Journal of Uncertain Systems
Vol.3, No.3, pp.232-240, 2009
Online at: www.jus.org.uk

Rough Set and Data Analysis in Decision Tables

J.H. Nasiri, M. Mashinchi*
[1]Islamic Azad University, Branch of Bonab, Bonab, Iran
[2]Department of Statistics, Faculty of Mathematics and Computer Sciences
Shahid Bahonar University of Kerman University, Kerman, Iran

Received 21 May 2008; Revised 23 August 2008

Abstract

The starting point of rough set theory is an information system. The information system contains data about objects of interest characterized in terms of some attributes. If we distinguish condition and decision attributes in the information system, then such a system is called a decision table. To every subset of attributes we associate a set of formulas. The decision table contains a set of decision rules. For every decision rule we associate a conditional probability which is called certainty factor. In this paper, we define the so called very positive, positive, boundary and negative regions for a decision based on the certainty factor. These regions divide the set of formulas on condition attributes according to certainty factor of decision rules. Then, some interesting relations among these divided sets of formulas are proved which are useful in decision making. Finally, a numerical example is given to show the potential application of our theoretical findings regarding selection of candidates to a school.

Keywords: rough sets, information system, decision table, certainty factor, coverage factor

1 Introduction

The concept of rough set was originally proposed by Pawlak [1, 9, 10, 11]. Its philosophy is based on the assumption that to every object of the universal set, we associate some information. Objects characterized by some information are indiscernible in view of the available information about them. The indiscernibility relation generated in this way is the mathematical basis of rough set theory [2, 3, 4].

Rough set theory approach has fundamental importance in the areas of machine learning, knowledge acquisition, decision analysis, and knowledge discovery from a database. It has been successfully applied in many real-life problems in medical [19, 22], pharmacology [5, 6], engineering, banking [20, 21] and others. The rough set approach to data analysis has many important advantages, to name a few are [6, 7, 10, 18]

- providing efficient algorithms for finding hidden patterns in data;
- finding minimal sets of data (reduction);
- evaluating significance of data;
- generating sets of decision rules from data;
- offering straight forward interpretation of obtained results;
- being easy to understand.

A rough set is basically an approximation representation of a given set in terms of two subsets derived from a crisp partition defined on the universal set. The two subsets are called a lower and an upper approximation [13, 17].

Information system is a useful concept for classification of data. It is a 4-tuple consisting of a finite universal set, a finite set of attributes, a finite set as a domain of attributes and finally an information function [2, 6, 7]. In each information system the pair of an attribute and its value is called a formula [12]. A problem related to many practical applications of information systems is whether the whole set of attributes is always necessary to define a given partition of a universe? This problem leads to define the so-called a reduce as a subset of attributes that are useful in partitioning universal set with indiscernibility relation [8, 18, 23].

Decision table is a special case of an information system. In this case, we distinguish two disjoint classes of attributes called condition (C) and decision (D) attributes. The expression *if* Φ *then* ψ is called a decision rule, where

* Corresponding author. Email: mashinchi@mail.uk.ac.ir (M. Mashinchi).

Φ and ψ belong to *C* and *D*, respectively [12, 13, 14, 16]. Then, certainty factor and coverage factor are defined for every decision rule. The two factors show correctness and consistency of a decision rule [3, 4, 12], based on which the classification of decision rules is done.

Throughout, we quote clearly references for the material presented in the text if available, otherwise they are all new. This paper is organized as follows. In Section 2, we recall basic concepts on rough sets and information systems. We prove, in Theorems 1 and 2, that lower and upper approximations are the best such bounds for a given set. In Section 3, we review the concept of decision table. We recall definitions of certainty factor, coverage factor and classification of decision rules based on the two factors. Moreover, we define very positive, positive, boundary and negative regions for a decision based on the certainty factor. Then, we present the analytical properties of these regions in Theorems 3-8. In Section 4, an example from [13] is resolved by results of this paper to show the potential application of our theoretical findings regarding selection of candidates to a school. Finally, a conclusion is given in Section 5.

2 Information System and Approximation of Sets

In this section we review some notions of rough sets which we will use throughout this paper. For details see the references.

Definition 1 ([14]). Any 4-tuple $S=<U,A_t,V,\rho>$ is called an *information system*, where U is a finite set of objects, A_t is a finite set of attributes, $V=\bigcup_{a\in A_t} v_a$, where v_a is a domain of the attributes, $\rho: U\times A_t \to V$ is a function called an information function by $\rho(x,a)=a(x)\in v_a$ for every $a\in A_t$ and $x\in U$. An information system is denoted by $S=(U,A_t)$.

Definition 2 ([14]). Let $S=(U,A_t)$ be an information system. Every $B\subseteq A_t$ generates a binary relation on ***U***, which is called an *indiscernibility relation*, and defined by $IND(B)=\{(x,y)\in U\times U : a(x)=a(y), \forall a\in B\}$. It is denoted by $IND(B)$ or simply by B.

