Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 507857, 19 pages
doi:10.1155/2012/507857

Research Article

Consistency of Probability Decision Rules and Its Inference in Probability Decision Table

Zheng Pei,[1] Li Zou,[2] Hamid Reza Karimi,[3] and Peng Shi[4,5]

[1] *School of Mathematics and Computer Engineering, Xihua University, Chengdu, Sichuan 610039, China*
[2] *School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China*
[3] *Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway*
[4] *Department of Computing and Mathematical Sciences, University of Glamorgan, Pontypridd CF37 1DL, UK*
[5] *School of Engineering and Science, Victoria University, Melbourne, VIC 8001, Australia*

Correspondence should be addressed to Hamid Reza Karimi, hamid.r.karimi@uia.no

Received 17 January 2012; Accepted 15 March 2012

Academic Editor: Zhan Shu

In most synthesis evaluation systems and decision-making systems, data are represented by objects and attributes of objects with a degree of belief. Formally, these data can be abstracted by the form (*objects; attributes; P*), where *P* represents a kind degree of belief between objects and attributes, such that, *P* is a basic probability assignment. In the paper, we provide a kind of probability information system to describe these data and then employ rough sets theory to extract probability decision rules. By extension of Dempster-Shafer evidence theory, we can get probabilities of antecedents and conclusion of probability decision rules. Furthermore, we analyze the consistency of probability decision rules. Based on consistency of probability decision rules, we provide an inference method to finish inference of probability decision rules, which can be used to decide the class of a new object x'. The conclusion points out that the inference method of the paper not only deals with precise information, but also imprecise or uncertain information as well.

1. Introduction

Processing uncertain or incomplete information may be placed in the sphere of artificial intelligence. The term "reasoning with uncertain or incomplete information" in the narrow sense means the way of representing a partial information that is available to a user about a fragment of reality and the way of processing such an information. In the broader sense, it is used to denote the interdisciplinary sphere of research concerned with the search for methods of modeling uncertain or incomplete knowledge. Those methods can refer

to any application domain and any level of knowledge; one seeks way of representing both object-level knowledge and meta-level knowledge, the latter being the knowledge about the former [1–5]. Nowadays, the main tools of processing uncertain or incomplete information are fuzzy set theory, probability theory, possibility theory, Dempster-Shafer evidence theory, and rough sets theory. The main advantage of fuzzy set theory is that the fuzzy set framework provides a lot of combination operators, which allows the user to adapt the processing scheme to the specificity of the data at hand [6–17]. The probabilistic theory has solid mathematic theory, but its inference cannot model uncertain measurement [18–20]. Numerical possibility distributions can encode special convex families of probability measures. The connection between possibility theory and probability theory is potentially fruitful in the scope of statistical reasoning, because variability of observations should be distinguished from incomplete information [21, 22]. Dempster-Shafer evidence theory has the ability to deal with ignorance and missing information [23]. In particular, it provides explicit estimations of imprecision and conflict between information from different sources. Indeed, probability theory may be seen as a limit of Dempster-Shafer evidence theory when it is assumed that there is no imprecision, and that only certainty has to be taken into account [24–27]. By using indistinguishability relations, rough sets theory can model and handle incomplete information and uncertain knowledge discovered from information system [2, 28–30].

In most synthesis evaluation and decision-making systems, data reflect the relation between objects and attributes with a degree of belief, that is, an object has an attribute with a degree of belief. Informally, the information system or decision tables with a degree of belief are called probability information system or probability decision tables. In this paper, we focus on probability decision tables, we analyze the consistency of probability decision rules which are extracted from the probability decision tables, and provide an inference method based on probability decision rules. The organization of this paper is as follows. In Section 2, we make briefly a review of extension of Dempster-Shafer evidence theory (DSEV). In Section 3, we provide a kind of probability information system and probability decision tables to represent objects and attributes with a degree of belief, and it is shown that our probability information system is extension of classical information system and a special case of interval-valued information system. In Section 4, we discuss how to extract probability decision rules from a probability decision table. In Section 5, we discuss consistency of probability decision rules. In Section 6, we provide a method to finish inference of probability decision rules. We conclude in Section 7.

2. Extension of DSEV

In DSEV, probability masses are allocated to subsets of a frame of discernment in contrast to Bayesian probability theory, in which only singletons are carrying probability masses. Such subsets with positive mass are called focal elements, and the family of focal elements is said to be the basic probability assignment. So, the DSEV can be seen to be generalization of the classic probability theory. Formally, the DSEV concerns itself with belief structures, which can be defined as follows: let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set of elements, and let m be a measure on the subsets of X such that (1) $0 \leq m(A) \leq 1$ for each $A \subset X$; (2) $m(\emptyset) = 0$; (3) $\sum_{A \subset X} m(A) = 1$. m is called a basic probability assignment function. Any subset A of X such that $m(A) > 0$ is called a focal element. m and its associated values are called by a belief structure. Two important functions play a significant role in DSEV: $\mathrm{Bel}(B) = \sum_{A \subset B} m(A)$

and $\mathrm{Pl}(B) = \sum_{A\cap B\neq\emptyset} m(A)$ for any subsets B and A of X. Bel is called a belief function, and Pl is called a plausibility function. $\mathrm{Bel}(B)$ measures the total amount of probability that must be distributed amongst the elements of B, and $\mathrm{Pl}(B)$ measures the maximal amount of probability that can be distributed among the elements in B. Denoted $\overline{B}$ is the complement of B, then we have

(i) $\mathrm{Pl}(B) \geq \mathrm{Bel}(B)$,

(ii) $\mathrm{Pl}(\emptyset) = \mathrm{Bel}(\emptyset) = 0$, $\mathrm{Pl}(X) = \mathrm{Bel}(X) = 1$,

(iii) $\mathrm{Pl}(B) = 1 - \mathrm{Bel}(\overline{B})$, $\mathrm{Bel}(B) = 1 - \mathrm{Pl}(\overline{B})$,

(iv) $\mathrm{Bel}(B) + \mathrm{Bel}(\overline{B}) \leq 1$, $\mathrm{Pl}(B) + \mathrm{Pl}(\overline{B}) \geq 1$.

In the classical probability model, the probability mass function is a mapping P from X to $[0,1]$, which indicates how the probability mass is assigned to the elements. Based on the probability mass function P, the set mapping $m' : 2^X \to [0,1]$ can be induced, where for each $A \subset X$, we have $m'(A) = \sum_{x\in A} P(x)$. Obviously, m' is a belief structure; furthermore, we have $\mathrm{Pl}(A) = \mathrm{Bel}(A)$. However, in DSEV, we know probabilities of the focal sets instead of probabilities of each element $x \in X$; hence, we are not able to calculate the probability $P(A)$ associated with the subsets of X, but instead to use the two measures $\mathrm{Bel}(A)$ and $\mathrm{Pl}(A)$, corresponding to a lower and upper bound on the unknown $P(A)$, that is, $\mathrm{Bel}(A) \leq P(A) \leq \mathrm{Pl}(A)$.

In [18], extending belief structure was proposed. Extending belief structure means using a belief structure defined on one frame to obtain a belief structure on another frame. Consider two frames X and Y, whose elements are possible answers to perhaps related questions. We say that an element $x \in X$ is compatible with an element $y \in Y$ if it is possible, relative to our knowledge and opinion, that x is the answer to the question considered by the frame X, and y is the answer to the question considered by the frame Y. Denote this as xRy. If for all $x \in X$ and all $y \in Y$, we have xRy, then we say that the two questions are independent.

Let X and Y be two frames, and $X \times Y$ is their Cartesian product, which consists of all pairs (x,y) with $x \in X$ and $y \in Y$, then the compatibility relation R is a subset of $X \times Y$ consisting of all pairs (x,y) for which xRy, that is, $(x,y) \in R$ if xRy. Specially, if X and Y are independent, then $R = X \times Y$. Any compatibility relation R over $X \times Y$ can be represented as a multivalued mapping: $G : X \to 2^Y$ such that $G(x) = \{y \mid (x,y) \in R\}$, where R is a compatibility relation. Assume that X and Y are two frames with a compatibility relation R and associated multivalued mapping G. Let Bel_X be a belief function defined on the frame X. The extension of Bel_X to Y as Bel_Y can be defined as follows:

$$\mathrm{Bel}_Y(B) = \mathrm{Bel}_X\left(\{x \mid (x,y) \in R \Longrightarrow y \in B\}\right) = \mathrm{Bel}_X(\{x \mid G(x) \subset B\}). \tag{2.1}$$

If $\{A_i\}$ are the focal elements of Bel_X where $m_X(A_i) = a_i$, then $\{B_i\}$, where $B_i = \bigcup_{x\in A_i} G(x)$, are the focal elements of Bel_Y and $m_Y(B_i) = \sum_k a_k$, over all k such that $B_k = B_i$. Let $H = \{x \mid G(x) \subset B\}$, then

$$\begin{aligned} \mathrm{Bel}_Y(B) &= \sum_{B_i\subset B} m_Y(B_i) = \sum_{\bigcup_{x\in A_i} G(x)\subset B}^{i} m_X(A_i), \\ \mathrm{Bel}_Y(B) &= \mathrm{Bel}_X(H) = \sum_{A_i\subset H} m_X(A_i). \end{aligned} \tag{2.2}$$

There exist two special cases of the extension: (1) assume that we know $\mathrm{Bel}_{S\times T}$, which is a belief function defined on $S \times T$, the marginal belief function of $\mathrm{Bel}_{S\times T}$ on S (or T) is defined as the extension of $\mathrm{Bel}_{S\times T}$ to S (or T), denoted as Bel_S (or Bel_T); (2) the extension of a belief function on the frame S to the frame $S \times T$. In (2.1), two extensions are expressed as follows.

