Knowledge-Based Systems 36 (2012) 39–50

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Evaluation of the decision performance of the decision rule set from an ordered decision table

Yuhua Qian [a,b], Jiye Liang [a,*], Peng Song [c], Chuangyin Dang [b], Wei Wei [a]

[a] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong
[c] School of Management, Shanxi University, Taiyuan, 030006 Shanxi, China

ARTICLE INFO

Article history:
Received 6 December 2011
Received in revised form 10 May 2012
Accepted 14 May 2012
Available online 1 June 2012

Keywords:
Ordered decision tables
Rough sets
Decision rules
Knowledge granulation
Decision evaluation

ABSTRACT

An ordered decision table is one of the most effective frameworks for the intelligent decision-making systems. As two classical measures, approximation accuracy and quality of approximation can be extended for evaluating the decision performance of an ordered decision table. However, from the viewpoint of evaluating the decision performance of a set of decision rules, these two measures are still not able to well measure the entire certainty and consistency of an ordered decision rule set. To overcome this deficiency, we first present three new measures for evaluating the decision performance of a decision-rule set extracted from an ordered decision table, and then analyze how each of these new measures depends on the condition granulation and the decision granulation of an ordered decision table. