A New Measure to Evaluate the Consistency of a Set of Decision Rules Extracted From a Decision Table

Nguyen Duc Thuan, *Member, IACSIT*, Nguyen Xuan Huy

***Abstract*— Two classical measures, approximation accuracy and consistency degree can be employed to evaluate the decision performance of a decision table. However, these two measures cannot give elaborate depictions of certainty and consistency of a decision table when their values equal to zero. Yuhua Qian et al have introduced three new measures for evaluating the decision performance of a decision–rule set extracted from a decision table, but these measures still have some limitations. In this paper, we propose a new measure and theoretically study to evaluate the consistency of a set of decision rules extracted from a decision table. Experimental analyses on two practical data sets also show our measure will be suited for evaluating the decision performance of a decision-rule than measures of Yuhua Qian et al.**

***Index Terms*— accuracy degree, consistency degree, decision table, decision evaluation, rough set theory**

I. INTRODUCTION

Rough set theory proposed by Pawlak has become a popular mathematical framework for the analysis of a vague description of an object, pattern recognition, image processing, feature selection, conflict analysis, decision support, datamining and knowledge discovery from large data set.

In recent years, how to evaluate the decision performance of a decision rule has become a very important issue in rough set theory. Many authors have proposed measures based on information entropy for this problem. Several other measures such as certainty measure and support measure are often used to evaluate a decision rule. However, all of the above measures are only defined for a single decision rule and are not suitable for measuring the decision performance of a rule set. Two measures: certainty measure and support measure, in some sense, could be regarded as measures for evaluating the decision performance of all decision-rules generated from a complete decision table (Pawlak). Nevertheless, they have some limitations. For instance, the certainty and consistency of a rule set could not be well characterized by the approximation accuracy and consistency degree when their values reach zero. As we know, when the approximation accuracy or consistency degree equal to zero, it is only implied that there is no decision rule with the certainty of one in the complete decision table. To overcome the shortcomings of the existing measures, Yuhua Qian et al have introduced three new measures for evaluating the decision performance of a decision–rule set extracted from a decision table. In their papers, they present properties of these three measures [1,2,3]. Although these measures have many good properties, the consistency measure has limitation: it is not monotony as the same classical measure. In this paper, we propose a new measure to evaluate the consistency of a set of decision rules extracted from a decision table to change the consistency measure of group authors.

The paper is organized as follows. Section 2 briefly introduces information system, decision table, partial relation decision rule, position region reduct. Section 3, presents three measures and properties. In section 4, our new consistency measure is introduced as well as its use on two data sets. At last, the paper is concluded with a summarization in section 5.

II. BASIC CONCEPTS

An information system is a pair S = (U, A), where U is a non-empty, finite set of objects and is called the universe and A is a non-empty, finite set of attributes. For each $a \in A$, $a : U \rightarrow V_a$, where V_a is the domain of *a*.

Each non-empty subset B ⊆ A determines an indiscernibility relation in the following way:

$$R_B = \{(x,y) \in UxU \mid a(x) = a(y), \forall a \in B\}$$

The relation R_B partitions U into some equivalence classes given by:

$U/R_B = \{[x]_B \mid x \in U\}$, *where* $[x]_B$ *denotes the equivalence class determined by x with respect to B, i.e.,*

$$[x]_B = \{y \in U \mid (x,y) \in R_B\}$$

A partial relation $\preceq$ on the family {U/B | B ⊆ A} is defined as follows:

$U / P \preceq U / Q$ *if and only if:* $\forall P_i \in U/P$, *there exists* $Q_j \in U/Q$ *such that* $P_i \subseteq Q_j$. *In this case, we say Q is coarser than P or P is finer than Q.*

A decision table is an information system S = (U, C ∪ D), with C ∩ D = ϕ), C is called a condition attribute set, a element of D is called a decision attribute, and D is called a decision attribute set. Suppose $U/C = \{X_1, X_2,.., X_m\}$ and U/D

Manuscript received April, 14, 2009.

Nguyen Duc Thuan is with the Department of Information Systems, Nha Trang University, Vietnam (phone: +84 58 905444448; fax: +84 58 3831147).

Nguyen Xuan Huy is with the Institute of Information Technology Vietnamese Academy of Science and Technology, Hanoi, Vietnam

= {Y_1, Y_2,.., Y_n}.

