

Inclusion degree with variable-precision model in analyzing inconsistent decision tables

Yu-Ru Syau¹ · Andrzej Skowron² · En-Bing Lin³ 

Received: 6 June 2016 / Accepted: 12 August 2016 / Published online: 18 August 2016
© Springer International Publishing Switzerland 2016

Abstract In this paper, we incorporate the concepts of evidence theory and variable precision rough set model (VP-model) to examine inconsistent decision tables. For the decision classes in a given decision table, we point out the relationship between approximation degree of dependency and the belief functions. We also introduce the notion of weakly discernable decision tables, and show how to obtain the discernibility threshold of a given weakly discernable decision table. We prove that weakly discernibility is a necessary condition of relative consistency in a given decision table.

Keywords Rough sets · Belief functions · Variable precision rough set models · Inclusion degrees · Inconsistent decision tables · Relative discernibility

1 Introduction

Nowadays, we live in a world submerged with more information than ever before, and the information is growing faster. On the other hand, we encounter imperfect decision-relevant information in some situations. Although big information systems appear everywhere, data has begun to accumulate to the point where a new and special

event is taking place. Today we deal with big data which does not necessarily mean good data. And that, as an increasing number of experts are saying more insistently, means big data does not automatically yield good analytics. If the data is imperfect, out of context or otherwise contaminated, it can lead to decisions that could undermine the competitiveness of an enterprise or damage the personal lives of individuals. Therefore, knowledge representation plays an important role in dealing with many aspects of problem solving. This includes handling imperfect knowledge. More precisely, we study inconsistent decision tables that can further analyze imperfect knowledge. One of the most important concepts an intelligent system needs to understand is the concept of knowledge. It may or may not be perfect. Also, one wants to know what knowledge is needed to achieve particular goals, and how that knowledge can be obtained. So, one of the important problems along this line is to seek an appropriate approach to analyze imperfect knowledge. The problem related to imperfect knowledge or an inconsistent decision table has been investigated by many researchers in different areas. Our idea is to apply a granular computing approach (Pedrycz and Chen 2014, 2015; Skowron et al. 2016), rough set theory, to cope with imperfect data analysis. Rough set theory is proposed by Pawlak (1982) and is based on fundamental set theory and has been applied extensively in many different aspects of theoretical and applied research areas.

Regarding the inclusion degree, several papers have done various aspects of research. Polkowski and Skowron (1996) proposed rough mereology as a foundation for approximate reasoning about complex objects. Yao et al. (2015) adopted the Bayesian decision-theoretic analysis to provide a systematic method for determining the precision parameters by using some notions of costs and risks. Zhang

✉ En-Bing Lin
enbing.lin@cmich.edu

¹ Department of Information Management, National Formosa University, Huwei, Yunlin 63201, Taiwan

² Institute of Mathematics, Warsaw University, Banacha 2, 02-097 Warsaw, Poland

³ Department of Mathematics, Central Michigan University, Mt. Pleasant, MI 48859, USA

and Leung (1996) proposed a generalized notion of inclusion degree in the context of a partially ordered set. Gomolinska (2008) obtained two rough inclusion functions (RIFs in short), different from the standard one and from each other. With every RIF, she associated a mapping which is in some sense complementary to it. She used these complementary mappings (co-RIFs) to define certain metrics. While the distance functions may directly be used to measure the degree of dissimilarity of sets of objects, their complementary mappings are useful in measuring the degree of mutual similarity of sets.

In our work, we use a table to represent some collections of data in the real world. Using the concept of lower and upper approximations in rough set theory together with the variation of parameter to describe the characterizations of the objects, we are able to come up with some variable precision models through the given decision information. It was first introduced by Ziarko (1993), the tool of variable precision and later was used by many others. However, this is the first time, we derive better analysis of equivalence classes by using infimum and supremum of inclusion degrees of equivalence classes in a given set with smaller (or greater) values than 0.5 instead of minimum and maximum, respectively. In fact, the family of inclusion degrees of equivalence classes in a given set with smaller (or greater) values than 0.5 may be empty. In what follows, we first provide some preliminary backgrounds in rough set theory and variable precision model (VP-model), under decision tables, and the definitions of positive region, boundary region and relative reduct for a decision table in the VP-model in Sect. 2. In Sect. 3, we provide the relationship between the approximation degree of dependency and the belief function in evidence theory. We then provide the intrinsic properties of positive regions and upper approximations of decision classes using the threshold of the decision table in Sect. 4. We also present several examples in Sect. 4. We show a very useful tool in understanding the condition classes in positive regions in Sect. 5. We show a clear relationship between consistency and positive regions of a given decision table in Sect. 6. Several concluding remarks are presented in Sect. 7.

2 Preliminaries

Let U be a finite and nonempty set, known as the universe of discourse. We use the symbol “ \subseteq ” (“ \subset ”) to denote set inclusion (strict set inclusion, respectively). The cardinality of a set $S \subseteq U$, denoted $|S|$, is the number of elements in S . The power set of U is the collection $2^U = \{S \mid S \subseteq U\}$.