Definition 3 ([14]). Let $S=(U,A_t)$ be an information system and $B\subseteq A_t$. The pair $A=(U,IND(B))$ is called an *approximation space*.

Definition 4 ([14]). Let $A=(U,IND(B))$ be an information system and $B\subseteq A_t$. Then $[x]_B=\{y\in U: (x,y) \in IND(B)\}$ is the equivalent class of $x\in U$ in $IND(B)$. It is called atom set or elementary set. The set of atoms will be denoted by *U/IND(B)*.

Definition 5 ([12]). Let $A=(U,IND(B))$ be an approximation space. Any union of elementary sets in A is called a *composed set* in A. The family of all composed sets in A is denoted by $Com(A)$.

Definition 6 ([9]). Let $A=(U,IND(B))$ be an approximation space and $X\subseteq U$. The *upper approximation* of X in A is defined as $\overline{A}(X)=\{x\in U: [x]_B\cap X\neq\phi\}$.

Definition 7 ([9]). Let $A=(U,IND(B))$ be an approximation space and $X\subseteq U$. The *lower approximation* of X in A is defined as $\underline{A}(X)=\{x\in U: [x]_B\subseteq X\}$.

We can give the following theorems.

Theorem 1. Let $A=(U,IND(B))$ be an approximation space and $X\subseteq U$. Then $\overline{A}(X)$ is the least upper bound of all composed sets containing *X*.

Proof: The proof follows from the following three steps.

(1) For every $x\in\overline{A}(X)$, we have $x\in[x]_B$, so $\overline{A}(X)\subseteq\bigcup_{x\in\overline{A}(X)}[x]_B$. Now, let $x\in\overline{A}(X)$. Then for every $y\in[x]_B$, we have $y\in\overline{A}(X)$. It follows that $\bigcup_{x\in\overline{A}(X)}[x]_B\subseteq\overline{A}(X)$. Therefore, $\overline{A}(X)\in Com(A)$.

(2) For all $x\in X$, we have $x\in[x]_B$, so $[x]_B\cap X\neq\phi$. By Definition 6, we have $x\in\overline{A}(X)$. Therefore, $X\subseteq\overline{A}(X)$.

(3) Let $T\in Com(A)$ be such that $X\subseteq T$. Then for all $x\in\overline{A}(X)$, we have $[x]_B\cap T\neq\phi$. So, $[x]_R\subseteq T$. Hence, $x\in T$. Therefore, for every $T\in Com(A)$ such that $X\subseteq T$, we have $\overline{A}(X)\subseteq T$.

Theorem 2. Let $A=(U,IND(B))$ be an approximation space and $X\subseteq U$. Then $\underline{A}(X)$ is the greatest lower bound of all composed sets contained in *X*.

Proof: The proof follows from the following three steps.

(1) For every $x \in \underline{A}(X)$, we have $x \in [x]_B$, so $\underline{A}(X) \subseteq \bigcup_{x\in \underline{A}(X)} [x]_B$. Let $x \in \underline{A}(X)$. Then for every $y \in [x]_R$, we have $y \in \underline{A}(X)$. It follows that $\bigcup_{x\in \underline{A}(x)} [x]_B \subseteq \underline{A}(X)$. Therefore, $\underline{A}(X) \in Com(A)$.

(2) For every $x \in \underline{A}(X)$, we have $x \in [x]_B$. Therefore, $[x]_B \subseteq X$. By Definition 7, we have $x \in X$. Therefore, $\underline{A}(X) \subseteq X$.

(3) Let $T \in com(A)$ such that $T \subseteq X$. Then for every $x \in T$, we have $[x]_B \subseteq T$, So $[x]_B \subseteq X$. By Definition 7 we have $x \in \underline{A}(X)$. It means that $T \subseteq \underline{A}(X)$. Therefore, for every $T \in Com(A)$ such that $T \subseteq X$, we have $T \subseteq \underline{A}(X)$.

Due to Theorems 1 and 2, we call $\overline{A}(X)$ and $\underline{A}(X)$ the best upper and the best lower approximation of X in A, respectively.

Definition 8 ([9]). Let $A = (U, IND(B))$ be an approximation space. The boundary set of $X \subseteq U$ in A is defined as

$$Bnd_A(X) = \overline{A}(X) - \underline{A}(X).$$

Definition 9 ([1]). Let $A = (U, IND(B))$ be an approximation space. A pair $(L, U) \in P(U) \times P(U)$ is called a rough *set* in A, where $L = \underline{A}(X)$, $U = \overline{A}(X)$ for some $X \subseteq U$ and $P(U)$ is the power set of U.

Due to the granularity of knowledge, rough sets cannot be characterized by using available knowledge. Therefore, for every rough set ***X***, we associate two *crisp* sets, called *lower* and *upper approximation*. Intuitively, the lower approximation of ***X*** consists of all elements that *surely* belong to ***X***, the upper approximation of ***X*** consists of all elements that *possibly* belong to ***X***, and the *boundary region* of ***X*** consists of all elements that cannot be classified uniquely to the set or its complement, by employing the available knowledge [15]. A pictorial presentation for Definition 9 is given in Figure 1.

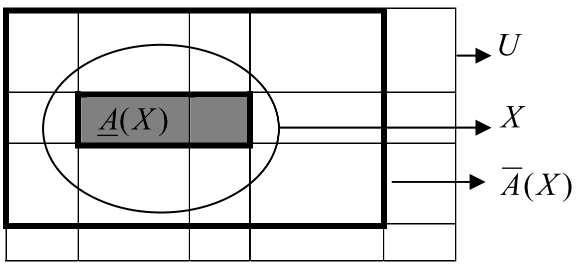

Figure 1: A rough set

Definition 10 ([23]). Let $S = (U, A_t)$ be an information system and $B \subseteq A_t$. The least B' that $B' \subseteq B$ such that $IND(B) = IND(B')$ is called a *reduce* in B. The set of all reduces in B is denoted by $RED(B)$**.**

Remark 1. Let $S = (U, A_t)$ and $B \subseteq A_t$. Then B may have more than one reduce. In other words, reduce of a set is not unique.

Definition 11 ([23]). Let $S = (U, A_t)$ be an information system and $B \subseteq A_t$. Then the intersection of all reduces of B is called the *core* of B, i.e., $Core(B) = \bigcap RED(B)$.

Remark 2. The core is a collection of the most significant attributes in the system.

3 Decision Table

In this section we are going to review the concepts of decision table and decision rules. Also, we introduce some definitions and theorems related to decision table. These points come in handy in decision making and improve the quality of decisions.

Definition12 ([12]). Let $S = (U, A_t)$ be an information system. If there are $C, D \subseteq A_t$ such that $C \cap D = \phi$ and $C \cup D = A_t$. Then S is called a *decision table,* which we denote by $S = (U, C, D)$. We call them C, D condition and decision attributes, respectively.

Definition 13 [12]. Let $S = (U, A_t)$ be an information system and $B \subseteq A_t$. Then

1. The pair *(a,v)* is called a *formula* on B, where $a \in B$, $v \in v_a$ and $\|(a, v)\|_s$ is defined as

$$\|(a, v)\|_s = \{x \in U : a(x) = v\}.$$

2. Let $\Phi=(a,v)$ and $\Psi=(a',v')$ be two formulas, where $a,a'\in B$ and $v\in v_a, v'\in v_{a'}$. Then $\Phi\vee\Psi$, $\Phi\wedge\Psi$ and $\sim\Phi$ are formulas on B and

$$\|\Phi\vee\Psi\|_S=\{x\in U: a(x)=v \text{ or } a'(x)=v'\},$$

$$\|\Phi\wedge\Psi\|_S=\{x\in U: a(X)=v \text{ and } a'(X)=v'\},$$

$$\|\sim\Phi\|_S=\{x\in U: a(x)\neq v\}.$$

3. Every formula on B is produced by 1 and 2 above.

The family of all formulas on B is denoted by $For(B)$.

Now we can prove the following theorem.

Theorem 3. Let $S=(U,A_t)$ be an information system and $B\subseteq A_t$. Then for any $\Phi,\Psi\in For(B)$, we have

1. $\|\Phi\vee\Psi\|_S=\|\Phi\|_s\cup\|\Psi\|_s$.
2. $\|\Phi\wedge\Psi\|_S=\|\Phi\|_S\cap\|\Psi\|_S$.
3. $\|\sim\Phi\|_S=U-\|\Phi\|_S$.

Proof: By Definition 13, we have

(1) $\|\Phi\vee\Psi\|_S=\{x\in U: a(x)=v \text{ or } a'(x)=v'\}=\{x\in U: a(x)=v\}\cup\{x\in U: a'(x)=v'\}=\|\Phi\|_s\cup\|\Psi\|_s$.

(2) $\|\Phi\wedge\Psi\|_S=\{x\in U: a(X)=v \text{ and } a'(X)=v'\}=\{x\in U: a(X)=v\}\cap\{x\in U: a'(x)=v'\}=\|\Phi\|_S\cap\|\Psi\|_S$.

(3) $\|\sim\Phi\|_S=\{x\in U: a(x)\neq v\}=\{x\in U: x\notin\|\Phi\|_S\}=U-\|\Phi\|_S$.

The proof of theorem is complete.

Definition 14 ([12]). Let $S=(U,C,D)$ be a decision table, $\Phi\in For(C)$ and $\Psi\in For(D)$. The expression *if* Φ *then* Ψ is called a *decision rule* and is denoted by $\Phi\to\Psi$.