A condition class $X_i \in U/C$ is said to be consistent if $d(x) = d(y)$, $\forall x,y \in X_i$ and $\forall d \in D$; a decision class $Y_j \in U/D$ is said to be conversely consistent if $a(x)=a(y)$, $\forall x,y \in Y_j$ and $\forall a \in C$.

A decision table S = (U, C $\cup$ D) is said to be consistent if every condition classes $X_i \in U/C$ is consistent. It is easy to see that if U/C $\preceq$ U/D then S = (U, C $\cup$ D) is said to be consistent.

One can define if U/D $\preceq$ U/C, then S is said to be conversely consistent.

Let S = (U, C $\cup$ D) be a decision table, $X_i \in U/C$, $Y_j \in U/D$ and $X_i Ç Y_j \neq \emptyset$. By des(X_i) and des(Y_j), we denote the descriptions of the equivalence classes X_i and Y_j in the decision table S. A decision rule is formally defined as:

$$Z_{ij} : des(X_i) ® des(Y_j).$$

Certainty measure and support measure of decision rule Z_{ij} are defined as follows:

$$\mu(Z_{ij}) = |X_i \cap Y_j| / |X_i| \text{ and } s(Z_{ij}) = |X_i \cap Y_j| / |U|$$

It is clear that $\mu(Z_{ij})$ and $s(Z_{ij})$ of a decision rule Z_{ij} falls into the interval $[\frac{1}{|U|}, 1]$. We also denote $|Z_{ij}|$ by $|X_i Ç Y_j|$.

Pawlak proposed consistency measure of a decision table S is

$$C_C(D) = \frac{\Sigma_{i=1}^{n} \left|\underline{C}Y_i\right|}{|U|}$$

It is clear that S is a decision conversely consistent table and consistency table: $C_C(D) = 0$.

C-positive region of D is determined:

$$POS_C(D) = \bigcup_{Y \in U|IND(D)} \underline{C}(Y) = \bigcup_{i=1}^{n} \underline{C}(Y_i)$$

If B $\subseteq$ C satisfies the following conditions:

1. $POS_C(D) = POS_B(D)$
2. "a Î C-B, $POS_C(D) = POS_{C-\{a\}}(D)$

B is a positive reduct of D with respect to C.

III. Decision rule and decision performance measurement in decision tables

Yuhua Qian et al have introduced three new measures: certainty measure, consistency measure and support measure for evaluating the decision performance of a decision–rule set extracted from a decision table, they were defined as follows:

Definition 3.1 [1]: Let S = (U, C$\cup$D) be a decision table, and RULE ={Z_{ij}| Z_{ij}: des(X_i)$\rightarrow$ des(Y_j), $X_i \in U/C$, $Y_j \in U/D$}, Certainty measure α of S is defined as:

$$a(S) = \sum_{i=1}^{m} \sum_{j=1}^{n} s(Z_{ij}) m(Z_{ij}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|^2}{|U|\left|X_i\right|}$$

Consistency measure b of S is defined as:

$$b(S) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} [1 - \frac{4}{|X_i|} \sum_{j=1}^{N_i} \left|X_i \cap Y_j\right| m(Z_{ij})(1 - m(Z_{ij}))]$$

Support measure g of S is defined as:

$$g(S) = \sum_{i=1}^{m} \sum_{j=1}^{n} s^2(Z_{ij}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|^2}{|U|^2}$$

Where $m(Z_{ij})$ and $s(Z_{ij})$ are the certainty degree and support degree of the rule Z_{ij} respectively.

IV. New measures of Decision rule and decision performance measurement in decision tables

In this section, we introduce new measures of decision rule and decision performance measurement in decision tables in rough set theory. Furthermore, we prove some important properties of them. The comparisons of values of measures with the numbers of features in two data sets from UCI Repository of machine learning databases are presented.

Definition 4.1: Let S = (U, C$\cup$D) be a decision table, and RULE ={Z_{ij} | Z_{ij}: des(X_i)$\rightarrow$ des(Y_j), $X_i \in U/C$, $Y_j \in U/D$}, Certainty measure $\tau\alpha$ of S is defined as:

$$ta(S) = 1 - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{\left|X_i\right|} \frac{\left|X_i \cap Y_j^C\right|}{|U|}$$

Consistency measure tb of S is defined as:

$$tb(S) = 1 - \frac{n}{n-1} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{\left|X_i\right|} \frac{\left|X_i \cap Y_j^C\right|}{|U|}$$

Where $m(Z_{ij})$ and $s(Z_{ij})$ are the certainty degree and support degree of the rule Z_{ij} respectively.

