $\dfrac{|\underline{R_B}(\varnothing)|}{|U|} = 0;$

(2) $\dfrac{|\overline{R_B}(U)|}{|U|} = 1,$

which satisfy (1) and (2) of the axioms. Consider a collection $X_1, X_2, ..., X_n \subseteq U$, we have

$$\frac{|\underline{R_B}(X_1 \cup X_2 \cup ... \cup X_n)|}{|U|}$$

$$\geq \frac{|\underline{R_B}(X_1) \cup \underline{R_B}(X_2) \cup ... \cup \underline{R_B}(X_n)|}{|U|}$$

$$= \sum_i \frac{|\underline{R_B}(X_i)|}{|U|} - \sum_{i<j} \frac{|\underline{R_B}(X_i) \cap \underline{R_B}(X_j)|}{|U|} + ...+$$

$$(-1)^{n+1} \frac{|\underline{R_B}(X_1) \cap \underline{R_B}(X_2) \cap ... \cap \underline{R_B}(X_n)|}{|U|}$$

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

Thus, we proved that $Bel_B(X)$ is the belief function of U. And one can obtain directly that $Pl_B(X)$ is the plausibility function since the duality between $Bel_B(X)$ and $Pl_B(X)$. So the theorem was proved. □

Corollary 3.7. *Let $I = (U, A, \tau)$ be a tolerance information system and $C \subseteq B \subseteq A$. For any $X \subseteq U$,*

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

Example 3.8. From Example 2.6, for $X = \{x_1, x_3, x_4\}$, we have got

$$\underline{R_A}(X) = \{x_1, x_3\}, \qquad \overline{R_A}(X) = \{x_1, x_2, x_3, x_4, x_6\}.$$

So we calculate

$$Bel_A(X) = \frac{|\underline{R_A}(X)|}{|U|} = \frac{2}{6},$$

$$Pl_A(X) = \frac{|\overline{R_A}(X)|}{|U|} = \frac{5}{6}.$$

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

$$[x_1]_B = \{x_1, x_3, x_4\};$$
$$[x_2]_B = \{x_2, x_4, x_5, x_6\};$$
$$[x_3]_B = \{x_1, x_3, x_4\};$$
$$[x_4]_B = \{x_1, x_2, x_3, x_4, x_5, x_6\};$$
$$[x_5]_B = \{x_2, x_4, x_5, x_6\};$$
$$[x_6]_B = \{x_2, x_4, x_5, x_6\}.$$

So

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

Thus, we get

$$Bel_B(X) = \frac{|\underline{R_B}(X)|}{|U|} = \frac{2}{6};$$

$$Pl_B(X) = \frac{|\overline{R_B}(X)|}{|U|} = 1.$$

Hence, the following is obvious.

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

4. Attribute reduction in tolerance information systems

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

Definition 4.1. Let $I = (U, A, \tau)$ be a tolerance information system. Then

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

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

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

Theorem 4.2. *Let $I = (U, A, \tau)$ be a tolerance information system and $B \subseteq A$. Then the following holds.*

(1) B is a classical consistent set of I if and only if B is a belief consistent set of I.

(2) B is a classical reduction of I if and only if B is a belief reduction of I.

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

$$[x]_B \subseteq X \Leftrightarrow [x]_A \subseteq X.$$

Then by the definition of lower approximation, we can have

$$x \in \underline{R_B}(X) \Leftrightarrow x \in \underline{R_A}(X), x \in U.$$

Hence $\underline{R_B}(X) = \underline{R_A}(X)$ for any $X \in U/R_A$. By Theorem 3.6, it follows that $Bel_B(X) = Bel_A(X)$ for any $X \in U/R_A$.

Thus B is a belief consistent set of I.

Conversely, if B is a belief consistent set of I, that is,

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

i.e.,

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

Then for any $x \in U$ we have

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

By Corollary 3.7 and Proposition 2.5, we obtain

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

So by the definition of lower approximation we have

$$\{y|[y]_A \subseteq [x]_A\} = \{y|[y]_B \subseteq [x]_A\}, \text{ for any } x, y \in U.$$

That is to say

$$[y]_A \subseteq [x]_A \Leftrightarrow [y]_B \subseteq [x]_A, \text{ for any } x, y \in U.$$

By above equation, we let $y = x$, then obtain $[x]_A \subseteq [x]_A \Leftrightarrow [x]_B \subseteq [x]_A$. Hence, we have $[x]_B \subseteq [x]_A$ for all $x \in U$. Therefore, by Proposition 2.3 we conclude that $[x]_B = [x]_A$ for any $x \in U$.