Remark 3 ([2]). Let $S=(U,C,D)$ be a decision table. We define a probability distribution P_U on U, where

$$P_U(x)=\frac{1}{card(U)} \quad \text{for } x\in U.$$

Obviously, we have $P_U(X)=card(X)/card(U)$ for all $X\subseteq U$.

Definition 15 ([12]). Let $S=(U,C,D)$ be a decision table and $\Phi\to\Psi$ a decision rule in S. The *certainty factor* of this rule is defined as

$$Cer_S(\Phi,\Psi)=\frac{card(\|\Phi\wedge\Psi\|_S)}{card(\|\Phi\|_S)}.$$

It is obvious that $0\le Cer_S(\Phi,\Psi)\le 1$ for every $\Phi\to\Psi$. This coefficient is widely used in data mining and is called *confidence coefficient* as well [10].

Definition 16 ([12]). Let $S=(U,C,D)$ be a decision table and $\Phi\to\Psi$ a decision rule in S. *The coverage factor* of this rule is defined as

$$Cov_S(\Phi,\Psi)=\frac{card(\|\Phi\wedge\Psi\|_S)}{card(\|\Psi\|_S)}.$$

It is obvious that $0\le Cov_S(\Phi,\Psi)\le 1$ for every $\Phi\to\Psi$.

The certainty factors in S can be interpreted as the frequency of objects having Ψ in the set of objects with the property Φ and the coverage factor as the frequency of objects with the property Φ in the set of objects with the property Ψ.

Definition 17 ([12]). Let $S=(U,C,D)$ be a decision table and $\Phi\to\Psi$ a decision rule in S.

1. If $Cer_S(\Phi,\Psi)=1$, then $\Phi\to\Psi$ is called a *certain rule.*
2. If $0<Cer_S(\Phi,\Psi)<1$, then $\Phi\to\Psi$ is called an *uncertain rule.*

Decision rules, which are in fact logical implications, constitute a logical counterpart of approximations, whereas the uncertain rules correspond to the boundary region.

If we want to prove properties of the data, the topological language of approximations is the right tool. However, in order to describe the patterns in the data for practical cases, the logical language of implications is the proper one. We can prove some of these logical consequences in the following theorems.

Theorem 4. Let $S=(U,C,D)$ be a decision table and $\Phi\to\Psi$ a decision rule in S. Then $Cer_S(\Phi,\Psi)=Cov_S(\Phi,\Psi)$ if and only if $card(\|\Phi\|_S)=card(\|\Psi\|_S)$.

Proof: (1) Suppose that $Cer_S(\Phi,\Psi)=Cov_S(\Phi,\Psi)$, by Definitions 15 and 16, we have

$$\frac{card(\|\Phi\wedge\Psi\|_s)}{card(\|\Phi\|_s)}=\frac{card(\|\Phi\wedge\Psi\|_s)}{card(\|\Psi\|_s)}.$$

Since the numerators of fractions are equal, the denominators must be equal. This means that $card(\|\Phi\|_s)=card(\|\Psi\|_s)$.

(2) Suppose that $card(\|\Phi\|_s)=card(\|\Psi\|_s)$ and moreover assume that $\|\Phi\|_s$ and $\|\Psi\|_s$ are not empty. If they are empty then the rule of $\Phi\to\Psi$ does not exist. Therefore,

$$\frac{card(\|\Phi\wedge\Psi\|_s)}{card(\|\Phi\|_s)}=\frac{card(\|\Phi\wedge\Psi\|_s)}{card(\|\Psi\|_s)}.$$

Hence by Definitions 15 and 16, we have $Cer_S(\Phi,\Psi)=Cov_S(\Phi,\Psi)$.

Therefore the proof follows by (1) and (2).

Theorem 5. Let $S=(U,C,D)$ be a decision table and $\Phi\to\Psi$ a decision rule such that $\|\Phi\|_S\subseteq\|\Psi\|_S$, then $Cov_S(\Phi,\Psi)\le 1$.

Proof: By assumption of theorem, we have $\|\Phi\|_S\subseteq\|\Psi\|_S$. Therefore, $\|\Phi\wedge\Psi\|_s=\|\Phi\|_s$. It follows that $card(\|\Phi\wedge\Psi\|_s)=card(\|\Phi\|_s)$. This means $card(\|\Phi\wedge\Psi\|_s)/card(\|\Psi\|_s)\le 1$. By Definition 16, we have $Cov_S(\Phi,\Psi)\le 1$.

Theorem 6. Let $S=(U,C,D)$ be a decision table, $\Phi_1,\Phi_2\in For(C)$ and $\Psi\in For(D)$. Then

1. $Cov_S(\Phi_1\wedge\Phi_2,\Psi)\le Cov_S(\Phi_i,\Psi)$ for $i=1,2$.
2. $Cov_S(\Phi_1\vee\Phi_2,\Psi)\ge Cov_S(\Phi_i,\Psi)$ for $i=1,2$.

Proof: (1) We know that $\|\Phi_1\wedge\Phi_2\|_s\subseteq\|\Phi_i\|_s$ for $i=1,2$. By Theorem 3, we have

$$\|(\Phi_1\wedge\Phi_2)\wedge\Psi\|_s=\|\Phi_1\wedge\Phi_2\|_s\cap\|\Psi\|_s\subseteq\|\Phi_i\|_s\cap\|\Psi\|_s=\|\Phi_i\wedge\Psi\|_s.$$

By Definition 16 and the above relations we have

$$Cov_S(\Phi_1\wedge\Phi_2,\Psi)\le Cov_S(\Phi_i,\Psi)\text{ for } i=1,2.$$

(2) We know that $\|\Phi\|_i\subseteq\|\Phi_1\vee\Phi_2\|_s$ for $i=1,2$. By Theorem 3, we have

$$\|(\Phi_1\vee\Phi_2)\wedge\Psi\|_s=\|\Phi_1\vee\Phi_2\|_s\cap\|\Psi\|_s\supseteq\|\Phi_i\|_s\cap\|\Psi\|_s=\|\Phi_i\wedge\Psi\|_s.$$

By Definition 16 and the above relations, we have

$$Cov_S(\Phi_1\vee\Phi_2,\Psi)\ge Cov_S(\Phi_i,\Psi),\ i=1,2.$$

In order to reveal more logical consequences, we give the following definition with some interpretations.

Definition 18. Let $S=(U,C,D)$ be a decision table and $\Psi\in For(D)$. Then we define $Vpos_S(\Psi)=\{\Phi\in For(C):Cer_S(\Phi,\Psi)=1\}$, $Pos_S(\Psi)=\{\Phi\in For(C):Cer_S(\Phi,\Psi)>0\}$, $Bnd_S(\Psi)=\{\Phi\in For(C):0<Cer_S(\Phi,\Psi)<1\}$, and $Neg_S(\Psi)=\{\Phi\in For(C):Cer_S(\Phi,\Psi)=0\}$.

In connection with decision Ψ, they are called *Very Positive, Positive, Boundary and Negative regions* of $For(C)$, respectively. The interpretation of these sets is as follows.

- $Vpos_S(\Psi)$ is the set of all conditions that certainly conclude decision Ψ.
- $Pos_S(\Psi)$ is the set of all conditions that conclude decision Ψ, certainly or probably.
- $Bnd_S(\Psi)$ is the set of all conditions that probably conclude decision Ψ.
- $Neg_S(\Psi)$ is the set of conditions that never conclude decision Ψ.

Actually these four sets divide the set $For(C)$ into four regions according to certainty factor of $\Phi_i\to\Psi$, where $\Phi_i\in For(C)$. Therefore Definition 18 can be viewed in Figure 2.

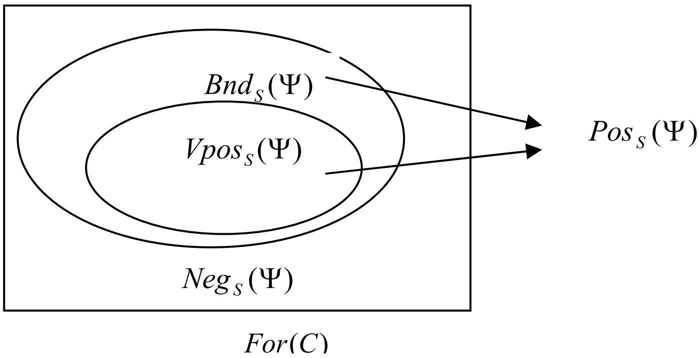

Figure 2: The four region in Definition 18

Now we can give the following theorems.

Theorem 7. Let $S=(U,C,D)$ be a decision table and $\Psi \in For(D)$. Then we have

(1) $Vpos_S(\Psi) \subseteq Pos_S(\Psi)$.

(2) $Vpos_S(\Psi) \cup Bnd_S(\Psi) = Pos_S(\Psi)$.

(3) $Vpos_S(\Psi) \cap Bnd_S(\Psi) \cap Neg_S(\Psi) = \phi$.

(4) $Vpos_S(\Psi) \cup Bnd_S(\Psi) \cup Neg_S(\Psi) = For(C)$.

Proof: (1) Let $\Phi \in for(C)$ and $\Phi \in Vpos_s(\Psi)$. By Definition 18, we have $Cer_s(\Phi,\Psi)=1$, so $Cer_s(\Phi,\Psi)>0$. Therefore, $\Phi \in Pos_s(\Psi)$. Hence the proof is complete.