Remark 4.1: For support measure, we use support measure of group authors (Yuhua Qian et al.). It means support measure γ of S.

Theorem 4.1: Let S = (U, C$\cup$D) be a decision table, and RULE ={Z_{ij} | Z_{ij}: des(X_i)$\rightarrow$ des(Y_j), $X_i \in U/C$, $Y_j \in U/D$}, certainty measure $\tau\alpha$ of S is certain measure α.

Proof: It is easy to see that

$$\left|X_i \cap Y_j^C\right| = \left|X_i\right| - \left|X_i \cap Y_j\right|.$$

Therefore, we have

$$ta(S) = 1 - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{\left|X_i\right|} \frac{\left|X_i \cap Y_j^C\right|}{|U|}$$

$$= 1 - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{\left|X_i\right|} (\frac{\left|X_i\right| - \left|X_i \cap Y_j\right|}{|U|})$$

$$= 1 - \sum_{i=1}^{m} \sum_{j=1}^{n} (\frac{\left|X_i \cap Y_j\right|}{|U|} - \frac{\left|X_i \cap Y_j\right|^2}{|U|\left|X_i\right|})$$

$$= 1 - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{|U|} + \sum_{i=1}^{m} \sum_{j-1}^{n} \frac{\left|X_i \cap Y_j\right|^2}{\left|X_i\right||U|}$$

Since $\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{|U|} = 1$. Therefore, we can get that:

$$1 - \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left|X_i \cap Y_j\right|}{|U|} + \sum_{i=1}^{m} \sum_{j-1}^{n} \frac{\left|X_i \cap Y_j\right|^2}{\left|X_i\right||U|} = \sum_{i=1}^{m} \sum_{j-1}^{n} \frac{\left|X_i \cap Y_j\right|^2}{\left|X_i\right||U|}$$

=α(S). This completes the proof.

Remark 4.2: As we see theorem 4.1, τα measure is α measure. So formula of τα has a common component with the other measures. It is convenient for programming when we use these measures concurrently.

We can see the important properties of τα or α measure in [1,2,3,4]. We can list as follows:

Theorem 4.2 [1]: (Extreme) Let S = (U, C∪D) be a decision table, and RULE ={Z_{ij}| Z_{ij}: des(X_i)→ des(Y_j), X_i∈ U/C, Y_j∈ U/D}

For any Z_{ij}∈ RULE, If μ(Z_{ij})=1, then the measure τα(S) achieves its maximum value 1.

If m=1 and n= |*U*|, then the measure τα(S) achieves its minimum value $\frac{1}{|U|}$.

Theorem 4.3 [1]: Let S_1=(U, C_1∪D_1), S_2=(U, C_2∪D_2) be two conversely consistent decision tables. If U/C_1=U/C_2 and U/$D_2 \preceq$ U/D_1 then τα(S_1) ≥ τα(S_2).

Theorem 4.4 [4]: Let S= (U, C∪D), S'=(U, B∪D) be two decision tables, where S is consistent, B⊆C. If B is a positive-region reduct of C, then

τα(S) = τα(S')

Theorem 4.5 [4]: Let S= (U, C∪D), S'=(U,B∪D) be two decision tables, B⊆C. If B is a positive-region reduct of C, then

$$\tau\alpha(S) \geq \tau\alpha(S')$$

Now, we introduce measure τβ and some results which have been found by us:

Theorem 4.6: (Extreme) Let S = (U, C∪D) be a decision table, and RULE ={Z_{ij}| Z_{ij}: des(X_i)→ des(Y_j), X_i∈ U/C, Y_j∈ U/D}

1. For every Z_{ij}∈ RULE, If μ(Z_{ij})=1, then the measure τβ (*S*) achieves its maximum value 1.

2. If m=1 and n= |*U*|, the measure τβ(*S*) achieves its minimum value 0.

Proof: (1) From μ(Z_{ij})=1⇔

$$\frac{|X_i \cap Y_j|}{|X_i|} = 1 \Leftrightarrow \frac{|X_i \cap Y_j^C|}{|U|} = 0$$

$$\Rightarrow \frac{|X_i \cap Y_j|}{|U|}\frac{|X_i \cap Y_j^C|}{|U|} = 0, \forall i, j$$

$$\Rightarrow \tau\beta(S) = 1 - \frac{n}{n-1}\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|X_i|}\frac{|X_i \cap Y_j^C|}{|U|} = 1$$