The inclusion degree of a nonempty set $X \subseteq U$ in a set $Y \subseteq U$ is defined as

$$I(X, Y) = \frac{|X \cap Y|}{|X|}. \quad (1)$$

Let us define $Pr : 2^U \rightarrow [0, 1]$ as follows:

$$Pr(X) = \frac{|X|}{|U|}, \quad \forall X \subseteq U. \quad (2)$$

The greatest lower bound (inf or infimum) and least upper bound (sup or supremum) of a subset S of the unit interval $[0, 1]$ will be denoted by $\inf S$ and $\sup S$, respectively. From the definitions of inf and sup, we have

$$\inf \emptyset = 1 \quad \text{and} \quad \sup \emptyset = 0. \quad (3)$$

Remark If $\inf S \in S$, then we also denote it by $\min S$ and call it the minimum of S , and if $\sup S \in S$, then we also denote it by $\max S$ and call it the maximum of S . However, nothing is an element of the empty set and the empty set would not have a minimum. In fact, we will deal with the family of inclusion degrees of equivalence classes in a given set with smaller (or greater) values than 0.5, which may be empty. We therefore introduce infimum and supremum to include all the possibilities.

2.1 VP-models under decision tables

The knowledge representation in the rough set model is often structured in a decision table which is a 4-tuple $(U, Q = C \cup D, V, f)$, where U is a nonempty finite universe, C is a nonempty finite set of condition attributes, D is a nonempty finite set of decision attributes, $C \cap D = \emptyset$, $V = \bigcup_{q \in Q} V_q$ and V_q is a domain of the attribute q , and

$$f : U \times Q \rightarrow V$$

is an information function such that $f(x, q) \in V_q$ for every $x \in U$ and $q \in Q$.

Every nonempty subset P of condition attributes C , or decision attributes D , generates an equivalence relation on U , denoted by \hat{P} and defined as follows (Pawlak 1988).

$$\hat{P} = \{(x, y) \in U \times U \mid \forall q \in P, f(x, q) = f(y, q)\}. \quad (4)$$

Let $P^* = \{P_1, P_2, \dots, P_{|P^*|}\}$ denote the partition on U induced by equivalence relation \hat{P} . Each member of D^* will be called a decision class. The decision table $(U, C \cup D, V, f)$ is consistent if $\hat{C} \subseteq \hat{D}$; otherwise, the decision table is inconsistent (Pawlak 1987).

For any $X \subseteq U$, we can define the P -lower approximation $\underline{P}(X)$ and P -upper approximation $\overline{P}(X)$ of X , in the classical rough set model as follows (An et al. 1996):

$$\underline{P}(X) = \cup\{P_i \in P^* \mid P_i \subseteq X\} = \cup\{P_i \in P^* \mid I(P_i, X) = 1\}, \quad (5)$$

$$\bar{P}(X) = \cup\{P_i \in P^* \mid P_i \cap X \neq \emptyset\} = \cup\{P_i \in P^* \mid I(P_i, X) > 0\}. \tag{6}$$

Let β be a parameter such that $0.5 < \beta \leq 1$. For $P_i \in P^*$ and $X \subseteq U$, we define

$$P_i \subseteq^\beta X \quad \text{if and only if } I(P_i, X) \geq \beta, \tag{7}$$

$$P_i \cap^\beta X \neq \emptyset \quad \text{if and only if } I(P_i, U - X) < \beta. \tag{8}$$

Then, we can define the P^β -lower approximation $\underline{P}^\beta(X)$ and P^β -upper approximation $\bar{P}^\beta(X)$ of X , in the VP-model under the threshold β , as follows (An et al. 1996):

$$\underline{P}^\beta(X) = \cup\{P_i \in P^* \mid P_i \subseteq^\beta X\} = \cup\{P_i \in P^* \mid I(P_i, X) \geq \beta\}, \tag{9}$$

$$\begin{aligned} \bar{P}^\beta(X) &= \cup\{P_i \in P^* \mid P_i \cap^\beta X \neq \emptyset\} \\ &= \cup\{P_i \in P^* \mid I(P_i, X) > 1 - \beta\}. \end{aligned} \tag{10}$$

Evidently, we have

$$\underline{P}^\beta(\emptyset) = \bar{P}^\beta(\emptyset) = \emptyset \quad \text{and} \quad \underline{P}^\beta(U) = \bar{P}^\beta(U) = U, \tag{11}$$

$$\underline{P}(X) = \underline{P}^1(X) \subseteq \underline{P}^\beta(X) \subseteq \bar{P}^\beta(X) \subseteq \bar{P}^1(X) = \bar{P}(X), \tag{12}$$

$$\bar{P}^\beta(X) = U - \underline{P}^\beta(U - X). \tag{13}$$

2.2 Positive region and reduct

Approximations of equivalence classes are obtained by using these concepts of lower approximations, upper approximations, boundary regions and positive regions. We extend the concepts to VP models by providing similar extensions via parameters. To obtain optimal approximations, we introduce reducts of decision classes.

Definition 1 Given a decision table $(U, C \cup D, V, f)$ and a parameter $\beta \in (0.5, 1]$, let $D^* = \{D_1, D_2, \dots, D_{|D^*|}\}$.

- (Inuiguchi 2005; Zhou and Miao 2011; Ziarko 1993) The quality of classification, or β -approximation degree of dependency, of D w.r.t. a nonempty $A \subseteq C$ is defined as:

$$\gamma_A^\beta(D) = \frac{|\text{POS}_A^\beta(D^*)|}{|U|} = \sum_{j=1}^{|D^*|} \frac{|\underline{A}^\beta(D_j)|}{|U|} \tag{14}$$

where

$$\text{POS}_A^\beta(D^*) = \bigcup_{D_j \in D^*} \underline{A}^\beta(D_j) \tag{15}$$

is called the β -positive region of D^* w.r.t. A .

- A nonempty $B \subseteq C$ is called a β -reduct of C w.r.t. D iff it is a minimum subset of C such that $\gamma_B^\beta(D) = \gamma_C^\beta(D)$.

Here “iff” is Halmos’ convention for “if and only if”.

3 Evidence theory and VP-models in decision tables

Let $(U, C \cup D, V, f)$ be a decision table. To each nonempty $A \subseteq C$, we associate a basic probability assignment (bpa) (Lin 1998; Shafer 1976; Skowron and Grzymala-Busse 1994) $m_A : 2^U \rightarrow [0, 1]$ defined as follows:

$$m_A(X) = \begin{cases} Pr(X) & \text{if } X \in A^* \\ 0 & \text{elsewhere.} \end{cases}$$

Then, by replacing “ \subseteq ” and “ \cap ” with “ \subseteq^β ” and “ \cap^β ”, respectively, in the definitions of the belief and plausibility function generated by the bpa m_A , we obtain the generalized notions of β -belief and β -plausibility functions generated by the bpa m_A as follows (Syau and Lin 2015): for any $X \subseteq U$,

$$\text{Bel}_A^\beta(X) = \sum_{A_i \in A^* : A_i \subseteq^\beta X} m_A(A_i), \quad \text{Pl}_A^\beta(X) = \sum_{A_i \in A^* : A_i \cap^\beta X \neq \emptyset} m_A(A_i). \tag{16}$$

In addition, motivated by our earlier paper (Syau and Lin 2015), we also have

$$\begin{aligned} \text{Bel}_A^\beta(X) &= \sum_{A_i \in A^* : A_i \subseteq^\beta X} m_A(A_i) = \sum_{A_i \in A^* : A_i \subseteq^\beta X} Pr(A_i) \\ &= \sum_{A_i \in A^* : A_i \subseteq^\beta X} \frac{|A_i|}{|U|} = \frac{|\cup\{A_i \in A^* \mid A_i \subseteq^\beta X\}|}{|U|} \\ &= Pr(\underline{A}^\beta(X)), \end{aligned} \tag{17}$$

$$\begin{aligned} \text{Pl}_A^\beta(X) &= \sum_{A_i \in A^* : A_i \cap^\beta X \neq \emptyset} m_A(A_i) = \sum_{A_i \in A^* : A_i \cap^\beta X \neq \emptyset} Pr(A_i) \\ &= \sum_{A_i \in A^* : A_i \cap^\beta X \neq \emptyset} \frac{|A_i|}{|U|} = \frac{|\cup\{A_i \in A^* \mid A_i \cap^\beta X \neq \emptyset\}|}{|U|} \\ &= Pr(\bar{A}^\beta(X)). \end{aligned} \tag{18}$$

The duality

$$\text{Pl}_A^\beta(X) = 1 - \text{Bel}_A^\beta(U - X) \tag{19}$$

follows immediately from (13), (17) and (18).

Let $D^* = \{D_1, D_2, \dots, D_{|D^*|}\}$. According to (2), (14) and (17), we obtain

$$\begin{aligned} \gamma_A^\beta(D) &= \sum_{j=1}^{|D^*|} \frac{|\underline{A}^\beta(D_j)|}{|U|} = \sum_{j=1}^{|D^*|} Pr(\underline{A}^\beta(D_j)) \\ &= \sum_{j=1}^{|D^*|} Bel_A^\beta(D_j). \end{aligned} \tag{20}$$

This leads to the following:

Theorem 1 Given a decision table $(U, C \cup D, V, f)$ and a parameter $\beta \in (0.5, 1]$, let $D^* = \{D_1, D_2, \dots, D_{|D^*|}\}$. For any nonempty $A \subseteq C$, we have

$$\gamma_A^\beta(D) = \gamma_C^\beta(D) \iff \sum_{j=1}^{|D^*|} Bel_A^\beta(D_j) = \sum_{j=1}^{|D^*|} Bel_C^\beta(D_j). \tag{21}$$

4 Relative discernibility of decision classes

As an immediate consequence of (9) and (10), we have the following:

Lemma 1 Given a decision table $(U, C \cup D, V, f)$ and a parameter $\beta \in (0.5, 1]$, let $D^* = \{D_1, D_2, \dots, D_{|D^*|}\}$. For every decision class $D_j \in D^*$, we have

$$\underline{C}^\beta(D_j) \subseteq \underline{C}^{\beta'}(D_j) \subseteq \overline{C}^{\beta'}(D_j) \subseteq \overline{C}^\beta(D_j), \quad \forall \beta' \in (0.5, \beta]. \tag{22}$$

Ziarko (1993) states that a decision class $D_j \in D^*$ is said to be β -discernable if

$$\underline{C}^\beta(D_j) = \overline{C}^\beta(D_j). \tag{23}$$

According to Ziarko (1993), a decision class which is not discernable for every $\beta \in (0.5, 1]$ will be called absolutely indiscernible. A decision class $D_k \in D^*$ is absolutely indiscernible iff its absolute boundary

$$M(D_k) = \cup\{C_i \in C^* : I(C_i, D_k) = 0.5\} \neq \emptyset. \tag{24}$$

A decision class which is not absolutely indiscernible will be referred to as weakly discernable. More precisely, a decision class $D_j \in D^*$ is weakly discernable iff $\underline{C}^\beta(D_j) = \overline{C}^\beta(D_j)$ for some $\beta \in (0.5, 1]$. The greatest value of β which makes D_j discernable is referred to as discernibility threshold. Ziarko (1993) also gives a proposition which provides the computation procedures for discernibility thresholds. Considering this proposition together with inclusion degrees and the remark in Sect. 2, we obtain the following.

Lemma 2 Given a decision table $(U, C \cup D, V, f)$, let $C^* = \{C_1, C_2, \dots, C_{|C^*|}\}$ and $D^* = \{D_1, D_2, \dots, D_{|D^*|}\}$. If a decision class $D_j \in D^*$ is weakly discernable and its discernibility threshold is equal to ζ_j . Then

$$\begin{aligned} \zeta_j &= \min\{\eta_j, \lambda_j\}, \quad \text{where} \\ \eta_j &= \inf\{I(C_i, D_j) \mid C_i \in C^* \ \& \ I(C_i, D_j) > 0.5\}, \\ \lambda_j &= 1 - \sup\{I(C_i, D_j) \mid C_i \in C^* \ \& \ I(C_i, D_j) < 0.5\}. \end{aligned} \tag{25}$$

$$\tag{26}$$

Example 1 Let us consider an example of an inconsistent decision table $(U, C \cup D, V, f)$ as shown in Table 1, where $C = \{a, b, c\}$ is the set of condition attributes, $D = \{d\}$ is the decision attribute.

From this table, we have

1. $U = \{x_1, x_2, \dots, x_8\}$.
2. $V_a = V_b = V_c = V_d = \{1, 2\}$.
3. $C^* = \{C_1, C_2, C_3\}$, $D^* = \{D_1, D_2\}$, where

$$C_1 = \{x_1, x_5, x_7\}, \quad C_2 = \{x_2, x_3, x_6, x_8\}, \quad C_3 = \{x_4\},$$

$$D_1 = \{x_1, x_2, x_4, x_6, x_7, x_8\}, \quad D_2 = \{x_3, x_5\}.$$

According to (1), we have

$$\begin{aligned} I(C_1, D_1) &= \frac{2}{3}, \quad I(C_1, D_2) = \frac{1}{3}, \\ I(C_2, D_1) &= \frac{3}{4}, \quad I(C_2, D_2) = \frac{1}{4}, \\ I(C_3, D_1) &= 1, \quad I(C_3, D_2) = 0. \end{aligned} \tag{27}$$

This gives the following:

- (i) $\{I(C_i, D_1) \mid C_i \in C^* \ \& \ I(C_i, D_1) < 0.5\} = \emptyset,$

$$\frac{2}{3} = \min\{I(C_i, D_1) \mid C_i \in C^* \ \& \ I(C_i, D_1) > 0.5\},$$

$$\underline{C}^\beta(D_1) = \underline{C}^{\frac{2}{3}}(D_1) = U, \quad \forall \beta \in \left(0.5, \frac{2}{3}\right], \tag{28}$$

$$\overline{C}^\beta(D_1) = U, \quad \forall \beta \in (0.5, 1]. \tag{29}$$

It follows from (28) and (29) that the decision class D_1 is weakly discernable and its discernibility threshold is equal to $\frac{2}{3}$. This example clearly

Table 1 Exemplary decision

U	a	b	c	d
x_1	1	2	1	1
x_2	2	2	1	1
x_3	2	2	1	2
x_4	1	1	2	1
x_5	1	2	1	2
x_6	2	2	1	1
x_7	1	2	1	1
x_8	2	2	1	1

validates that our method works and indicates that why we use supremum instead of maximum in (26).

$$(ii) \quad \{I(C_i, D_2) \mid C_i \in C^* \ \& \ I(C_i, D_1) > 0.5\} = \emptyset, \quad \frac{2}{3} \\ = 1 - \max \{I(C_i, D_2) \mid C_i \in C^* \ \& \ I(C_i, D_1) < 0.5\}, \\ \underline{C}^\beta(D_2) = \emptyset, \quad \forall \beta \in (0.5, 1], \quad (30)$$

$$\overline{C}^\beta(D_2) = \emptyset, \quad \forall \beta \in (0.5, \frac{2}{3}]. \quad (31)$$

It follows from (30) and (31) that the decision class D_2 is weakly discernible and its discernibility threshold is equal to $\frac{2}{3}$. Similarly to part (i), this example clearly validates that our method works and indicates that why we use infimum instead of minimum in (25).