(1) Let $X = S \times T$ and $Y = S$ (or T), then we have a multivalued mapping

$$G : S \times T \longrightarrow 2^S \quad \text{or} \quad G : S \times T \longrightarrow 2^T, \tag{2.3}$$

in which $G((s,t)) = \{s\}$ or $G((s,t)) = \{t\}$. Thus, we have a belief function

$$\mathrm{Bel}_Y(B) = \mathrm{Bel}_{S\times T}(\{(s,t) \mid G((s,t)) \subset B\}). \tag{2.4}$$

Furthermore, $\{(s,t) \mid G((s,t)) \subset B\} = \bigcup_{s\in B}(\{s\} \times T) = B \times T$, or $\{(s,t) \mid G((s,t)) \subset B\} = \bigcup_{t\in B}(S \times \{t\}) = S \times B$.

(2) The marginal belief function on S (or T) is

$$\mathrm{Bel}_Y(B) = \mathrm{Bel}_{S\times T}(B \times T) \quad \text{or} \quad \mathrm{Bel}_Y(B) = \mathrm{Bel}_{S\times T}(S \times B). \tag{2.5}$$

If $\{A_i\}$ are focal elements of $\mathrm{Bel}_{S\times T}$ with $m_{S\times T}(A_i) = a_i$, we get m_S of Bel_S (or m_T of Bel_T) as $m_S(B_i) = a_i$ (*or* $m_T(B_i) = a_i$), where $B_i = \bigcup_{(s,t)\in A_i} G((s,t))$, that is, B_i is projection of A_i onto S (or T).

Assume that Bel_S is a belief function on S, and let $\mathrm{Bel}_{S\times T}$ be the extension of Bel_S onto $S \times T$ that we know nothing about the answer on T. In this case, we call $\mathrm{Bel}_{S\times T}$ the minimal extension of Bel_S; in particular, we assume that every answer in T is compatible with any answer in S; thus, compatible function is

$$G : S \longrightarrow 2^{S\times T}, \qquad G(s) = \{(s,t) \mid \forall t \in T\}. \tag{2.6}$$

So, we have extension of Bel_S onto $S \times T$ as $\mathrm{Bel}_{S\times T}(B) = \mathrm{Bel}_S(\{s \mid G(s) \subset B\})$. Assume that $B = A \times T$, then as for all $s \in A$, we have $G(s) \in B$, and as for all $s \overline{\in} A$, we have $G(s) \not\subseteq B$. Hence, $\{s \mid G(s) \subset B\} = A$, and $\mathrm{Bel}_{S\times T}(A \times T) = \mathrm{Bel}_S(A)$. Assume that $B \neq A \times T$, then for all $s \in S$, $G(s) = \{s\} \times T \not\subseteq B$; thus, $\{s \mid G(s) \subset B\} = \emptyset$, and $\mathrm{Bel}_{S\times T}(B) = \mathrm{Bel}_S(\emptyset) = 0$. If we have $m_S(A_i) = a_i$, we can get $\mathrm{Bel}_{S\times T}(B_i) = a_i$, where $B_i = \bigcup_{s\in A_i} G(s)$. Since $G(s) = \{s\} \times T$, then $B_i = A_i \times T$, that is, B_i is the set consisting of each element in A_i coupled with each element in T.

Assume that we have two independent sources of evidence as the location of the special element in X, which have associated belief structures m_1 and m_2, respectively. The problem is to find a combined belief structure m over X reflecting the "ANDing" of the two pieces of evidence. Let m_1 and m_2 be two belief structures on X with focal elements $A_1, \ldots, A_k$ and $B_1, \ldots, B_p$, respectively, then their combination, denoted by $m = m_1 \oplus m_2$, is a belief structure over X such that (1) $m(A \neq \emptyset) = \sum_{A_i\cap B_j=A} m_1(A_i)m_2(B_j)/(1-K)$; (2) $m(\emptyset) = 0$; (3) $K = \sum_{A_i\cap B_j=\emptyset} m_1(A_i)m_2(B_j)$. The focal elements of m are all sets A such that $A_i \cap B_j = A$. Because m satisfies commutativity and associativity, we have $m = m_1 \oplus m_2 \oplus \cdots \oplus m_r$ for r belief structures on X. $1 - K$ indicates the normalization factor needed to assure $\sum m(A) = 1$.

If $K = 0$, then no normalization is required. If $K = 1$, then we cannot obtain $m(A \neq \emptyset)$ based on the above method (1). In this case, the belief structures are completely conflicting, and we need another method for combining evidence [18, 20].

3. Probability Information Systems

Information systems, sometimes called data tables, attribute value systems, condition action tables, knowledge representation systems, and so forth are used for representing knowledge and have been popularly used in artificial intelligence [2]. Formally, a pair $\Omega = (U, A)$ is called information systems. Where U is a nonempty, finite set called the universe and A a nonempty, finite set of attributes, that is, $a : U \to V_a$ for $a \in A$, where V_a is called the value set of a, elements of U are called objects and interpreted as cases, states, processes, patients, and observations. Attributes are interpreted as features, variables, characteristic, and conditions. As a special case of information systems, a decision table has the form $\Delta = (U, A \cup \{b\})$, where $b \overline{\in} A$ is a distinguished attribute called decision (attribute). The elements of A are called conditions (attributes).

The relation between an object $x \in U$ and a value $v_a \in V_a$ of an attribute $a \in A$ is certain, that is, $a : U \to V_a$ is a function, and either $a(x) = v_a$ or $a(x) \neq v_a$ is true. In real practice, the relation between objects and attribute values is uncertain, that is, $a(x) = v_a$ with a degree of belief $p_{(a(x)=v_a)} \in [0, 1]$. If $p_{(a(x)=v_a)} = 1$, it means that $a(x) = v_a$. If $p_{(a(x)=v_a)} = 0$, it means that $a(x) \neq v_a$. If $p_{(a(x)=v_a)} \in (0, 1)$, it means that $a(x) = v_a$ with uncertainty. In this paper, we limit the degree of belief in probability of $a(x) = v_a$.

Definition 3.1. A triple $\Omega = (U, A, P)$ is called a probability information system, where U is a nonempty, finite set called the universe, A a nonempty, finite set of attributes, and $P = \{P_a \mid a \in A\}$, $P_a : U \times A \to [0, 1]$ such that for each $x \in U$, $\sum_{v_a \in V_a} P_a(x, v_a) = 1$, $P_a(x, v_a) = p_{(a(x)=v_a)}$ means that object x has v_a of a with probability $p_{(a(x)=v_a)}$.

In Definition 3.1, for any fixed object $x \in U$, $\{P_a(x, v_a) \mid v_a \in V_a\}$ is a probability density function on V_a of a. Hence, a probability information system is also understood by an information system with a probability density function on the value set of each attribute for each object. An example of a probability information system is shown in Table 1, in which, for object 1 and solar energy, (high, 0.2) means that object 1 has high solar energy with probability 0.2, that is, $P_S((1, \text{high})) = 0.2$. For object 1 and residual CO_2, (high, 0) means that object 1 has not high residual CO_2, that is, $P_R((1, \text{high})) = 0$. For object 3 and temperature, (low, 1) means that object 4 has low value, that is, $P_T((1, \text{low})) = 1$.

In a probability information system, for any object $x \in U$ and $a \in A$, we denote

$$\Lambda_a^x = \{v_a \in V_a \mid P_a((x, v_a)) \neq 0\}. \tag{3.1}$$

If $|\Lambda_a^x| = 1$ ($|\Lambda_a^x|$ is cardinality of Λ_a^x), then there exists unique $v_a \in V_a$ such that $P_a((x, v_a)) = 1$ due to $\sum_{v_a \in V_a} P_a(x, v_a) = 1$. In this case, the probability information system is reduced to an information system because for each object x and $a \in A$, there exists only one $v_a \in V_a$ such that $a(x) = v_a$. From this point of view, probability information systems are extension of information systems. If $|\Lambda_a^x| > 1$, then the probability information system becomes an interval-valued information system when all probabilities $P_a((x, v_a)) \neq 0$ are not considered, in such case, for each object x and $a \in A$, we have $a(x) = \Lambda_a^x$. Hence, a probability information system

Table 1: An example of probability information system.