(2) By definition τβ(S), we can write it as:

$$\tau\beta(S) = 1 - \frac{n}{n-1}\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|X_i|}\frac{|X_i \cap Y_j^C|}{|U|}$$

$$= 1 - \frac{n}{n-1}(1 - \tau\alpha(S)) = 1 + (\tau\alpha(S) - 1)\frac{n}{n-1}$$

If m=1 and n= |*U*|, then the measure τα(*S*) achieves its minimum value $\frac{1}{|U|} = \frac{1}{n}$ (Theorem 4.2). Hence, the measure τβ(*S*) achieves its minimum value 0.

$$\tau\beta(S)_{\min} = 1 + (\frac{1}{n} - 1)(\frac{n}{n-1}) = 0$$

This completes the proof.

The monotony of the measure τβ on the conversely consistent decision tables can be found in the following theorems:

Theorem 4.7: Let S_1=(U, C_1∪D_1), S_2=(U, C_2∪D_2) be two conversely consistent decision tables. If U/C_1=U/C_2 and U/$D_2 \preceq$ U/D_1 then τβ(S_1) ≥ τβ(S_2).

Proof: From Definition 4.1 and theorem 4.3. The monotony of the measure τβ and τα are the same. We also directly prove it:

Since U/C_1 = U/C_2, S_1 and S_2 be two conversely consistent decision tables. Suppose:

U/C_1 = U/C_2 = {X_1, X_2,.., X_m}

U/D_1 = {Y_1, Y_2,.., Y_n}

If U/$D_2 \preceq$ U/D_1, then U/D_2 = $\bigcup_{j=1}^{n}\{\bigcup_{k=1}^{t_j} Y_j^k\}$, where

$$Y_j = \bigcup_{k=t_{j-1}+1}^{t_j} Y_j^k$$, let t_0 = 0, t_n=s, Z_k= Y_j^k

It follows that U/D_2= {Z_1, Z_2,.., Z_s}.

We can get that:

$$|X_i \cap Y_j||X_i \cap Y_j^C|$$

$$= \left|X_i \cap (\bigcup_{k=t_{j-1}+1}^{t_j} Y_j^k)\right|\left|X_i \cap (\bigcup_{k=t_{j-1}+1}^{t_j} Y_j^k)^C\right|$$

$$\leq \left|\bigcup_{k=t_{j-1}+1}^{t_j}(X_i \cap Y_j^k)\right|\left|\bigcup_{k=t_{j-1}+1}^{t_j}(X_i \cap (Y_j^k)^C\right|$$

$$= \sum_{k=t_{j-1}+1}^{t_j}|X_i \cap Y_j^k||X_i \cap (Y_j^k)^C|$$

So, we know that if

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|X_i|}\frac{|X_i \cap Y_j^C|}{|U|} \leq \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{s}\frac{|X_i \cap Y_j^k|}{|X_i|}\frac{|X_i \cap (Y_j^k)^C|}{|U|}$$

THEN

$$1 - \frac{n}{n-1}\sum_{i-1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|U|}\frac{|X_i \cap Y_j^C|}{|X_i|}$$

$$\geq 1 - \frac{n}{n-1}\sum_{i=1}^{m}\sum_{k=1}^{s}\frac{|X_i \cap Z_k|}{|U|}\frac{|X_i \cap Z_k^C|}{|X_i|}$$

In other words, we have τβ (S_1) ≥ τβ (S_2). This completes the proof.

This theorem states that the consistency measure τβ of a conversely consistent decision table increases with its decision classes becoming finer with all μ(*Zij*).

Lemma 4.1: Let S = (U, C∪D) be a decision table, and 2 non-empty sets X, Y $\subseteq$U. Suppose X= $\bigcup_{j=1}^{k} X_j$, $X_p \cap X_q = \emptyset$ for any p≠q, i.e. $\{X_1, X_2, .., X_k\}$ is a partition of X , then

$$\frac{|X \cap Y|}{|U|}\frac{|X \cap Y^C|}{|X|} \geq \sum_{j=1}^{k} \frac{|X_j \cap Y|}{|U|}\frac{|X_j \cap Y^C|}{|X|}$$

And, if $|X_p \cap Y||X_q \cap Y^C| = 0$, for any p≠q and p,q = 1, 2, ..,k, we have "=".