In what follows, we consider a given decision tale, $(U, C \cup D, V, f)$. Let

$$C^* = \{C_1, C_2, \dots, C_{|C^*|}\}, \quad D^* = \{D_1, D_2, \dots, D_{|D^*|}\}, \\ M(D_j) = \cup\{C_i \in C^* : I(C_i, D_j) = 0.5\}, \\ H(D_j) = \cup\{C_i \in C^* : I(C_i, D_j) \geq 0.5\}, \\ \eta_j = \inf\{I(C_i, D_j) \mid C_i \in C^* \ \& \ I(C_i, D_j) > 0.5\}, \\ \lambda_j = 1 - \sup\{I(C_i, D_j) \mid C_i \in C^* \ \& \ I(C_i, D_j) < 0.5\}, \\ \zeta_j = \min\{\eta_j, \lambda_j\}, \\ \eta = \min\{\eta_1, \eta_2, \dots, \eta_{|D^*|}\}, \quad \lambda = \min\{\lambda_1, \lambda_2, \dots, \lambda_{|D^*|}\}, \\ \zeta = \min\{\eta, \lambda\} = \min\{\zeta_1, \zeta_2, \dots, \zeta_{|D^*|}\}. \quad (32)$$

Then, by Lemma 2 and Example 1, we obtain the following:

Lemma 3 *Given a decision table $(U, C \cup D, V, f)$, and the notations η, λ defined in (32), for each $D_j \in D^*$, we have*

$$\underline{C}^\beta(D_j) = \underline{C}^{\eta_j}(D_j), \quad \forall \beta \in (0.5, \eta_j] \quad (33)$$

$$\overline{C}^\beta(D_j) = \overline{C}^{\lambda_j}(D_j) = H(D_j), \quad \forall \beta \in (0.5, \lambda_j] \quad (34)$$

If $M(D_j) = \cup\{C_i \in C^* : I(C_i, D_j) = 0.5\} = \emptyset$, then the decision class D_j is weakly discernible and its discernibility threshold is equal to ζ_j . That is,

$$\underline{C}^\beta(D_j) = \overline{C}^\beta(D_j), \quad \forall \beta \in (0.5, \zeta_j]. \quad (35)$$

Considering an improvement of Cheng et al. (2015), we define relative discernibility of decision tables as follows.

Definition 2 A decision table $(U, C \cup D, V, f)$ is said to be weakly discernible iff all its decision classes $D_1, D_2, \dots, D_{|D^*|}$ are weakly discernible, or equivalently, iff for each $D_j \in D^*$,

$$M(D_j) = \cup\{C_i \in C^* : I(C_i, D_j) = 0.5\} = \emptyset. \quad (36)$$

The greatest value of β which makes the decision table $(U, C \cup D, V, f)$ discernable will be referred to as discernibility threshold.

Theorem 2 *Given a decision table $(U, C \cup D, V, f)$, and the notations η, λ defined in (32), we have*

1. $\text{POS}_C^\beta(D^*) = \text{POS}_C^\eta(D^*)$, $\forall \beta \in (0.5, \eta]$.
2. for each $D_j \in D^*$,

$$\overline{C}^\beta(D_j) = \overline{C}^\lambda(D_j), \quad \forall \beta \in (0.5, \lambda].$$

If the decision table is weakly discernible, then its discernibility threshold is equal to $\zeta = \min\{\eta, \lambda\}$. That is,

$$\text{POS}_C^\beta(D^*) = \bigcup_{D_j \in D^*} \underline{C}^\beta(D_j) = \bigcup_{D_j \in D^*} \overline{C}^\beta(D_j), \quad \forall \beta \in (0.5, \zeta], \quad (37)$$

or equivalently, for each $D_j \in D^*$,

$$\overline{C}^\beta(D_j) = \underline{C}^\beta(D_j), \quad \forall \beta \in (0.5, \zeta]. \quad (38)$$

To illustrate the above concepts, we present below an example of a weakly discernible decision table (Table 2).