(2) Let $\Phi \in for(C)$ and $\Phi \in Vpos_s(\Psi) \cup Bun_s(\Psi)$. On the other words, $\Phi \in Vpos_s(\Psi)$ or $\Phi \in Bun_s(\Psi)$. By Definition 18, we have $Cer_s(\Phi,\Psi)>0$. Hence, $\Phi \in Pos_s(\Psi)$.

Now we suppose that $\Phi \in Pos_s(\Psi)$, again from Definition 18 we can conclude that $Cer_s(\Phi,\Psi)>0$. It means that $\Phi \in Vpos_s(\Psi)$ or $\Phi \in Bnd_s(\Psi)$. Hence we have done.

(3) Let $\Phi \in Vpos_S(\Psi) \cap Bnd_S(\Psi) \cap Neg_S(\Psi)$. Then by Definition 18, we have $Cer_s(\Phi,\Psi)=1$ and $Cer_s(\Phi,\Psi)=0$. Two parts of them conclude that, there is no such Φ. Therefore, $Vpos_S(\Psi) \cap Bnd_S(\Psi) \cap Neg_S(\Psi)$ is empty, the assertion (3) is valid.

(4) From Definition 18, it is obvious that $Vpos_S(\Psi) \cup Bnd_S(\Psi) \cup Neg_S(\Psi) \subseteq For(C)$.

On the other hand, let $\Phi \in For(C)$. In this case, by Definition 15, we have $0 \le Cer_s(\Phi,\Psi) \le 1$ and $\Phi \in Vpos_s(\Psi) \cup Pos_s(\Psi) \cup Neg_S(\Psi)$. This completes the proof.

Theorem 8. Let $S=(U,C,D)$ be a decision table and $\Psi \in For(D)$. Then we have

(1) $\underline{C}(\|\Psi\|_S) = \bigcup_{\Phi \in Vpos_S(\Psi)} \|\Phi\|_S$.

(2) $\overline{C}(\|\Psi\|_S) = \bigcup_{\Phi \in Pos_S(\Psi)} (\|\Phi\|_S)$.

(3) $Bnd_S(\|\Psi\|_S) = \bigcup_{\Phi \in Bnd_S(\Psi)} (\|\Phi\|_S)$.

Proof. First note that U is partitioned by condition and decision attributes as the following partition classes respectively

$$U/C = \{\|\Phi\|_s : \Phi \in For(C)\}, U/D = \{\|\Psi\|_s : \Psi \in For(D)\}.$$

(1) By Definitions 7, 15 and 18, we have

$$\underline{C}(\|\Psi\|_s) = \bigcup_{\|\Phi\|_s \subseteq \|\Psi\|_s} \|\Phi\|_s = \bigcup_{cer_s(\Phi,\Psi)=1} \|\Phi\|_s = \bigcup_{\Phi \in Vpos_S(\Psi)} (\|\Phi\|_S).$$

(2) By Definitions 6, 15 and 18, we have

$$\overline{C}(\|\Psi\|_s) = \bigcup_{\|\Phi\|_s \cap \|\Psi\|_s \neq \phi} \|\Phi\|_s = \bigcup_{0 < cer_s(\Phi,\Psi) \le 1} \|\Phi\|_s = \bigcup_{\Phi \in Pos_S(\Psi)} \|\Phi\|_S.$$

(3) By Definitions 6, 8 and 18, we have

$$Bnd_S(\|\Psi\|_S) = \overline{C}(\|\Phi\|_s) - \underline{C}(\|\Psi\|_s) = \bigcup_{0<cer_S(\Phi,\Psi)\le 1} \|\Phi\|_S - \bigcup_{cer_S(\Phi,\Psi)=1} \|\Phi\|_S = \bigcup_{0<cer_s(\Phi,\Psi)<1} \|\Phi\|_s = \bigcup_{\Phi\in Bnd_s(\Psi)} (\|\Phi\|_S).$$

Now we consider the following example from [13] and resolve it by our theoretical results presented in this paper.

4 Example

To illustrate the rough analysis we consider a simple case of selection of candidates to a school. The candidates to the school have submitted their application packages with secondary school certification, curriculum vitae and opinion from previous school, for consideration by an admission committee based on these documents, the candidates were described using seven criteria together with corresponding scales ordered from the best to the worst value, given below

c_1 -Score in mathematics: $v_{c_1} = \{5,4,3\}$, c_2 -Score in physics: $v_{c_2} = \{5,4,3\}$,

c_3 -Score in English: $v_{c_3} = \{5,4,3\}$, c_4 - Mean in other subjects: $v_{c_4} = \{5,4,3\}$,

c_5 -Type of secondary school: $v_{c_5} = \{1,2,3\}$, c_6 -Motivation: $v_{c_6} = \{1,2,3\}$,

c_7 -Opinion from previous school: $v_{c_7} = \{1,2,3\}$, d -Decision of committee: $v_d = \{A,R\}$.