Proof: Similar to idea of Lemma 4.3 in [5], we can prove it as follows:

Since, $X_p Ç X_q$=Æ for any p¹q, we have

$$\frac{|X \cap Y|}{|U|}\frac{|X \cap Y^C|}{|X|} = \frac{\left|\bigcup_{j=1}^{k} X_i \cap Y\right|}{|U|}\frac{\left|\bigcup_{j=1}^{k} X_i \cap Y^C\right|}{|X|}$$

$$= \sum_{j=1}^{k}\frac{|Xj \cap Y|}{|U|}\sum_{j=1}^{k}\frac{|X_j \cap Y^C|}{|X|} = \sum_{j=1}^{k}\sum_{p=1}^{k}\frac{|X_j \cap Y|}{|U|}\frac{|X_p \cap Y^C|}{|X|}$$

$$\geq \sum_{j=1}^{k}\frac{|X_j \cap Y|}{|U|}\frac{|X_j \cap Y^C|}{|X|}.$$

It is clearly, if $|X_p \cap Y||X_q \cap Y^C| = 0$, for any p≠q and p,q = 1, 2, ..,k, we have "=". This completes the proof.

Theorem 4.8: Let S_1=(U, $C_1 \cup D_1$), S_2=(U, $C_2 \cup D_2$) be two conversely consistent decision tables. If U/D_1=U/D_2 and U/$C_2 \preceq$ U/C_1 then $\tau\beta(S_1) \leq \tau\beta(S_2)$.

Proof: Suppose U/D_2 = $\{Y_1, Y_2,.., Y_n\}$,U/C_1 = $\{X_1, X_2,.., X_m\}$,

If $X_i = \bigcup_{k=t_{i-1}+1}^{t_i} X_i^k \Rightarrow$ U/$C_2 = \bigcup_{i=1}^{m}\{\bigcup_{k=t_i+1}^{t_i} X_i^k\}$, let t_0=0,

From lemma 4.1, for each j=1,..,n and i=1,..,m, we have:

$$\frac{|X_i \cap Y_j|}{|U|}\frac{|X_i \cap Y_j^C|}{|X_i|} \geq \sum_{k=t_{i-1}+1}^{t}\frac{|X_i^k \cap Y_j|}{|U|}\frac{|X_i^k \cap Y_j|}{|X_i|},$$

let t_0=0. Thus,

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|U|}\frac{|X_i \cap Y_j^C|}{|X_i|}$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=t_{i=1}+1}^{t_i}\frac{|X_i^k \cap Y_j|}{|U|}\frac{|X_i^k \cap Y_j^C|}{|X_i|}$$

Then, we can get that

$$1-\frac{n}{n-1}\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i \cap Y_j|}{|U|}\frac{|X_i \cap Y_j^C|}{|X_i|}$$

$$\leq 1-\frac{n}{n-1}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=t_{i=1}+1}^{t_i}\frac{|X_i^k \cap Y_j|}{|U|}\frac{|X_i^k \cap Y_j^C|}{|X_i|}.$$

Therefore, we have $\tau\beta(S_1) \leq \tau\beta(S_2)$. This completes the proof.

Lemma 4.2: Let S_1=(U, C$\cup D_1$), S_2=(U, B$\cup D_2$) be two decision tables, if C$\subseteq$B then U/B $\preceq$ U/C then $\tau\beta(S_2) \leq \tau\beta(S_1)$.

And, $\tau\beta(S_2)=\tau\beta(S_1)$ if S_1, S_2 are consistent decision tables.

Proof: We only consider the case $\tau\beta(S_2)=\tau\beta(S_1)$ (*).

when U/B $\preceq$ U/C we have $\forall X_i \in$ U/C, X_i can write $X_i = \bigcup_{k=1}^{t_i} X_i^k$, where $X_i^k \in$U/B. From lemma 4.1 if $|X_i^k \cap Y_j||X_i^l \cap Y_j^C| = 0$, ∀i,j, k≠l we have $\tau\beta(S_2) = \tau\beta(S_1)$. So $|X_i^k \cap Y_j||X_i^l \cap Y_j^C| = 0$ if $|X_i^k \cap Y_j||X_i^l \cap Y_j^C| = 0$ (1), or $|X_i^l \cap Y_j||X_i^k \cap Y_j^C| = 0$ (2).