Example 2 Let us consider a decision table $(U, C \cup D, V, f)$ as shown in Table 1, where $C = \{a, b, c\}$ is the set of condition attributes, $D = \{d\}$ is the decision attribute. From this table, we have

1. $U = \{x_1, x_2, \dots, x_{11}\}$.
2. $V_a = V_b = V_c = \{1, 2\}$, $V_d = \{1, 2, 3\}$.
3. $C^* = \{C_1, C_2, C_3\}$, $D^* = \{D_1, D_2, D_3\}$, where

Table 2 Exemplary decision

U	a	b	c	d
x_1	1	1	1	1
x_2	1	2	2	2
x_3	1	1	2	2
x_4	1	2	2	1
x_5	1	1	1	2
x_6	1	2	2	1
x_7	1	1	1	3
x_8	1	1	2	3
x_9	1	1	2	3
x_{10}	1	2	2	1
x_{11}	1	1	2	3

$$\begin{aligned} C_1 &= \{x_1, x_5, x_7\}, & C_2 &= \{x_2, x_4, x_6, x_{10}\}, \\ C_3 &= \{x_3, x_8, x_9, x_{11}\}, & D_1 &= \{x_1, x_4, x_6, x_{10}\}, \\ D_2 &= \{x_2, x_3, x_5\}, & D_3 &= \{x_7, x_8, x_9, x_{11}\}. \end{aligned}$$

According to (1), we obtain

$$\begin{aligned} I(C_1, D_1) &= \frac{1}{3}, & I(C_1, D_2) &= \frac{1}{3}, & I(C_1, D_3) &= \frac{1}{3}, \\ I(C_2, D_1) &= \frac{3}{4}, & I(C_2, D_2) &= \frac{1}{4}, & I(C_2, D_3) &= 0, \\ I(C_3, D_1) &= 0, & I(C_3, D_2) &= \frac{1}{4}, & I(C_3, D_3) &= \frac{3}{4}. \end{aligned} \quad (39)$$

Then, according to (5) and (6), we obtain

$$\underline{C}^\beta(D_1) = \begin{cases} C_2, & \text{if } 0.5 < \beta \leq \frac{3}{4} \\ \emptyset, & \text{if } \frac{3}{4} < \beta \leq 1, \end{cases} \quad (40)$$

$$\underline{C}^\beta(D_2) = \emptyset, \quad \forall \beta \in (0.5, 1] \quad (41)$$

$$\underline{C}^\beta(D_3) = \begin{cases} C_3, & \text{if } 0.5 < \beta \leq \frac{3}{4} \\ \emptyset, & \text{if } \frac{3}{4} < \beta \leq 1, \end{cases} \quad (42)$$

$$\overline{C}^\beta(D_1) = \begin{cases} C_2, & \text{if } 0.5 < \beta \leq \frac{2}{3} \\ C_1 \cup C_2, & \text{if } \frac{2}{3} < \beta \leq 1, \end{cases} \quad (43)$$

$$\overline{C}^\beta(D_2) = \begin{cases} \emptyset, & \text{if } 0.5 < \beta \leq \frac{2}{3} \\ C_1, & \text{if } \frac{2}{3} < \beta \leq \frac{3}{4} \\ U, & \text{if } \frac{3}{4} < \beta \leq 1, \end{cases} \quad (44)$$

and

$$\overline{C}^\beta(D_3) = \begin{cases} C_3, & \text{if } 0.5 < \beta \leq \frac{2}{3} \\ C_1 \cup C_3, & \text{if } \frac{2}{3} < \beta \leq 1. \end{cases} \quad (45)$$

Equations (40)–(45) give

$$\overline{C}^\beta(D_1) = \underline{C}^\beta(D_1) = C_2, \quad \forall \beta \in \left(0.5, \frac{2}{3}\right], \quad (46)$$

$$\overline{C}^\beta(D_2) = \underline{C}^\beta(D_2) = \emptyset, \quad \forall \beta \in \left(0.5, \frac{2}{3}\right], \quad (47)$$

$$\overline{C}^\beta(D_3) = \underline{C}^\beta(D_3) = C_3, \quad \forall \beta \in \left(0.5, \frac{2}{3}\right], \quad (48)$$

and

$$\begin{aligned} \text{POS}_C^\beta(D^*) &= \bigcup_{D_j \in D^*} \underline{C}^\beta(D_j) = \text{POS}_C^{\frac{2}{3}}(D^*) \\ &= C_2 \cup C_3, \quad \forall \beta \in \left(0.5, \frac{3}{4}\right]. \end{aligned} \quad (49)$$

According to (32) and (39), we have

$$\begin{aligned} \eta_1 &= \frac{3}{4}, & \lambda_1 &= 1 - \frac{1}{3} = \frac{2}{3}, & \zeta_1 &= \min\{\eta_1, \lambda_1\} = \frac{2}{3}, \\ \eta_2 &= \inf \emptyset = 1, & \lambda_2 &= 1 - \frac{1}{3} = \frac{2}{3}, & \zeta_2 &= \min\{\eta_2, \lambda_2\} = \frac{2}{3}, \\ \eta_3 &= \frac{3}{4}, & \lambda_3 &= 1 - \frac{1}{3} = \frac{2}{3}, & \zeta_3 &= \min\{\eta_3, \lambda_3\} = \frac{2}{3}, \\ \eta &= \min\{\eta_1, \eta_2, \eta_3\} = \frac{3}{4}, & \lambda &= \min\{\lambda_1, \lambda_2, \lambda_3\} = \frac{2}{3}, \\ \zeta &= \min\{\eta, \lambda\} = \min\{\zeta_1, \zeta_2, \zeta_3\} = \frac{2}{3}. \end{aligned} \quad (50)$$

This example validates the results of Lemma 3 and Theorem 2.