Fact	Solar energy (S)	Volcanic activity (V)	Residual CO_2 (R)	Temperature (T)
1	(High, 0.2)	(High, 0.6)	(High, 0)	(High, 0.8)
	(Medium, 0.8)	(Medium, 0.3)	(Medium, 0.3)	(Medium, 0.2)
	(Low, 0)	(Low, 0.1)	(Low, 0.7)	(Low, 0)
2	(High, 0.9)	(High, 0.8)	(High, 0.7)	(High, 0.8)
	(Medium, 0.1)	(Medium, 0.2)	(Medium, 0.3)	(Medium, 0.1)
	(Low, 0)	(Low, 0)	(Low, 0)	(Low, 0.1)
3	(High, 0.1)	(High, 0)	(High, 0)	(High, 0)
	(Medium, 0.2)	(Medium, 0.3)	(Medium, 0.2)	(Medium, 0)
	(Low, 0.7)	(Low, 0.7)	(Low, 0.8)	(Low, 1)

Table 2: A probability decision table.

$U \setminus A \cup \{b\}$	a_1	$\cdots$	a_m	b
x_1	$(v_{a_1}^1, p_{a_1}^{x_1}(v_{a_1}^1))$	$\cdots$	$(v_{a_m}^1, p_{a_m}^{x_1}(v_{a_m}^1))$	$(v_b^1, p_b^{x_1}(v_b^1))$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$(v_{a_1}^{r_1}, p_{a_1}^{x_1}(v_{a_1}^{r_1}))$	$\cdots$	$(v_{a_m}^{r_m}, p_{a_m}^{x_1}(v_{a_m}^{r_m}))$	$(v_b^s, p_b^{x_1}(v_b^s))$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$(v_{a_1}^1, p_{a_1}^{x_n}(v_{a_1}^1))$	$\cdots$	$(v_{a_m}^1, p_{a_m}^{x_n}(v_{a_m}^1))$	$(v_b^1, p_b^{x_n}(v_b^1))$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$(v_{a_1}^{r_1}, p_{a_1}^{x_n}(v_{a_1}^{r_1}))$	$\cdots$	$(v_{a_m}^{r_m}, p_{a_m}^{x_n}(v_{a_m}^{r_m}))$	$(v_b^s, p_b^{x_n}(v_b^s))$

of Definition 3.1 is a special case of interval-valued information system. Deference between them is $a(x) = \Lambda_a^x$ in interval-valued information system, but $a(x) = \Lambda_a^x$ with a probability density function on Λ_a^x in probability information system.

Based on probability information system, probability decision tables have the form $\Delta = (U, A \cup \{b\}, P)$, A is called conditions, and b is called decision. Let $P = \{P_c \mid c \in A \cup \{b\}\}$, $P_c : U \times (A \cup \{b\}) \to [0,1]$ such that for each $x \in U$ and $c \in A \cup \{b\}$, $\sum_{v_c \in V_c} P_c^x((x, v_c)) = 1$. For simplicity, denote V_c and P_c as V_a and P_a, respectively, when $c \in A$, and V_c and P_c as V_b and P_b when $c = b$. For any object $x \in U$ and $c \in A \cup \{b\}$, denote $\Lambda_c^x = \{v_c \in V_c \mid P_c((x, v_c)) \neq 0\}$. If $|\Lambda_c^x| = 1$, then probability decision table is reduced to classical decision tables. If $|\Lambda_a^x| > 1$, then probability decision table becomes interval-valued decision tables when all probabilities $P_a((x, v_a)) \neq 0$ are not considered.

Let $U = \{x_1, x_2, \ldots, x_n\}$ and $A = \{a_1, \ldots, a_m\}$. For each $x_i \in U$ and attribute value set V_{a_j} of attribute $a_j \in A$, $V_{a_j} = \{v_{a_j}^1, \ldots, v_{a_j}^{r_j}\}$, denote $P_{a_j}^{x_i}$ as a probability density function on V_{a_j}, and denote $P_b^{x_i}$ as a probability density function on decision value set $V_b = \{v_b^1, \ldots, v_b^s\}$. Formally, a probability decision table is shown in Table 2.

In Table 2, $p_{a_j}^{x_i}(v_{a_j}^{r_j'})$ is the probability of "x_i that has the value $v_{a_j}^{r_j'}$ of a_j", where $r_j' \in \{1, 2, \ldots, r_j\}$, and for each x_i, $\sum_{r_j'=1}^{r_j} p_{a_j}^{x_i}(v_{a_j}^{r_j'}) = 1$. For each x_i and $c \in A \cup \{b\}$, denote

$$p_{a_j}^{x_i}\left(v_{a_j}^{r_j'}\right) = \max\left\{p_{a_j}^{x_i}\left(v_{a_j}^1\right), p_{a_j}^{x_i}\left(v_{a_j}^2\right), \ldots, p_{a_j}^{x_i}\left(v_{a_j}^{r_j}\right)\right\}. \tag{3.2}$$

Table 3: The maximum probability decision table of Table 2.

$U \setminus A \cup \{b\}$	a_1	a_2	$\cdots$	a_m	b
x_1	$(v_{a_1}^{r_1'}, p_{a_1}^{x_1}(v_{a_1}^{r_1'}))$	$(v_{a_2}^{r_2'}, p_{a_2}^{x_1}(v_{a_2}^{r_2'}))$	$\cdots$	$(v_{a_m}^{r_m'}, p_{a_m}^{x_1}(v_{a_m}^{r_m'}))$	$(v_b^{s'}, p_b^{x_1}(v_b^{s'}))$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$(v_{a_1}^{r_1'}, p_{a_1}^{x_n}(v_{a_1}^{r_1'}))$	$(v_{a_2}^{r_2'}, p_{a_2}^{x_n}(v_{a_2}^{r_2'}))$	$\cdots$	$(v_{a_m}^{r_m'}, p_{a_m}^{x_n}(v_{a_m}^{r_m'}))$	$(v_b^{s'}, p_b^{x_n}(v_b^{s'}))$

Then we can get a reduced probability decision table shown in Table 3, which is called the maximum probability decision table in this paper.

4. Probability Decision Rules

In a decision table $\Delta = (U, A \cup \{b\})$, decision rules can be extracted and formalized as follows:

$$\tau \longrightarrow b = i, \tag{4.1}$$

where τ is a formula, which is generated by some $v_i \in V$ finitely using connectives $\vee$ or $\wedge$, $V = \bigcup_{a \in A} V_a$, and $i \in V_b$. For more details about decision rules, reader can consult [2, 28–30]. In this section, we discuss some properties and formalization of probability decision rules based on Table 3.

Firstly, we can use rough set theory to deal with the information of Table 3 without considering probability $p_{a_j}^{x_i}(v_{a_j}^{r_j'})$; for simplicity, denote

$$\begin{aligned} V_{a_j}' &= \left\{ v_{a_j} \mid \exists x_i \in U, P_{a_j}^{x_i}\left(v_{a_j}\right) = \max\left\{ P_{a_j}^{x_i}\left(v_{a_j}^1\right), \ldots, P_{a_j}^{x_i}\left(v_{a_j}^{r_j}\right) \right\} \right\}, \\ V_b' &= \left\{ v_b^k \mid \exists x_i \in U, P_b^{x_i} v_b^k = \max\left\{ P_b^{x_i}\left(v_b^1\right), \ldots, P_b^{x_i}\left(v_b^s\right) \right\} \right\}. \end{aligned} \tag{4.2}$$

Then attributes of Table 3 can be represented as follows:

$$\begin{aligned} a_j : U \longrightarrow V_{a_j}', \qquad a_j(x_i) = v_{a_j}, \\ b : U \longrightarrow V_b', \qquad b(x_i) = v_b^k. \end{aligned} \tag{4.3}$$

We can define an equivalence relation on U as follows:

$$x_i \sim x_i' \quad \text{iff } a_j(x_i) = a_j\left(x_i'\right). \tag{4.4}$$

Then lower and upper approximation, reduction of attributes, decision rules, consistent decision tables, and so on can be discussed. However, in the paper, the problem that we need to deal with is how to get probability of every decision rule. By using rough set theory, we can get decision rules as follows:

$$\tau^k \longrightarrow b = v_b^k, \tag{4.5}$$

where $v_b^k \in V_b'$, $\tau^k = \wedge_{l=1}^{q} v_{a_l'}$, $v_{a_l'} \in V_{a_l'}'$, a_l' and $a_l' \in \{a_1, \ldots, a_m\}$. Let $U_{\tau^k} \subseteq U$ is the equivalence class decided by τ^k, $U_{v_b^k} \subseteq U$ is the equivalence class decided by v_b^k. for all $x_i \in U_{\tau^k}$ and for all $v_{a_l'} \in \tau^k$, $P_{a_l'}^{x_i}(v_{a_l'}^{r_l'})$ is the probability of "x_i has $v_{a_l'}^{r_l'}$ of a_l'". Let $|U_{\tau^k}| = T$, then we have T independent sources of evidence over $U_{\tau^k} \times V_{a_l'}$, and belief structure of each source of evidence is as follows:

$$m_s\left(U_{\tau^k}, v_{a_l'}^1\right) = P_{a_l'}^{x_s}\left(v_{a_l'}^1\right), \ldots, m_s\left(U_{\tau^k}, v_{a_l'}^{r_l}\right) = P_{a_l'}^{x_s}\left(v_{a_l'}^{r_l}\right), \tag{4.6}$$

where $s = 1, \ldots, T$. Based on (4.6), we can obtain a combined belief structure m over $U_{\tau^k} \times V_{a_l'}$, that is, $m = m_1 \oplus \cdots \oplus m_T$ as follows [19]:

$$P_{U_{\tau^k}}^{a_l'} = \oplus_{x_s \in U_{\tau^k}} P_{a_l'}^{x_s}, \tag{4.7}$$

$$P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_l'}\right) = \frac{\prod_{x_s \in U_{\tau^k}} P_{a_l'}^{x_s}\left(v_{a_l'}^{r_l'}\right)}{1 - K}, \tag{4.8}$$

$$1 - K = \sum_{v_{a_l'}^{r_l'} \in V_{a_l'}} \prod_{x_s \in U_{\tau^k}} P_{a_l'}^{x_s}\left(v_{a_l'}^{r_l'}\right). \tag{4.9}$$