If we have (1), then $|X_i^k \cap Y_j| = 0$ or $|X_i^l \cap Y_j^C| = 0$, or $|X_i^k \cap Y_j| = 0, |X_i^l \cap Y_j^C| = 0$.

If $|X_i^k \cap Y_j| = 0 \Rightarrow |X_i^k \cap Y_j^C| \neq 0 \Rightarrow |X_i^l \cap Y_j| = 0$. Since X_i^k , X_i^l are the same, we can say that, $\tau\beta(S_2)=\tau\beta(S_1)$ holds, when $\exists j_0$, ∀k=1..t_i, $X_i^k \subseteq Y_j$. In other words, S, S' are consistent decision tables. In this case, we also have $m(Z_{ij}) = m(Z_{ij}^k)$, ∀j,k, where $X_i = \bigcup_{k=1}^{t_i} X_i^k$.

Theorem 4.9: Let S_1=(U, C$\cup D_1$), S_2=(U, B$\cup D_2$) be two decision tables, if S is consistent decision table and B$\subseteq$C. If B is a positive reduct of D with respect to C, then

$$tb(S_2) = tb(S_1)$$

Proof: It is straightforward from definition positive reduct and lemma 4.2

From lemma 4.2, we also have theorem:

Theorem 4.10: Let S_1=(U, C$\cup D_1$), S_2=(U, B$\cup D_2$) be two decision tables, if B$\subseteq$C and B is a positive reduct of D with respect to C, then

$$tb(S_2)^3 tb(S_1)$$

For general decision tables, to illustrate the differences between the consistency measure τβ and the consistency measures: β and $C_C(D)$, we have downloaded two data sets from UCI Repository of machine learning databases [6], which described in table 1. All condition attributes and decision attributes in two data sets are discrete.

Here, we compare the consistency measure τβ with consistency measures: β and $C_C(D)$ on 2 data sets: Tic-tac-toe, Dermatology. The comparisons of values of three measures with the numbers of features in these 2 data sets are shown in Tables 2-3 and figs 1-2.

TABLE 1 DATA SETS DESCRIPTION.

Data sets	Samples	Condition features	Decision classes
Tic-tac-toe	958	9	2
Dermatology	366	33	6

TABLE 2 $c_c(D)$, β AND τβ WITH DIFFERENT NUMBERS OF FEATURES IN THE DATA SET TIC-TAC-TOE;

Measure / Features	$c_c(D)$	b	tb
1	0.0000	0.1114	0.1114
2	0.0000	0.1322	0.1322
3	0.1253	0.2827	0.2827
4	0.1628	0.3300	0.3300
5	0.4186	0.5832	0.5832
6	0.7766	0.8000	0.8000
7	0.9436	0.9436	0.9436
8	1.0000	1.0000	1.0000
9	1.0000	1.0000	1.0000

TABLE 3 $c_c(D)$, β AND τβ WITH DIFFERENT NUMBERS OF FEATURES IN THE DATA SET DERMATOLOGY

Measure / Features	$c_c(D)$	b	tb
1	0.0000	0.3350	0.0854
2	0.0109	0.3164	0.1581
3	0.0437	0.2821	0.2960
6	0.6066	0.6826	0.7661

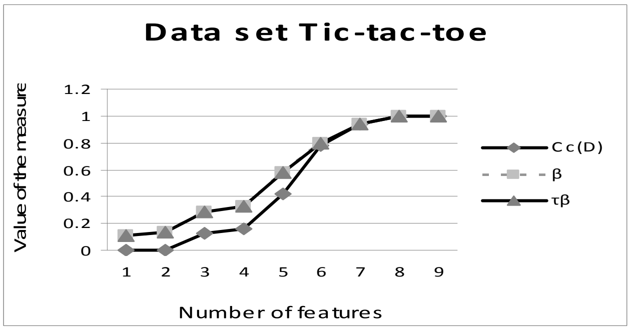

Figure 1. Variation of the consistency measures: τβ, β and consistency degree with the number of features (data set Tic-Tac-Toe)