5 Inclusion degrees in computing β -positive region

Let $\beta \in (0.5, 1]$. As it is indicated in (14) that

$$\underline{C}^\beta(D_j) \cap \underline{C}^\beta(D_k) = \emptyset, \quad \forall D_j, D_k \in D^* \quad (j \neq k). \quad (51)$$

Notice that for each $C_i \in C^*$, we have

$$\begin{aligned} I(C_i, D_1) + I(C_i, D_2) + \cdots + I(C_i, D_{|D^*|}) \\ = \frac{|C_i \cap D_1|}{|C_i|} + \frac{|C_i \cap D_2|}{|C_i|} + \cdots + \frac{|C_i \cap D_{|D^*|}|}{|C_i|} = 1. \end{aligned} \quad (52)$$

This, combined with (9) and (51), leads to the following:

Lemma 4 Let $C_i \in C^*$.

- For $\beta \in (0.5, 1]$, one and only one of the following two cases: case 1. $C_i \subseteq \text{POS}_C^\beta(D^*)$, or case 2. $C_i \cap \text{POS}_C^\beta(D^*) = \emptyset$ must occur.
- $C_i \subseteq \text{POS}_C^\beta(D^*)$ for some $\beta \in (0.5, 1]$ iff there exists (one and only one) $D_j \in D^*$ with $I(C_i, D_j) > 0.5$. In this case, we have

$$I(C_i, D_k) < 0.5, \quad \forall D_k \in D^* - \{D_j\}. \quad (53)$$
- If $I(C_i, D_j) \leq 0.5$ for all $D_j \in D^*$, or if there exists $D_j \in D^*$ with $I(C_i, D_j) = 0.5$, then we must have

$$C_i \cap \text{POS}_C^\beta(D^*) = \emptyset, \quad \forall \beta \in (0.5, 1].$$

6 Relative consistency of decision tables

Following from Example 2, the β -positive regions, $POS_C^\beta(D^*)$ for all $\beta \in (0.5, 1]$, of the weakly discernable decision table can be expressed as:

$$POS_C^\beta(D^*) = \begin{cases} C_2 \cup C_3 & \text{if } \beta \in (0.5, \frac{3}{4}], \\ \emptyset & \text{if } \beta \in (\frac{3}{4}, 1]. \end{cases}$$

We indeed have $POS_C^\beta(D^*) \subset U$ for all $\beta \in (0.5, 1]$. We have thus seen that weakly discernibility of a decision table $(U, C \cup D, V, f)$ is not able to ensure that $POS_C^\beta(D^*) = U$. This gives a counterexample to Proposition 3.2 of Cheng et al. (2015). The above observation leads to the following definition:

Definition 3 A decision table $(U, C \cup D, V, f)$ is said to be β -consistent ($\beta \in (0.5, 1]$) if its β -positive region of D^* w.r.t. C equals the universe U . That is,

$$POS_C^\beta(D^*) = U. \tag{54}$$

A decision table which is not consistent for every $\beta \in (0.5, 1]$ will be called absolutely inconsistent.

Using Item 3 of Lemma 4, it can be easily seen that if a decision table $(U, C \cup D, V, f)$ is absolutely inconsistent iff there exists $C_i \in C^*$ such that

$$I(C_i, D_j) \leq 0.5, \quad \forall D_j \in D^*, \tag{55}$$

or equivalently,

$$C_i \cap POS_C^\beta(D^*) = \emptyset, \quad \forall \beta \in (0.5, 1]. \tag{56}$$

A decision table which is not absolutely inconsistent will be referred to as weakly consistent. More precisely, a decision table $(U, C \cup D, V, f)$ is weakly consistent iff for each $C_i \in C^*$ there exists (one and only one) $D_j \in D^*$ with $I(C_i, D_j) > 0.5$. The greatest value of β which makes $(U, C \cup D, V, f)$ consistent will be referred to as consistency threshold; hence, by Theorem 2, we obtain the following:

Theorem 3 Given a weakly consistent decision table $(U, C \cup D, V, f)$, and the notation η defined in (32), we have

$$POS_C^\beta(D^*) = POS_C^\eta(D^*) = U, \quad \forall \beta \in (0.5, \eta]. \tag{57}$$

According to Theorem 3 and (22), we have immediate consequences as follows.