Then the set of condition attributes is $C = \{c_1,c_2,c_3,c_4,c_5,c_6,c_7\}$. Fifteen candidates with rather different application packages have been sorted by the committee after due consideration. They create the set of examples. The set of decision attributes is $D = \{A,R\}$, where A stands for an admission and R for a rejection. The information is represented in Table 1.

Table 1: Decision table composed of example

Condition / *Candidate*	c_1	c_2	c_3	c_4	c_5	c_6	c_7	*Decision d*
X_1	4	4	4	4	2	2	1	A
X_2	3	3	4	3	2	1	1	R
X_3	3	4	3	3	1	2	2	R
X_4	5	3	5	4	2	1	2	A
X_5	4	4	5	4	2	2	2	A
X_6	3	4	3	3	2	1	3	R
X_7	4	4	5	4	2	2	2	A
X_8	4	4	4	4	2	2	2	A
X_9	4	4	4	4	2	2	2	R
X_{10}	5	3	5	4	2	1	2	A
X_{11}	5	4	4	4	1	1	2	A
X_{12}	5	3	4	4	2	2	2	A
X_{13}	4	3	3	3	3	2	2	R
X_{14}	3	3	4	3	2	3	3	R
X_{15}	4	5	5	4	2	1	1	A

We have $Y_A = \{X_1, X_4, X_5, X_7, X_8, X_{10}, X_{11}, X_{12}, X_{15}\}$ and $Y_R = \{X_2, X_3, X_6, X_9, X_{13}, X_{14}\}$.

The lower approximation and upper approximation of Y_A and Y_R are as follows

$$\underline{C}Y_A = \{X_1, X_4, X_5, X_7, X_{10}, X_{11}, X_{12}, X_{15}\},\ \overline{C}Y_A = \{X_1, X_4, X_5, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{15}\}.$$

Moreover $Bnd_C(Y_A)=\{X_8,X_9\}$. Also, we have $\underline{C}Y_R=\{X_2,X_3,X_6,X_{13},X_{14}\}$, $\overline{C}Y_R=\{X_2,X_3,X_6,X_8,X_9,X_{13},X_{14}\}$, and $Bnd_C(y_R)=\{X_8,X_9\}$.

The next step of the rough set analysis of the decision table is construction of minimal subset of independent criteria ensuring the same quality of sorting as whole set C, i. e., the reduces of C. In our case, there are three such reduces: $RED_Y^1(C)=\{c_2,c_3,c_6,c_7\}$, $RED_Y^2(C)=\{c_1,c_3,c_7\}$, and $RED_Y^3(C)=\{c_2,c_3,c_5,c_7\}$

It can be said that the committee took the fifteen sorting decisions taking into account the criteria from one of reduces and discarded all the remaining criteria. Let us notice that criterion c_4 has no influence at all on the decision because it is not represented in any reduce.

We suppose that the committee has chosen reduce $RED_Y^2(C)$ composed of c_1, c_3, c_7, i.e., score in mathematics and English and opinion from previous school. Now the decision table can be reduced to criteria represented in $RED_Y^2(C)$. The decision rules generated from the reduced decision table have the following form:

r_1) $(c_1,5)\mapsto(d,A)$.

r_2) $(c_3,5)\mapsto(d,A)$.

r_3) $(c_1,4)\wedge(c_7,1)\mapsto(d,A)$.

r_4) $(c_1,4)\wedge(c_3,4)\wedge(c_7,2)\mapsto(d,A)$.

r_5) $(c_1,4)\wedge(c_3,4)\wedge(c_7,2)\mapsto(d,R)$.

r_6) $(c_1,3)\mapsto(d,R)$, and r_7) $(c_3,3)\mapsto(d,R)$.

Five rules are certain and two rules are uncertain. We have certainty factor and coverage factor of rules in Table 2.

Table 2: Certainty and coverage factors

Rule	Certainty	Coverage
1	1	0.44
2	1	0.55
3	1	0.33
4	0.5	0.11
5	0.5	0.11
6	1	0.66
7	1	0.5

From the certainty factors of decision rules, we can conclude:

1. If score in mathematics is 5, then the candidate is admitted certainly.
2. If score in physics is 5 then the candidate is admitted certainly.
3. If score in mathematics is 4 and opinion from previous school be 1, then the candidate is admitted certainly.
4. If score in mathematics is 4 and score in English is 4 and opinion from previous school is 2, then the possibility that the candidate admitted is equal to 0.5.
5. If score in mathematics is 4 and score in English is 4 and opinion from previous school is 2, then the possibility that the candidate rejected is equal to 0.5.
6. If score in mathematics is 3, then the candidate is rejected certainly.
7. If score in English is 3, then the candidate is rejected certainly.