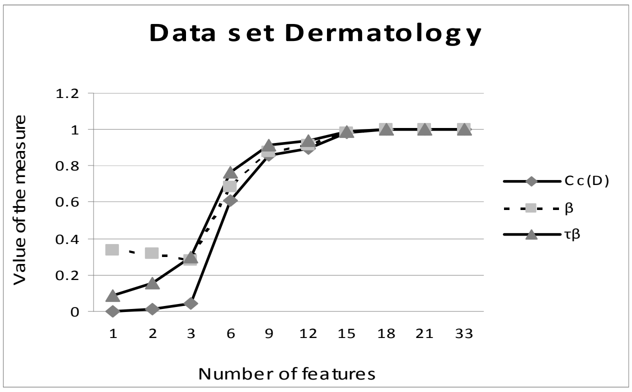

Figure 2. Variation of the consistency measures: τβ, β and consistency degree with the number of features (data set Dermatology)

It can be seen from tables 2-3 that the value of consistency measure τβ is biggest for the same number of selected features. However, τβ and $C_C(D)$ are the same monotony on all number of selected features, but β is not, it increases with its decision classes becoming finer when $\forall\mu(Z_{ij})\leq 1/2$, and decreases with its decision class becoming finer when $\forall\mu(Z_{ij})\geq 1/2$ and it decreases with its condition classes becoming finer when $\forall\mu(Z_{ij})\leq 1/2$, and increases with its decision class becoming finer when $\forall\mu(Z_{ij})\geq 1/2$ (We also can see this through theorem 6 and theorem 7 in [1]).

V. CONCLUSION

In this paper, we introduce two measures for evaluating a decision rule or decision table such as certainty measure and consistency measure. Note: certainty measure by Yuahua Qian et al is identical to the certainty measure of mine. We have proposed our measures to overcome the limitations of previous measures. Our measures have simple formulas, and they have common components. So, they are convenient for programming when we use these measures concurrently. We have proved theorems and properties of our measures. The experimental analyses on the 2 practical decision tables show that these new measures are adequate for evaluating the decision performance of a decision–rule set extracted from a decision table in rough set theory.

REFERENCES

[1] Yuhua Qian, Jiye Liang, Deyeu Li, Haiyun Zhang, Chuangyin Dang, Measures of evaluating the decision performace of a decision table in rough set theory, Information Sciences 178(2008) 181 – 202.
[2] Yuhua Qian, Jiye Liang, Chuangyin Dang, Converse approximation and rule extraction from decision tables in rough set theory, Computers and Mathematics with Applications 55(2008) 1754-1765.
[3] Yuhua Qian, Chuangyin Dang, Jiye Liang, Haiyun Zhang, Jianmin Ma, On the valuation of the decision performance of an incomplete decision table, Data & Knowledge Engineering 65(2008) 373-400.
[4] Wei Wei, Jiye Liang, Yuhua Qian, Change Mechanism of a Decision Table's Decision Performance Caused by Attribute Reductions, Granular Computing, 2008, GrC 2008, IEEE Internation Conference p 638 – 643.
[5] Nguyen Thanh Tung, Nong Thi Hoa, A new method construction of decision tree based on rough set theory, Proceedings of National Conference, Hue 12-13, 2008, 229-241.
[6] The UCI machine learning repository, <http://mlearn.ics.uci.edu/MLRepository.html>.
[7] Z. Pawlak, Rough sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, London 1991.

First Author Name: Nguyen Duc Thuan was born in Hue, Vietnam, 1962
Brief Biographical History:
-1985, Bachelor in Mathematics- Lecturer of Hue University, Vietnam.
-1998, Master in Information Technology, Lecturer of Nha Trang University, Vietnam.
-2006, Ph.D Student in Institute of Information Technology, Vietnamese academy of Science and Technology.
Current research: Rough set, Datamining, Distributed Database...
Second Author name Nguyen Xuan Huy was born in Hai Duong, Vietnam, 1944.
Brief Biographical History:
- 1973, Bachelor in Mathematics of Leningrad (Russia)
- 1982, 1990, Doctor and Post Doc in Information Technology of Russian Academy of Science and Technology
- At present, Associate Professor of Vietnamese Academy of Science and Technology

IACSIT
International Association of Computer Science and Information Technology
WWW.IACSIT.ORG