For attribute a_l', $P_{a_l'}^{x_s}(v_{a_l'})$ is the maximum probability that x_s has the value $v_{a_l'}$ of a_l', so, for all $v_{a_l'}^{r_l'} \in V_{a_l'}$ and for all $x_s \in U_{\tau^k}$, $P_{a_l'}^{x_s}(v_{a_l'}) \geq P_{a_l'}^{x_s}(v_{a_l'}^{r_l'})$, that is,

$$\prod_{x_i \in U_{\tau^k}} P_{a_l'}^{x_i}\left(\left(x_i, v_{a_l'}\right)\right) \geq \prod_{x_i \in U_{\tau^k}} P_{a_l'}^{x_i}\left(\left(x_i, v_{a_l'}^{r_l'}\right)\right). \tag{4.10}$$

This means that $P_{U_{\tau^k}}^{a_l'}(v_{a_l'})$ is the maximum possibility of all $P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_l'})$. Similarly, we can get the probability of each $v_b^{k'}$ as follows:

$$P_{U_{v_b^k}}^{b} = \oplus_{x_i \in U_{v_b^k}} P_b^{x_i}, \tag{4.11}$$

$$P_{U_{v_b^k}}^{b}\left(v_b^{k'}\right) = \frac{\prod_{x_s \in U_{v_b^k}} P_b^{x_s}\left(v_b^{k'}\right)}{1 - K}, \tag{4.12}$$

$$1 - K = \sum_{v_b^{k'} \in V_b} \prod_{x_s \in U_{v_b^k}} P_b^{x_s}\left(v_b^{k'}\right), \tag{4.13}$$

where $P_{U_{v_b^k}}^{b}(v_b^k)$ is the maximum possibility of $P_{U_{v_b^k}}^{b}(v_b^{k'})$.

According to (4.8) and (4.12), we can get a probability decision rule as follows:

$$\left(\tau^k, P_{U_{\tau^k}}^{\tau^k}\right) \longrightarrow b = \left(v_b^k, P_{U_{v_b^k}}^{b}\left(v_b^k\right)\right), \tag{4.14}$$

where $(\tau^k, P^{\tau^k}_{U_{\tau^k}}) = \wedge_{l=1}^{q}(v_{a'_l}, P^{a'_l}_{U_{\tau^k}}(v_{a'_l}))$, $q \le m$. The decision rule (4.5) is a special case of the probability decision rule (4.14), that is, if every $P^{a'_l}_{U_{\tau^k}}(v_{a'_l}) = 1$ and $P^{b}_{U_{v^k_b}}(v^k_b) = 1$. Obviously, for every $a'_l \in A$,

$$\sum_{r'_l=1}^{r_l} P^{a'_l}_{U_{\tau^k}}\left(v^{r'_l}_{a'_l}\right) = 1, \qquad \sum_{k'=1}^{s} P^{b}_{U_{v^k_b}}\left(v^{k'}_{b}\right) = 1. \tag{4.15}$$

Theorem 4.1. *Let $x_i \in U_{\tau^k}$ if $P^{x_i}_{a'_l}((x_i, v_{a'_l})) = 1$, then $P^{a'_l}_{U_{\tau^k}}(v_{a'_l}) = 1$.*

Proof. By $P^{x_i}_{a'_l}((x_i, v_{a'_l})) = 1$, we know that for all $v^{r'_l}_{a'_l} \in V_{a'_l}$, $v^{r'_l}_{a'_l} \neq v_{a'_l}$, $P^{x_i}_{a'_l}((x_i, v^{r'_l}_{a'_l})) = 0$, so $1 - K = \sum_{v^{r'_l}_{a'_l} \in V_{a'_l}} \prod_{x_i \in U_{\tau^k}} P^{x_i}_{a'_l}((x_i, v^{r'_l}_{a'_l})) = \prod_{x_i \in U_{\tau^k}} P^{x_i}_{a'_l}((x_i, v_{a'_l}))$, by (4.8),

$$P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right) = \frac{\prod_{x_i \in U_{\tau^k}} P^{x_i}_{a'_l}\left(\left(x_i, v_{a'_l}\right)\right)}{1 - K} = 1. \tag{4.16}$$

□

Theorem 4.2. *Let x' be a new object, then*

$$\sum_{v^{r'_j}_{a'_l} \in V_{a'_l}} \left(P^{x'}_{a'_l}\left(v_{a'_l}\right) - P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)\right) P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right) < 0 \tag{4.17}$$

if and only if $P^{a'_l}_{U_{\tau^k} \cup \{x'\}}(v_{a'_l}) < P^{a'_l}_{U_{\tau^k}}(v_{a'_l})$.

Proof. According to (4.7), we know that $P^{a'_l}_{U_{\tau^k} \cup \{x'\}} = \oplus_{x_i \in U_{\tau^k} \cup \{x'\}} P^{x_i}_{a'_l} = (\oplus_{x_i \in U_{\tau^k}} P^{x_i}_{a'_l}) \oplus P^{x'}_{a'_l} = P^{a'_l}_{U_{\tau^k}} \oplus P^{x'}_{a'_l}$. So, we have

$$\begin{aligned} P^{a'_l}_{U_{\tau^k} \cup \{x'\}}\left(v_{a'_l}\right) - P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right) &= \frac{P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right) P^{x'}_{a'_l}\left(v_{a'_l}\right)}{\sum_{v^{r'_j}_{a'_l} \in V_{a'_l}} P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right) P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)} - P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right) \\ &= P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right)\left(\frac{P^{x'}_{a'_l}\left(v_{a'_l}\right)}{\sum_{v^{r'_j}_{a'_l} \in V_{a'_l}} P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right) P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)} - 1\right) \end{aligned}$$

$$= P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right)\frac{\sum_{v_{a_l'}^{r_j'}\neq v_{a_l'}}\left(P_{a_l'}^{x'}\left(v_{a_l'}\right)-P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)\right)P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)}{\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)}$$

$$= P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right)\frac{\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}\left(P_{a_l'}^{x'}\left(v_{a_l'}\right)-P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)\right)P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)}{\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)}, \tag{4.18}$$

and this means that if $P_{U_{\tau^k}\cup\{x'\}}^{a_l'}(v_{a_l'}) < P_{U_{\tau^k}}^{a_l'}(v_{a_l'})$, then

$$P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right)\frac{\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}\left(P_{a_l'}^{x'}\left(v_{a_l'}\right)-P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)\right)P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)}{\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}^{r_j'}\right)P_{a_l'}^{x'}\left(v_{a_l'}^{r_j'}\right)} < 0, \tag{4.19}$$

that is, $\sum_{v_{a_l'}^{r_j'}\in V_{a_l'}}(P_{a_l'}^{x'}(v_{a_l'}) - P_{a_l'}^{x'}(v_{a_l'}^{r_j'}))P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_j'}) < 0$. Converse is obvious. □

Corollary 4.3. *If* $P_{a_l'}^{x'}((x', v_{a_l'})) = \max_{v_{a_l'}^{r_j'}\in V_{a_l'}}\{P_{a_l'}^{x_i}((x_i, v_{a_l'}^{r_j'}))\}$, *then*

$$P_{U_{\tau^k}\cup\{x'\}}^{a_l'}\left(v_{a_l'}\right) \geq P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right). \tag{4.20}$$

According to the maximum possibility, Corollary 4.3 means that the more the elements of U_{τ^k} are, the more the probabilities of $v_{a_l'}$ of τ^k are. Intuitively, the elements of U_{τ^k} reflect the degree of belief of a decision rule; the more the elements of U_{τ^k} are, the more the degree of belief of a decision rule is. Theorem 4.2 means that, sometimes, adding an element to U_{τ^k} can make the degree of belief of a decision rule decrease. According to the condition of Theorem 4.2, one can see that $P_{a_l'}^{x'}(v_{a_l'})$ is not the maximal probability. From the probability point of view, x' may not be in U_{τ^k}; if x' is forced in U_{τ^k}, then the degree of belief of a decision rule will decrease.

5. Consistency of a Probability Decision Rule

From the logic point of view, if there exists a valuation v such that $v(a) = 1$ and $v(b) = 1$, then $a \to b$ is satisfiable, otherwise, $a \to b$ is not satisfiable. We discuss consistency of probability decision rule similar to satisfiability of $a \to b$.

Definition 5.1. In probability decision rule (4.14), for a new object x', if for all $v_{a_l'}$, $l = 1, \ldots, q$, x' is such that

$$P^{a_l'}_{U_{\tau k}\cup\{x'\}}\left(v_{a_l'}\right) \geq P^{a_l'}_{U_{\tau k}}\left(v_{a_l'}\right), \qquad P^{b}_{U_{v_b^k}\cup\{x'\}}\left(v_b^k\right) \geq P^{b}_{U_{v_b^k}}\left(v_b^k\right), \tag{5.1}$$

then there exists consistency between x' and the probability decision rule, and we also call antecedents and conclusion of the probability decision rule consistent. If there does not exist such x', then we call antecedents and conclusion of the decision rule inconsistent.

By Theorem 4.2, we know that

$$\sum_{v^{r_j'}_{a_l'}\in V_{a_l'}} \left(P^{x'}_{a_l'}\left(v_{a_l'}\right) - P^{x'}_{a_l'}\left(v^{r_j'}_{a_l'}\right)\right) P^{a_l'}_{U_{\tau k}}\left(v^{r_j'}_{a_l'}\right) \geq 0, \qquad \sum_{s'=1}^{s}\left(P^{x'}_{b}\left(v_b^k\right) - P^{x'}_{b}\left(v_b^{s'}\right)\right) P^{b}_{U_{v_b^k}}\left(v_b^{s'}\right) \geq 0 \tag{5.2}$$

if and only if (5.1) is satisfied. Obviously, if $P^{x'}_{a_l'}(v_{a_l'})$ and $P^{x'}_{b}(v_b^k)$ are maximum probability assignments, respectively, that is, if they satisfy the condition of Corollary 4.3, then (5.1) is true. So, using maximum probability assignments to get a probability decision rule, its antecedents and conclusion are consistent, and for a new object x' and its probability assignments on V_{a_j}, if we use maximum probability assignment to decide the class of x', then x' and the probability decision rule are consistent. However, sometimes, there exists a case: for a new object x', according to maximum probability assignment, we get each attribute value $v^{r_{j'}'}_{a_l'}$ and $v_b^{k'}$, but there does not exist decision rule such that its antecedents and conclusion match $v^{r_{j'}'}_{a_l'}$ and $v_b^{k'}$, respectively. In this case, using (5.2), we can choose a probability decision rule such that x' and the probability decision rule are consistent. If there exist more than one decision rule, then x' is included in the decision rule such that its conclusion is $\vee P^{b}_{U_{v_b^k}\cup\{x'\}}(v_b^k)$.

Corollary 5.2. *A new object x' and decision rule* (4.14) *are consistent if and only if the probability assignments of x' are such that* (5.2).

Theorem 5.3. *For a new object x', let*

$$\sum_{s'=1}^{s} P^{x'}_{b}\left(v_b^{s'}\right) P^{b}_{U_{v_b^k}}\left(v_b^{s'}\right) = \alpha \sum_{v^{r_j'}_{a_l'}\in V_{a_l'}} P^{a_l'}_{U_{\tau k}}\left(v^{r_j'}_{a_l'}\right) P^{x'}_{a_l'}\left(v^{r_j'}_{a_l'}\right) \tag{5.3}$$

if $\alpha \leq P^{x'}_{b}(v_b^k)/P^{x'}_{a_l'}(v_{a_l'})$ and $P^{a_l'}_{U_{\tau k}\cup\{x'\}}(v_{a_l'}) \geq P^{a_l'}_{U_{\tau k}}(v_{a_l'})$, then $P^{b}_{U_{v_b^k}\cup\{x'\}}(v_b^k) \geq P^{b}_{U_{v_b^k}}(v_b^k)$.

Proof. By $P^{a_l'}_{U_{\tau k}\cup\{x'\}}(v_{a_l'}) \geq P^{a_l'}_{U_{\tau k}}(v_{a_l'})$, we know that

$$\sum_{v^{r_j'}_{a_l'}\in V_{a_l'}} \left(P^{x'}_{a_l'}\left(v_{a_l'}\right) - P^{x'}_{a_l'}\left(v^{r_j'}_{a_l'}\right)\right) P^{a_l'}_{U_{\tau k}}\left(v^{r_j'}_{a_l'}\right) \geq 0. \tag{5.4}$$

On the other hand,

$$
\begin{aligned}
P^b_{U_{v^k_b}\cup\{x'\}}\left(v^k_b\right) - P^b_{U_{v^k_b}}\left(v^k_b\right) &= P^b_{U_{v^k_b}}\left(v^k_b\right)\left(\frac{P^{x'}_b\left(v^k_b\right)}{\sum_{s'=1}^{s} P^{x'}_b\left(v^{s'}_b\right) P^b_{U_{v^k_b}}\left(v^{s'}_b\right)} - 1\right) \\
&= P^b_{U_{v^k_b}}\left(v^k_b\right)\left(\frac{P^{x'}_b\left(v^k_b\right)}{\alpha \sum_{v^{r'_j}_{a'_l}\in V_{a'_l}} P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right) P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)} - 1\right) \\
&= P^b_{U_{v^k_b}}\left(v^k_b\right)\frac{\sum_{v^{r'_j}_{a'_l}\in V_{a'_l}}\left(P^{x'}_b\left(v^k_b\right)/\alpha - P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right)\right) P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)}{\sum_{v^{r'_j}_{a'_l}\in V_{a'_l}} P^{a'_l}_{U_{\tau^k}}\left(v^{r'_j}_{a'_l}\right) P^{x'}_{a'_l}\left(v^{r'_j}_{a'_l}\right)}.
\end{aligned}
\tag{5.5}
$$

By $\alpha \le P^{x'}_b(v^k_b)/P^{x'}_{a'_l}(v_{a'_l})$, get $P^{x'}_b(v^k_b)/\alpha \ge P^{x'}_{a'_l}(v_{a'_l})$, so we have $\sum_{v^{r'_j}_{a'_l}\in V_{a'_l}}(P^{x'}_b(v^k_b)/\alpha - P^{a'_l}_{U_{\tau^k}}(v^{r'_j}_{a'_l}))P^{x'}_{a'_l}(v^{r'_j}_{a'_l}) \ge \sum_{v^{r'_j}_{a'_l}\in V_{a'_l}}(P^{x'}_{a'_l}(v_{a'_l}) - P^{x'}_{a'_l}(v^{r'_j}_{a'_l}))P^{a'_l}_{U_{\tau^k}}(v^{r'_j}_{a'_l}) \ge 0$, that is, $P^b_{U_{v^k_b}\cup\{x'\}}(v^k_b) \ge P^b_{U_{v^k_b}}(v^k_b)$. □

6. Inference of Probability Decision Rules

Inspired by inference method of Zadeh [13], we provide an inference method of probability decision rules in this section, where the decision rule has the form $\wedge^q_{l=1}(v_{a'_l}, P^{a'_l}_{U_{\tau^k}}(v_{a'_l})) \to b = (v^k_b, P^b_{U_{v^k_b}}(v^k_b))$. Assume that we have a new element x' shown in Table 4, we infer a probability density function on V_b and decide the class of x' and its degree of belief. The inference process can be rewritten as follows:

$$
\begin{aligned}
&p : \bigwedge_{l=1}^{q}\left(v_{a'_l}, P^{a'_l}_{U_{\tau^k}}\left(v_{a'_l}\right)\right) \longrightarrow b = \left(v^k_b, P^b_{U_{v^k_b}}\left(v^k_b\right)\right), \\
&q : \text{give Table 4}, \\
&\overline{c : \qquad\qquad \left(v^{s'}_b, P^{s'}_{x',b}\right).}
\end{aligned}
\tag{6.1}
$$

In (6.1), the probability decision rule p has the universes of U_{τ^k} of antecedent and $U_{v^k_b}$ of conclusion. In (6.1), p can be rewritten as "if the degree of belief of $v_{a'_1}$ is $P^{a'_l}_{U_{\tau^k}}(v_{a'_l})$ and ... and the degree of belief of $v_{a'_q}$ is $P^{a'_q}_{U_{\tau^k}}(v_{a'_q})$, then the degree of belief of v^k_b is $P^b_{U_{v^k_b}}(v^k_b)$," q can be rewritten as "the degree of belief of attribute values of x' is Table 4"; the conclusion is "how

many are the degrees of belief of decision attribute values $v_b^{s'}$ of x'." The single condition probability decision rule of p has the form:

$$\begin{aligned} &p : \left(v_{a_l'}, P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right)\right) \longrightarrow b = \left(v_b^k, P_{U_{v_b^k}}^{b}\left(v_b^k\right)\right), \\ &q : \sum_{i=1}^{r_l} p_{x',a_l'}^{i} = 1, \\ &\overline{c : \qquad\qquad \left(v_b^{s'}, P_{x',b}^{s'}\right).} \end{aligned} \tag{6.2}$$