Thus B is a consistent set of I.

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

Theorem 4.3. *Let $I = (U, A, \tau)$ be a tolerance information system and $B \subseteq A$. Denote*

$$U/R_A = \{C_1, C_2, ..., C_n\}, \quad M = \sum_{i=1}^{n} Bel_A(C_i).$$

Then the following holds.

(1) *B is a classical consistent set of I if and only if $\sum\limits_{i=1}^{n} Bel_B(C_i) = M$.*

(2) *B is a classical reduction of I if and only if $\sum\limits_{i=1}^{n} Bel_B(C_i) = M$, and for any nonempty proper subset $B^{'} \subset B$, $\sum\limits_{i=1}^{n} Bel_{B'}(C_i) < M$ is true.*

Proof. (1) By Theorem 4.2, we know that B is a classical consistent set of I if and only if B is a belief consistent set of I. Thus B is classical consistent set of I if and only if $\sum\limits_{i=1}^{n} Bel_B(C_i) = \sum\limits_{i=1}^{n} Bel_A(C_i)$. That is to say B is classical consistent set of I if and only if $\sum\limits_{i=1}^{n} Bel_B(C_i) = M$.

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

Corollary 4.4. *Let $I = (U, A, \tau)$ be a tolerance information system and $B \subseteq A$. Denote*

$$U/R_B = \{C_1, C_2, ..., C_n\}, \quad M = \sum_{i=1}^{n} Bel_B(C_i),$$

then $M \geq 1$.

Example 4.5. Let consider the system in Example 2.6. Denote

$$C_1 = [x_1]_A = \{x_1, x_3\};$$
$$C_2 = [x_2]_A = \{x_2, x_4, x_5, x_6\};$$
$$C_3 = [x_3]_A = \{x_1, x_3, x_4\};$$
$$C_4 = [x_4]_A = \{x_2, x_3, x_4, x_6\};$$
$$C_5 = [x_5]_A = \{x_2, x_5, x_6\};$$
$$C_6 = [x_6]_A = \{x_2, x_4, x_5, x_6\}.$$

So it can be calculated that

$$M = \sum_{i=1}^{6} Bel_A(C_i) = \frac{11}{6}.$$

On the other hand, it can be computed that

$$\begin{aligned}
Bel_{\{a\}}(C_i) &= \frac{19}{6};\\
Bel_{\{b\}}(C_i) &= \frac{22}{6};\\
Bel_{\{c\}}(C_i) &= \frac{23}{6};\\
Bel_{\{a,b\}}(C_i) &= \frac{14}{6};\\
Bel_{\{a,c\}}(C_i) &= \frac{16}{6};\\
Bel_{\{b,c\}}(C_i) &= \frac{13}{6}.
\end{aligned}$$

Hence, from the above calculation and Theorem 4.3 we can see that I has been a unique belief reduction.

Theorem 4.6. *Let $I = (U, A, \tau)$ be a tolerance information system and $B \subseteq A$. If B is a classical consistent set of I, then B is a plausibility consistent set of I.*

Proof. Assume that B is a classical consistent set of I. For any $X \in U/R_A$, since $[x]_B = [x]_A$ for all $x \in U$, thus we know

$$[x]_B \cap X \neq \varnothing \Leftrightarrow [x]_A \cap X \neq \varnothing.$$

Then by the definition of upper approximation we have

$$x \in \overline{R_B}(X) \Leftrightarrow x \in \overline{R_A}(X), x \in U.$$

Hence $\overline{R_B}(X) = \overline{R_A}(X)$ for any $X \in U/R_A$. By Theorem 3.6, it follows that $Pl_B(X) = Pl_A(X)$ for any $X \in U/R_A$. Thus B is a plausibility consistent of I. The proof was completed. □

The theorem shows the classical consistent set is the plausibility consistent set in tolerance information systems. However, the reversion of Theorem 4.6 does not hold. And we can show this fact by the following example.