5 Conclusions

In this paper approximation space and basic concepts of rough set theory have been reviewed. Decision table and decision algorithms, with many applications recently, have been discussed on decision table. Here, we recalled certainty factor and coverage factor. These two coefficients express the accuracy of decision rules. We defined several regions that have different values of certainty factors and several interesting theoretical results were proved. Using these regions is helpful in decision making. It seems that, using weaker relations instead of indiscernibility relation make a better decision table. To show the potential application of our theoretical findings regarding selection of candidates to a school, a known example from the literature was resolved by results of this paper

Acknowledgments

The second author is partially supported by Fuzzy Systems and Applications Center of Excellence, Shahid Bahonar University of Kerman, Iran. The authors also would like to thank the anonymous referees for their valuable suggestions and comments which improved the quality of this paper.

References

[1] Biswas, R., and S. Nanda, Rough groups and rough subgroups, *Bulletin of the Polish Academy of Sciences Mathematics*, vol.42, no.3, pp.251-254, 1994.

[2] Intan, R., and M. Mukaidono, Degree of similarity in fuzzy partition, *AFSS*, pp.20-26, 2002.

[3] Intan, R., and M. Mukaidono, Generalization of rough membership function based on alpha-covering of the universe, *Proceeding of AFSS*, pp.129-135, 2002.

[4] Intan, R., M. Mukaidono, and Y.Y. Yao, Generalization of rough sets with alpha-coverings of the universe induced by conditional probability relations, *Proceedings of International Workshop on Rough Sets and Granular Computing*, pp.311-315, 2001.

[5] Krysinski, J., Rough set approach to the analyses of structure activity relationship of quaternary imidazolium components, *Arzneimittel Forschungl Drug Research*, vol.40, no.2, pp. 593-597,1990.

[6] Krysinski, J., Application of the rough sets theory to the analysis of structure-activity relationship of antimicrobial pyridinium components, *Die Pharmazie*, vol.50, pp.593-597, 1995.

[7] Kryzkiewicz, M., and H. Rybinski, Finding reducts in composed information system, *Proceedings of the International Workshop on Rough Sets and Knowledge Discovery*, pp.261-273, 1993.

[8] Lashin, E.F., and T. Medhat, Topological reduction of information systems, *Chaos, Solitons & Fractals*, vol.25, no.2, pp.277-286, 2005.

[9] Pawlak, Z., Rough sets, *J. Comput. Information Sciences*, vol.11, pp.341-345, 1982.

[10] Pawlak, Z., Rough sets and fuzzy sets, *Fuzzy Sests and Systems*, vol.17, pp.99-102, 1985.

[11] Pawlak, Z., Hard and soft sets, *Fuzzy Sets and Knowledge-Discovery*, pp.130-135, 1994.

[12] Pawlak, Z, Rough sets and intelligent data analysis, *Information Sciences*, vol.147, pp.1-12, 2002

[13] Pawlak, Z., and R. Slowinski, Decision analysis using rough sets, *International Transactions in Operational Research*, vol.1, pp.107-114, 1994.

[14] Pawlak, Z., and R. Slowinski, Rough set approach to multi–attribute decision analysis, *European Journal of Operational Research*, vol.72, pp.443-449, 1994.

[15] Pawlak, Z., and A. Skorwon, Rudiments of rough sets, *Information Sciences*, vol.177, no.1, pp.3-27, 2007.

[16] Qian, Y., J. Liang, *et al.*, Measures for evaluating the decision performance of a decision table in rough set theory, *Information Sciences*, vol.178, no.1, pp.181-202, 2008.

[17] Radzikowska, A.M., and E.E. Kerre, A comparative study of fuzzy rough sets, *Fuzzy Sets and Systems*, vol.126, pp.137-135, 2002.

[18] Skorwon, A., and C. Rauszer, The discernibility matrices and functions in information systems, Intelligence Decision Support, Handbook of Application and Advances of the Rough Set Theory, pp.311-362, 1992.

[19] Slowinski, R., Rough classification of HSV patients, In: Slowinski, R. (ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Set Theory, pp.77-93, 1992.

[20] Slowinski, R., and C. Zopounidis, Sough set sorting of firms according to bankruptcy risk, *Applying Multiple Criteria Aid for Decision to Environmental Management*, pp.339-357, 1994.

[21] Slowinski, R., and C. Zopounidis, Applications of the rough set approach to evaluation of bankruptcy risk, *International Journal of Intelligent Systems in According Finance and Management*, vol.4, pp.27-41, 1995.

[22] Tanaka, H., H. Ishibuchi, and T. Shigenaga, Fuzzy inference system based on rough sets and its applications to medical diagnostic, Intelligent Decision Support, Handbook of Applications and Advances of the Rough Set Theory, pp.111-117, 1992.

[23] Zhao, Y., Y. Yao, and F. Luo, Data analysis based on discernibility and indiscernibility, *Information Sciences*, vol.177, no.22, pp.4959-4979, 2007.