In the paper, domains of antecedents and conclusion of p are $\{(U_{\tau^k}, v'_{a_l'}) \mid v'_{a_l'} \in V_{a_1'}\}$ and $\{(U_{v_b^k}, v'_b) \mid v'_b \in V_b\}$, respectively. We rewrite probability density functions as follows:

$$\begin{aligned} &\widetilde{f^{a_1'}}_{U_{\tau^k}} : \left\{\left(X, v'_{a_l'}\right) \mid v'_{a_l'} \in V_{a_1'}\right\} \longrightarrow [0,1], \qquad \widetilde{f^{a_1'}}_{U_{\tau^k}}\left(X, v'_{a_l'}\right) = P_{U_{\tau^k}}^{a_l'}\left(v_{a_l'}\right), \\ &\widetilde{f^{b}}_{U_{v_b^k}} : \{(X, v'_b) \mid v'_b \in V_b\} \longrightarrow [0,1], \qquad \widetilde{f^{b}}_{U_{v_b^k}}(X, v'_b) = P_{U_{v_b^k}}^{b}(v'_b), \\ &\widetilde{f^{a_l'}}_{x'} : \left\{\left(X, v'_{a_l'}\right) \mid v'_{a_1'} \in V_{a_l'}\right\} \longrightarrow [0,1], \qquad \widetilde{f^{a_l'}}_{x'}\left(X, v'_{a_l'}\right) = p_{x',a_l'}^{v'_{a_l'}}. \end{aligned} \tag{6.3}$$

Then, (6.1) and (6.2) can be modified as

$$\begin{aligned} &p : \bigwedge_{l=1}^{q} \widetilde{f^{a_l'}}_{U_{\tau^k}} \longrightarrow \widetilde{f^{b}}_{U_{v_b^k}}, \\ &q : \bigwedge_{l=1}^{q} \widetilde{f^{a_l'}}_{x'}, \\ &\overline{c : \qquad\qquad \widetilde{f^{b}}_{x'},} \end{aligned} \tag{6.4}$$

$$\begin{aligned} &p : \widetilde{f^{a_1'}}_{U_{\tau^k}} \longrightarrow \widetilde{f^{b}}_{U_{v_b^k}}, \\ &q : \widetilde{f^{a_l'}}_{x'}, \\ &\overline{c : \qquad\qquad \widetilde{f^{b}}_{x'}.} \end{aligned} \tag{6.5}$$

Let $\widetilde{f^{a_1'}}_{U_{\tau^k}} = P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^1)/(X, v_{a_l'}^1) + \cdots + P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_l})/(X, v_{a_l'}^{r_l})$ and $\widetilde{f^{b}}_{U_{v_b^k}} = P_{U_{v_b^1}}^{b}(v_b^1)/(X, v_b^1) + \cdots + P_{U_{v_b^s}}^{b}(v_b^s)/(X, v_b^s)$, then the logical combinations of $\widetilde{f^{a_1'}}_{U_{\tau^k}}$ and $\widetilde{f^{b}}_{U_{v_b^k}}$ are given as follows: $\widetilde{f^{a_1'}}_{U_{\tau^k}} \oplus \widetilde{f^{b}}_{U_{v_b^k}} = \min\{1, P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_l'}) + P_{U_{v_b^k}}^{b}(v_b^{s'})\}/(v_{a_l'}^{r_l'}, v_b^{s'})$, $\widetilde{f^{a_1'}}_{U_{\tau^k}} \rightarrow \widetilde{f^{b}}_{U_{v_b^k}} = \min\{1, 1 - P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_l'}) + P_{U_{v_b^1}}^{b}(v_b^{s'})\}/(v_{a_l'}^{r_l'}, v_b^{s'})$, $\widetilde{f^{a_1'}}_{U_{\tau^k}} \times \widetilde{f^{b}}_{U_{v_b^k}} = \min\{P_{U_{\tau^k}}^{a_l'}(v_{a_l'}^{r_l'}), P_{U_{v_b^k}}^{b}(v_b^{s'})\}/(v_{a_l'}^{r_l'}, v_b^{s'})$, and

$(\widetilde{f^{a_1'}}_{U_{\tau^k}})' = (1 - P^{a_l'}_{U_{\tau^k}}(v^{r_l'}_{a_l'}))/(X, v^{r_l'}_{a_l'})$. Then, we can get the following relations on $\{(X, v'_{a_l'}) \mid v'_{a_l'} \in V_{a_1'}\} \times \{(X, v'_b) \mid v'_b \in V_b\}$:

$$R_1 = \widetilde{f^{a_1'}}_{U_{\tau^k}} \times \widetilde{f^b}_{U_{v_b^k}}, \tag{6.6}$$

$$R_2 = \left(\widetilde{f^{a_1'}}_{U_{\tau^k}} \times \widetilde{f^b}_{U_{v_b^k}}\right) \vee \left(\left(\widetilde{f^{a_1'}}_{U_{\tau^k}}\right)' \times \left(\widetilde{f^b}_{U_{v_b^k}}\right)'\right), \tag{6.7}$$

$$R_3 = \left(\left(\widetilde{f^{a_1'}}_{U_{\tau^k}}\right)' \oplus \widetilde{f^b}_{U_{v_b^k}}\right) \wedge \left(\widetilde{f^{a_1'}}_{U_{\tau^k}} \oplus \left(\widetilde{f^b}_{U_{v_b^k}}\right)'\right), \tag{6.8}$$

$$R_4 = \widetilde{f^{a_1'}}_{U_{\tau^k}} \longrightarrow \widetilde{f^b}_{U_{v_b^k}}, \tag{6.9}$$

$$R_5 = \left(\widetilde{f^{a_1'}}_{U_{\tau^k}} \longrightarrow \widetilde{f^b}_{U_{v_b^k}}\right) \vee \left(\left(\widetilde{f^{a_1'}}_{U_{\tau^k}}\right)' \longrightarrow \left(\widetilde{f^b}_{U_{v_b^k}}\right)'\right). \tag{6.10}$$

The inference for (6.5) can be expressed as follows:

$$\widetilde{f^b_{x'}} = \widetilde{f^{a_l'}_{x'}} \circ R_i, \tag{6.11}$$

where "$\circ$" operator is "max-min." According to (6.11), we can obtain $\widetilde{f^b_{x'}} = \widetilde{P^b_{x'}}(v_b^1)/(X, v_b^1) + \cdots + \widetilde{P^b_{x'}}(v_b^s)/(X, v_b^s)$; generally, maybe $P = \sum_{s'=1}^{s} \widetilde{P^b_{x'}}(v_b^{s'}) \neq 1$, however, every $\widetilde{P^b_{x'}}(v_b^{s'})$ $(s' = 1, 2, \ldots, s)$ of $\widetilde{f^b_{x'}}$ can be normalized as follows:

$$P^b_{x'}\left(v_b^{s'}\right) = \frac{\widetilde{P^b_{x'}}\left(v_b^{s'}\right)}{P}, \tag{6.12}$$

where P is normalized factor. By (6.12), we can get the probability density function on $\{(x', v_b^{s'}) \mid v_b^{s'} \in V_b\}$,

$$f^b_{x'} = \frac{P^b_{x'}(v_b^1)}{(x', v_b^1)} + \cdots + \frac{P^b_{x'}(v_b^s)}{(x', v_b^s)}. \tag{6.13}$$

According to the maximum probability of (6.13), we can decide in which class x' is included and its degree of belief.

Table 4: The information of x'.

	a_1	$\cdots$	a_m	b
x'	$(v_{a_1}^1, p_{x',a_1}^1)$	$\cdots$	$(v_{a_m}^1, p_{x',a_m}^1)$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$(v_{a_1}^{r_1}, p_{x',a_1}^{r_1})$	$\cdots$	$(v_{x',a_m}^{r_m}, p_{x',a_m}^{r_m})$	

Theorem 6.1. *In* (6.11)*, if* $R_i = R_1$, $P_{U_{\tau k}}^{a_l'}(v_{a_l'})$, $P_{U_{v_b^k}}^{b}(v_b^k)$, *and* $P_{x',a_l'}^{v_{a_l'}}$ *are maximum probabilities, respectively, then* $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^{b}(v_b^k))$ *is the maximum probability of* (6.13).

Proof. By (6.6) and (6.11), we know that $P_{x'}^b(v_b^k) = \max\{\min(P_{x',a_l'}^{v_{a_l'}^m}, \min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^k}}^{b}(v_b^k))) \mid v_{a_l'}^m \in V_{a_1'}\} = \max\{\min(P_{x',a_l'}^{v_{a_l'}^m}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^k}}^{b}(v_b^k)) \mid v_{a_l'}^m \in V_{a_1'}\} = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^{b}(v_b^k))$. Obviously, for all $v_{a_l'}^m \in V_{a_1'}$ and for all $v_b^{s'} \in V_b$,

$$\min\left(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}\left(v_{a_l'}\right), P_{U_{v_b^k}}^{b}\left(v_b^k\right)\right) \geq \min\left(P_{x',a_l'}^{v_{a_l'}^m}, P_{U_{\tau k}}^{a_l'}\left(v_{a_l'}^m\right), P_{U_{v_b^{s'}}}^{b}\left(v_b^{s'}\right)\right), \tag{6.14}$$

so we have $P_{x'}^b(v_b^k) \geq P_{x'}^b(v_b^{s'})$, that is, $P_{x'}^b(v_b^k)$ is the maximum probability of (6.13). □

According to Corollary 4.3, Theorem 6.1 shows that the inference conclusion can make x' and decision rule p consistent. For simplicity, in the rest of this section, we always assume that $P_{U_{\tau k}}^{a_l'}(v_{a_l'})$, $P_{U_{v_b^k}}^{b}(v_b^k)$, and $P_{x',a_l'}^{v_{a_l'}}$ are maximum probabilities, respectively.

Theorem 6.2. *If* $R_i = R_2$, $P_{U_{\tau k}}^{a_l'}(v_{a_l'}) \geq 0.5$, *and* $P_{U_{v_b^k}}^{b}(v_b^k) \geq 0.5$, *then* $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^{b}(v_b^k))$ *is the maximum probability of* (6.13).

Theorem 6.3. *If* $R_i = R_3$, *then* $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}) + P_{U_{v_b^k}}^{b}(v_b^k), 1 + P_{U_{\tau k}}^{a_l'}(v_{a_l'}) - P_{U_{v_b^k}}^{b}(v_b^k))$ *is the maximum probability.*

Theorem 6.4. *If* $R_i = R_4$, *then* $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}) + P_{U_{v_b^k}}^{b}(v_b^k))$ *is the maximum probability of* (6.13).

Proof. By (6.9) and (6.11), we know that $P_{x'}^b(v_b^k) = \max\{\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \min(1, 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) + P_{U_{v_b^k}}^{b}(v_b^k))) \mid v_{a_l'}^m \in V_{a_1'}\} = \max\{\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, 1, 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) + P_{U_{v_b^k}}^{b}(v_b^k)) \mid v_{a_l'}^m \in V_{a_1'}\} = \max\{\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) + P_{U_{v_b^k}}^{b}(v_b^k)) \mid v_{a_l'}^m \in V_{a_1'}\}$. When $v_{a_l'}^m \neq v_{a_l'}$, by $P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) \leq 0.5$, we know that $1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) + P_{U_{v_b^k}}^{b}(v_b^k) \geq 1$, so $\rho_{m,k} = P_{x',a_l'}^{v_{a_l'}^m} \leq \rho_k = \min(P_{x',a_l'}^{v_{a_l'}}, 1 -$

$P^{a_l'}_{U_{\tau k}}(v_{a_l'}) + P^{b}_{U_{v_b^k}}(v_b^k)) = P^{b}_{x'}(v_b^k)$. For all $s' \neq k$, $\rho_{s'} = \min(P^{v_{a_l'}}_{x',a_l'}, 1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}) + P^{b}_{U_{v_b^k}}(v_b^{s'}))$ and $\rho_{m,s'} = \min(P^{v_{a_l'}^m}_{x',a_l'}, 1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m) + P^{b}_{U_{v_b^k}}(v_b^{s'})) = P^{v_{a_l'}^m}_{x',a_l'}$. obviously, $\rho_k \geq \max\{\rho_{s'}, \rho_{m,s'}\} = P^{b}_{x'}(v_b^{s'})$, and $P^{b}_{x'}(v_b^k)$ is the maximum probability. □

Remark 6.5. According to (5.2), if $P^{x'}_{a_l'}(v_{a_l'})$ and $P^{x'}_{b}(v_b^k)$ are maximum probability assignments, respectively, then the probability decision rule is consistent. Theorems 6.1–6.4 show that we obtain maximum probability. Hence, using $R_1 - R_4$ ensures the consistency of decision rule.

Theorem 6.6. *If $R_i = R_5$, then $P^{b}_{x'}(v_b^k) = \min(P^{v_{a_l'}}_{x',a_l'}, \max(1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}) + P^{b}_{U_{v_b^k}}(v_b^k), 1 + P^{a_l'}_{U_{\tau k}}(v_{a_l'}) - P^{b}_{U_{v_b^k}}(v_b^k)))$ is the minimum probability of* (6.13).

Proof. By (6.10) and (6.11), we know that $P^{b}_{x'}(v_b^k) = \max\{\rho_{m,k} = \min(P^{v_{a_l'}^m}_{x',a_l'}, \max(\min(1, 1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m) + P^{b}_{U_{v_b^k}}(v_b^k)), \min(1, 1 + P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m) - P^{b}_{U_{v_b^k}}(v_b^k)))) \mid v_{a_l'}^m \in V_{a_1'}\}$. By $P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m) \leq 0.5$, we know that $\min(1, 1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m) + P^{b}_{U_{v_b^k}}(v_b^k)) = 1$, so $\rho_{m,k} = P^{v_{a_l'}^m}_{x',a_l'}$. $\rho_k = \min(P^{v_{a_l'}}_{x',a_l'}, \max(1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}) + P^{b}_{U_{v_b^k}}(v_b^k), 1 + P^{a_l'}_{U_{\tau k}}(v_{a_l'}) - P^{b}_{U_{v_b^k}}(v_b^k)))$, so we have $P^{b}_{x'}(v_b^k) = \rho_k$. For $P^{b}_{x'}(v_b^{s'})$, by $1 + P^{a_l'}_{U_{\tau k}}(v_{a_l'}) - P^{b}_{U_{v_b^k}}(v_b^{s'}) \geq 1$, we get for all $s' \neq k$, $P^{b}_{x'}(v_b^{s'}) = P^{v_{a_l'}}_{x',a_l'} \geq \rho_k = P^{b}_{x'}(v_b^k)$, $P^{b}_{x'}(v_b^k)$ is minimal probability of (6.13). □

Remark 6.7. According to (5.2), if $P^{x'}_{a_l'}(v_{a_l'})$ and $P^{x'}_{b}(v_b^k)$ are maximum probability assignments, respectively, then the probability decision rule is consistent. If we use R_5, Theorem 6.6 show that we obtain minimal probability. Hence, using R_5 cannot ensure the consistency of decision rule.

7. Conclusion

In this paper, we provide a kind of probability information system and probability decision tables to represent a degree of belief of "objects have attributes," formally, the kind of probability information system is extension of classical information system and a special case of interval-valued information system. Based on rough set theory, we discuss extraction of probability decision rules. Then we analyse consistency of probability decision rules. Finally, we provide a method to finish inference of probability decision rules.

Appendices

A. The Proof of Theorem 6.2

Proof. By (6.7) and (6.11), we know that $P^{b}_{x'}(v_b^k) = \max\{\rho_{m,k} = \min(P^{v_{a_l'}^m}_{x',a_l'}, \max(\min(P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m), P^{b}_{U_{v_b^k}}(v_b^k)), \min((1 - P^{a_l'}_{U_{\tau k}}(v_{a_l'}^m)), (1 - P^{b}_{U_{v_b^k}}(v_b^k))))) \mid v_{a_l'}^m \in V_{a_1'}\}$, where index

k of $\rho_{m,k}$ shows the k column of R_2, when $v_{a_l'}^m = v_{a_l'}$, noted by ρ_k. By $P_{U_{\tau k}}^{a_l'}(v_{a_l'}) \geq 0.5$, $P_{U_{v_b^k}}^b(v_b^k) \geq 0.5$, we get $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k)) \geq \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})), (1 - P_{U_{v_b^k}}^b(v_b^k)))$. So, $\rho_k = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k))$. By for all $v_{a_l'}^m \neq v_{a_l'}$, we have $P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) \leq P_{U_{\tau k}}^{a_l'}(v_{a_l'}) \leq 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)$, and $P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) \leq P_{U_{v_b^k}}^b(v_b^k)$, so $\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \max(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), 1 - P_{U_{v_b^k}}^b(v_b^k)))$ and obviously, $\rho_k \geq \rho_{m,k}$. So, we have $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k))$. By for all $v_b^{s'} \neq v_b^k$, we have $P_{x'}^b(v_b^{s'}) = \max\{\rho_{m,s'} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \max(\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^{s'}}}^b(v_b^{s'})), \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))))) \mid v_{a_l'}^m \in V_{a_1'}\}$. By $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^{s'}}}^b(v_b^{s'})) = P_{U_{v_b^{s'}}}^b(v_b^{s'}), \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))) = 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})$, we have $\rho_{s'} = \min(P_{x',a_l'}^{v_{a_l'}}, \max(P_{U_{v_b^{s'}}}^b(v_b^{s'}), 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})))$. Because for all $v_{a_l'}^m \neq v_{a_l'}$, $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^{s'}}}^b(v_b^{s'})) \leq \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'})))$, so $\rho_{m,s'} = \min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), (1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))) = P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)$, $P_{x'}^b(v_b^{s'}) = \max\{\rho_{s'}, \rho_{m,s'}\}$, and $P_{x'}^b(v_b^k) \geq P_{x'}^b(v_b^{s'})$, that is, $P_{x'}^b(v_b^k)$ is the maximum probability of (6.13). □