Example 4.7. Suppose $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, A consists of two attributes called b, c respectively. Their attribute values are given in Table 2. Let us define a mapping $\tau : 2^A \mapsto \mathcal{R}$(where $\mathcal{R}$ is a family of tolerance relations), $\forall B \in 2^A, R \in \mathcal{R}$

$$R = \{(x, y) || f_{l_i}(x) - f_{l_i}(y)| \leq 1\}, \forall l_i \in B.$$

Table 2.

U	b	c
x_1	2	1
x_2	2	2
x_3	1	2
x_4	1	3
x_5	3	2
x_6	2	3

If we denote $A = \{b, c\}$ and $B = \{b\}$, then from Table 2 we have

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

On the other hand, we have

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

Thus we can obtain

$$\begin{aligned}
\overline{R_B}(C_1) = \overline{R_A}(C_1) &= \{x_1, x_2, x_3, x_4, x_5, x_6\};\\
\overline{R_B}(C_2) = \overline{R_A}(C_2) &= \{x_1, x_2, x_3, x_4, x_5, x_6\};\\
\overline{R_B}(C_3) = \overline{R_A}(C_3) &= \{x_1, x_2, x_3, x_4, x_5, x_6\};\\
\overline{R_B}(C_4) = \overline{R_A}(C_4) &= \{x_1, x_2, x_3, x_4, x_5, x_6\};\\
\overline{R_B}(C_5) = \overline{R_A}(C_5) &= \{x_1, x_2, x_3, x_4, x_5, x_6\};\\
\overline{R_B}(C_6) = \overline{R_A}(C_6) &= \{x_1, x_2, x_3, x_4, x_5, x_6\}.
\end{aligned}$$

So

$$\begin{aligned}
&Pl_A(C_1) = 1, Pl_A(C_2) = 1, Pl_A(C_3) = 1,\\
&Pl_A(C_4) = 1, Pl_A(C_5) = 1, Pl_A(C_6) = 1;\\
&Pl_B(C_1) = 1, Pl_B(C_2) = 1, Pl_B(C_3) = 1,\\
&Pl_B(C_4) = 1, Pl_B(C_5) = 1, Pl_B(C_6) = 1.
\end{aligned}$$

Hence, we can see that B is a plausibility consistent set of the tolerance information system. But it is not a classical consistent set of the tolerance information system, because $R_B \neq R_A$.

5. Conclusion

On the basis of classical rough set theory which defines the lower and upper approximations by using an equivalence relation, some researchers proposed its extended model called tolerance rough set model by using a tolerance relation. We have discussed in this paper attribute reduction via the Dempster-Shafer theory of evidence in tolerance information systems. The concepts of belief reduction and plausibility reduction in tolerance information systems have been introduced and compared with the concepts of classical reduction and relative one, respectively. The results will help us to gain much more insights into the meaning of lower and upper approximations in rough set theory. We can see that the belief and plausibility functions in the Dempster-Shafer theory of evidence may be used to characterize the numeric aspects of uncertainty of rough sets and the evidence theory may provide a useful tool for practical applications of rough set data analysis. In this paper, we only discussed the issue of attribute reduction by the theory of evidence in tolerance information systems without decision. We will investigate their application for knowledge acquisition in the form of rule induction in our further study.

Acknowledgements. This work is supported by National Natural Science Foundation of China (No. 61105041, 71071124 and 11001227), Postdoctoral Science Foundation of China (No. 20100481331) and National Natural Science Foundation of CQCSTC (No. cstc2011jjA 40037).

References

[1] V. S. Ananthanarayana, M. M. Narasimha and D. K. Subramanian, Tree structure for efficient data mining using rough sets, Pattern Recogn. Lett. 24 (2003) 851–862.
[2] M. Beynon, Reducts within the variable precision rough set model:a further investigation, European J. Oper. Res. 134 (2001) 592–605.
[3] W. Cheng, Z. W. Mo and J. Wang, Notes on the lower and upper approximations in a fuzzy group and rough ideals in semigroups, Inform. Sci. 177 (2007) 5134–5140.
[4] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat. 38 (1967) 325–339.
[5] I. Düntsch and G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1998) 109–137.
[6] G. Jeon, D. Kim and J. Jeong, Rough sets attributes reduction based expert system in interlaced video sequences, IEEE Trans. Consumer Electronics 52 (2006) 1348–1355.
[7] R. Jensen and Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE T. Fuzzy Syst. 15 (2007) 73–89.
[8] J. Järiner, Approximations and rough sets based on tolerances, Inform. Sci. 177 (2001) 182–189.
[9] M. Kryszkiewicz, Rough set approach to incomplete information system, Inform. Sci. 112 (1998) 39–49.
[10] J. A. Pomykala, Rough sets and current trends in computing: about tolerance and similarity relations in information systems, Springer, 2002.
[11] Z. Pawlak, Rough sets, Int. J. Comput. Inform. Sci. 11 (1982) 341–356.
[12] Z. Pawlak, Rough set: Theoretical aspects of reasoning about data, Kluwer Academic Publishers, 1991.
[13] Z. Pawlak, Rough sets approach to Multi-attribute decision analysis, European J. Oper. Res. 72 (1994) 443–459.
[14] Z. Pawlak, Rough sets, decision algorithms and Bayes's theorem, Europe J. Oper. Res. 136(1) (2002) 181–189.

[15] Z. Pawlak, Some remarks on conflict analysis, European J. Oper. Res. 166 (2005) 649–654.
[16] Z. Pawlak and A. Skowron, Rudiments of rough sets, Inform. Sci. 177 (2007) 3–27.
[17] Y. H. Qian, J. Y. Liang, W. Pedrycz and C. Y. Dang, An efficient accelerator for attribute reduction from incomplete data in rough set framework, Pattern Recogn. 44 (2011) 1658–1670.
[18] Eric. C. C. Tsang, D.G. Chen, Daniel S. Yeung, X. Z. Wang and John E. T. Lee, Attributes reduction using fuzzy rough sets, IEEE T. Fuzzy Syst. 16 (2008) 1130–1141.
[19] G. Shafer, A mathematical theory of evidence, Princeton University Press, 1976.
[20] W. Z. Wu and W. X. Zhang, Connections between rough set theory and Dempster-Shafer theory of evidence, Int. J. Gen. Syst. 31 (2002) 405–430.
[21] W. Z. Wu, Attribute reduction based on evidence theory in incomplete decision systems, Inform. Sci. 178 (2008) 1355–1371.
[22] W. Z. Wu, M. Zhang, H. Z. Li and J. S. Mi, Knowledge reduction in random information systems via Dempster-Shafer theory of evidence, Inform. Sci. 174 (2005) 143–164.
[23] W. H. Xu, X. Y. Zhang, J. M. Zhong and W. X. Zhang, Attribute reduction in ordered information systems based on evidence theory, Knowledge and Information Systems 25 (2010) 169–184.
[24] B. Z. Xu, X. G. Hu and H. Wang, A general rough set model based on tolerance, International Academic Publishers, 2004.
[25] Y. Y. Yao and P. Lingras, Interpretations of berief functions in the theorey of rough sets, Inform. Sci. 104 (1998) 81–106.
[26] Z. Zheng, H. Hu and Z. Z. Shi, Tolerance relation based granular space, Springer Berlin, 2005.
[27] M. Zhang, L. D. Xu, W. X. Zhang and H. Z. Li, A rough set approach to knowledge reduction based on inclusion degree and evidence reasoning theory, Expert Syst. 20 (2003) 298–304.
[28] W. X. Zhang, W. Z. Wu, J. Y. Liang and D. Y. Li, Theory and methord of rough sets, Science Press, 2001.

Weihua Xu (chxuwh@gmail.com.)
School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, 400054, P. R. China and School of Management, Xi'an Jiaotong University, Xi'an , 710049, P. R. China

Wenxin Sun (sunxuxin520@163.com.)
School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, 400054, P. R. China

Qiaorong Wang (wangqiaoyun@cqut.edu.cn.)
School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, 400054, P. R. China

Yufeng Liu (liuyufeng@cqut.edu.cn.)
School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, 400054, P. R. China