Corollary 1 Given a weakly consistent decision table $(U, C \cup D, V, f)$, and the notations η, λ, ζ defined in (32), we have

$$POS_C^\beta(D^*) = U = \bigcup_{D_j \in D^*} \overline{C}^\beta(D_j), \quad \forall \beta \in (0.5, \zeta]. \tag{58}$$

Example 3 From Example 1, we have

$$POS_C^\beta(D^*) = U, \quad \forall \beta \in \left(0.5, \frac{2}{3}\right].$$

7 Concluding remarks

We have shown a close relationship between β -approximation degree of dependency and β -belief function of a given decision table. We further characterized a decision table by its positive regions with certain thresholds. We indicated how one finds a threshold for weakly discernable decision tables. In particular, positive regions with some thresholds provide a clear description for a weakly consistent decision table. By finding inclusion degrees, we are able to find thresholds and give rise to the characterization of decision tables. Our examples have demonstrated some good understandings of our theory in small scales. In fact, the approximation degree plays an interesting role in upper and lower approximations. This paper also showed the relationships between approximation degrees and discernibility thresholds. We have extended and improved other’s results (Ziarko 1993). It is anticipated to derive some algorithms for a large decision table. It would be interesting to connect with mutual entropy and Bayesian considerations (Düntsch and Gediga 2015) which may give rise to some algorithm to determine an approximately optimal set of predicting attributes. Future applications of our theory will include incorporating the inclusion degree in analyzing inconsistent decision tables on a Hadoop-based distributed platform.

Acknowledgments The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

References

An A, Shan N, Chan C, Cercone N, Ziarko W (1996) Discovering rules for water demand prediction: an enhanced rough-set approach. *Eng Appl Artif Intell* 9(6):645–653

Beynon M (2001) Reducts within the variable precision rough sets model: a further investigation. *Eur J Oper Res* 134:592–605

Beynon M, Driffield N (2005) An illustration of variable precision rough sets model: an analysis of the findings of the UK Monopolies and Mergers Commission. *Comput Oper Res* 32:1739–1759

Beynon M, Peel M (2001) Variable precision rough set theory and data discretisation: an application to corporate failure prediction. *Omega* 29(6):561–576

Cheng Y, Zhan W, Wu X, Zhang Y (2015) Automatic determination about precision parameter value based on inclusion degree with variable precision rough set model. *Inform Sci* 290:72–85

- Düntsch I, Gediga G (2015) PRE and variable precision models in rough set data analysis. In: Peters J et al (eds) Transactions on rough sets XIX. Springer, Berlin, pp 17–37
- Gomolinska A (2008) On certain rough inclusion functions. In: LNCS transactions on rough sets IX, LNCS, vol 5390, pp 35–55
- Inuiguchi M (2005) Several approaches to attribute reduction in variable precision rough set model. In: Torra V, Narukawa Y, Miyamoto S (eds) Modeling decisions for artificial intelligence. Springer, Berlin, pp 215–226
- Lin TY (1998) Granular computing on binary relations II: rough set representation and belief function. In: Skowron A, Polkowski L (eds) Rough sets and knowledge discovery. Physica-Verlag, Heidelberg, pp 122–140
- Pawlak Z (1991) Rough sets: theoretical aspects of reasoning about data (Chapters 1 and 2). Kluwer Academic Publishers, Dordrecht
- Pawlak Z (1987) Decision tables—a rough set approach. Bull EATCS 33:85–95
- Pawlak Z (1982) Rough sets. Int J Comput Inform Sci 11:341–356
- Pawlak Z, Wong SKN, Ziarko W (1988) Rough sets: probabilistic versus deterministic approach. Int J Man Mach Stud 29:81–95
- Pedrycz W, Chen SM (2014) Information granularity, big data, and computational intelligence. Springer, Heidelberg
- Pedrycz W, Chen SM (2015) Granular computing and decision-making: interactive and iterative approaches. Springer, Heidelberg
- Polkowski L, Skowron A (1996) Rough mereology: a new paradigm for approximate reasoning. Int J Approx Reason 15(4):333–365
- Shafer G (1976) Mathematical theory of evidence. Princeton University Press, Princeton
- Skowron A, Grzymala-Busse J (1994) From rough set theory to evidence theory. In: Yager RR, Fedrizzi M, Kacprzyk J (eds) Advances in the Dempster–Shafer theory of evidence. Wiley, New York, pp 193–236
- Skowron A, Jankowski A, Dutta S (2016) Interactive granular computing. Granul Comput 1(2):95–113
- Syau YR, Lin EB (2015) Evidence theory in incomplete information tables. Int J Mach Learn Comput 5:336–344
- Syau YR, Lin EB (2015) Discernibility thresholds and approximate dependency in analysis of decision tables. J Comput 10(6):412–417
- Wang JY, Zhou J (2009) Research of reduct features in the variable precision rough set model. Neurocomputing 72:2643–2648
- Yao Y, Greco S, Slowinski R (2015) Probabilistic rough sets, Springer handbook of computational intelligence. pp 387–411
- Zhang WX, Leung Y (1996) Theory of including degrees and its applications to uncertainty inferences. In: Proceedings of the 1996 Asian fuzzy systems symposium, soft computing in intelligent systems and information processing, pp 496–501
- Zhou J, Miao D (2011) β -Interval attribute reduction in variable precision rough set model. Soft Comput 15(8):1643–1656
- Ziarko W (1993) Variable precision rough set model. J Comput Syst Sci 46:39–59