B. The Proof of Theorem 6.3

Proof. By (6.7) and (6.11), we know that $P_{x'}^b(v_b^k) = \max\{\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \max(\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^k}}^b(v_b^k)), \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^k}}^b(v_b^k))))) \mid v_{a_l'}^m \in V_{a_1'}\}$, where index k of $\rho_{m,k}$ shows the k column of R_2, when $v_{a_l'}^m = v_{a_l'}$, noted by ρ_k. By $P_{U_{\tau k}}^{a_l'}(v_{a_l'}) \geq 0.5$, $P_{U_{v_b^k}}^b(v_b^k) \geq 0.5$, we get $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k)) \geq \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})), (1 - P_{U_{v_b^k}}^b(v_b^k)))$. So, $\rho_k = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k))$. By for all $v_{a_l'}^m \neq v_{a_l'}$, we have $P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) \leq P_{U_{\tau k}}^{a_l'}(v_{a_l'}) \leq 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)$, and $P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m) \leq P_{U_{v_b^k}}^b(v_b^k)$, so $\rho_{m,k} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \max(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), 1 - P_{U_{v_b^k}}^b(v_b^k)))$, and obviously, $\rho_k \geq \rho_{m,k}$. So, we have $P_{x'}^b(v_b^k) = \min(P_{x',a_l'}^{v_{a_l'}}, P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^k}}^b(v_b^k))$. By for all $v_b^{s'} \neq v_b^k$, we have $P_{x'}^b(v_b^{s'}) = \max\{\rho_{m,s'} = \min(P_{x',a_l'}^{v_{a_l'}^m}, \max(\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^{s'}}}^b(v_b^{s'})), \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))))) \mid v_{a_l'}^m \in V_{a_1'}\}$. By $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}), P_{U_{v_b^{s'}}}^b(v_b^{s'})) = P_{U_{v_b^{s'}}}^b(v_b^{s'}), \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))) = 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})$, we have $\rho_{s'} = \min(P_{x',a_l'}^{v_{a_l'}}, \max(P_{U_{v_b^{s'}}}^b(v_b^{s'}), 1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'})))$. Because for all $v_{a_l'}^m \neq v_{a_l'}$, $\min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), P_{U_{v_b^{s'}}}^b(v_b^{s'})) \leq \min((1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'})))$, so $\rho_{m,s'} = \min(P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m), (1 - P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)), (1 - P_{U_{v_b^{s'}}}^b(v_b^{s'}))) = P_{U_{\tau k}}^{a_l'}(v_{a_l'}^m)$, $P_{x'}^b(v_b^{s'}) = \max\{\rho_{s'}, \rho_{m,s'}\}$, and $P_{x'}^b(v_b^k) \geq P_{x'}^b(v_b^{s'})$ that is, $P_{x'}^b(v_b^k)$ is the maximum probability of (6.13). □

Acknowledgments

This work is partially supported by Sichun Key Laboratory of Intelligent Network Information Processing (SGXZD1002-10), National Natural Science Foundation (61175055, 61105059), Sichuan Key Technology Research and Development Program (2012GZ0019, 2011FZ0051), and Liaoning Excellent Talents in University (LJQ2011116).

References

[1] O. Khozaimya, A. Al-Dhaherib, and A. M. M. Sharif UllahA, "Decision-making approach using point-cloud-based granular information," *Applied Soft Computing*, vol. 11, no. 10, pp. 2576–2586, 2011.
[2] E. Orlowska, *Incomplete Information: Rough Set Analysis*, vol. 13, Physica, Heidelberg, Germany, 1998.
[3] Z. Pei, D. Ruan, J. Liu, and Y. Xu, *Linguistic Values Based Intelligent Information Processing: Theory, Methods, and Application*, Atlantis press & World Scientific, 2009.
[4] G. Resconi and A. J. van der Wal, "Morphogenic neural networks encode abstract rules by data," *Information Sciences*, vol. 142, no. 1, pp. 249–273, 2006.
[5] R. R. Yager, "A framework for reasoning with soft information," *Information Sciences*, vol. 180, no. 8, pp. 1390–1406, 2010.
[6] C. Basaran, M. H. Suzer, K. D. Kang, and X. Liu, "Robust fuzzy CPU utilization control for dynamic workloads," *Journal of Systems and Software*, vol. 83, no. 7, pp. 1192–1204, 2010.
[7] M. Doumpos and C. Zopounidis, "Preference disaggregation and statistical learning for multicriteria decision support: a review," *European Journal of Operational Research*, vol. 209, no. 3, pp. 203–214, 2011.
[8] J. H. Luo, X. Liu, Y. Zhang, D. X. Ye, and Z. Xu, "Fuzzy trust recommendation based on collaborative filtering for mobile ad-hoc networks," in *Proceedings of the 33rd Annual IEEE Conference on Local Computer Networks*, Montreal, QC, Canada, 2008.
[9] Z. Pei, Y. Xu, D. Ruan, and K. Qin, "Extracting complex linguistic data summaries from personnel database via simple linguistic aggregations," *Information Sciences*, vol. 179, no. 10, pp. 2325–2332, 2009.
[10] Z. Pei, X. Liu, and L. Zou, "Extracting association rules based on intuitionistic fuzzy sets," *International Journal of Innovative Computing, Information and Control*, vol. 6, no. 6, pp. 2567–2580, 2010.
[11] J. Rao, Y. D. Wei, J. Y. Gong, and C. Z. Xu, "DynaQoS: Model-free self-tuning fuzzy control of virtualized resources for QoS provisioning," in *Proceedings of the IEEE 19th International Workshop on Quality of Service*, 2009.
[12] P. C. Xiong and Y. S. Fan, "QoS-aware web service selection by a synthetic weight," in *Proceedings of the 4th International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 2, pp. 632–637, Hainan, China, 2007.
[13] L. A. Zadeh, "A note on Z-numbers," *Information Sciences*, vol. 181, no. 14, pp. 2923–2932, 2011.
[14] L. A. Zadeh, "Generalized theory of uncertainty (GTU)—principal concepts and ideas," *Computational Statistics & Data Analysis*, vol. 51, no. 1, pp. 15–46, 2006.
[15] L. A. Zadeh, "Probability measures of fuzzy events," *Journal of Mathematical Analysis and Applications*, vol. 23, pp. 421–427, 1968.
[16] L. A. Zadeh, "Fuzzy sets," *Information and Computation*, vol. 8, pp. 338–353, 1965.
[17] L. Zou, Z. Pei, X. Liu, and Y. Xu, "Semantic of linguistic truth-valued intuitionistic fuzzy proposition calculus," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 12, pp. 4745–4752, 2009.
[18] R. R. Yager, "On the fusion of imprecise uncertainty measures using belief structures," *Information Sciences*, vol. 181, no. 15, pp. 3199–3209, 2011.
[19] R. R. Yager, "Including probabilistic uncertainty in fuzzy logic controller modeling using Dempster-Shafer theory," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 8, pp. 1221–1230, 1995.
[20] R. R. Yager, "On the Dempster-Shafer framework and new combination rules," *Information Sciences*, vol. 41, no. 2, pp. 93–137, 1987.
[21] D. Dubois, "Possibility theory and statistical reasoning," *Computational Statistics & Data Analysis*, vol. 51, no. 1, pp. 47–69, 2006.
[22] D. Dubois and H. Prade, "On possibility theory, formal concept analysis and granulation survey," *Applied and Computational Mathematics*, vol. 10, no. 1, pp. 10–19, 2011.
[23] A. P. Dempster, "New methods for reasoning towards posterior distributions based on sample data," *Annals of Mathematical Statistics*, vol. 37, pp. 355–374, 1966.

[24] D. Fixsen and R. P. S. Mahler, "The modified Dempster-Shafer approach to classification," *IEEE Transactions on Systems, Man, and Cybernetics Part A*, vol. 27, no. 1, pp. 96–104, 1997.
[25] Y. Tang and Y. Xu, "Application of fuzzy naive Bayes and a real-valued genetic algorithm in identification of fuzzy model," *Information Sciences*, vol. 169, no. 3-4, pp. 205–226, 2005.
[26] Y. Tang and J. Zheng, "Generalized Jeffrey's rule of conditioning and evidence combining rule for a priori probabilistic knowledge in conditional evidence theory," *Information Sciences*, vol. 176, no. 11, pp. 1590–1606, 2006.
[27] J. Yen, "Generalizing the Dempster-Shafer theory to fuzzy sets," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 3, pp. 559–570, 1990.
[28] Z. Pawlak, "Rough sets, decision algorithms and Bayes' theorem," *European Journal of Operational Research*, vol. 136, no. 1, pp. 181–189, 2002.
[29] Z. Pawlak, "Rough sets and intelligent data analysis," *Information Sciences*, vol. 147, no. 1–4, pp. 1–12, 2002.
[30] Z. Pawlak and A. Skowron, "Rough sets and Boolean reasoning," *Information Sciences*, vol. 177, no. 1, pp. 41–73, 2007.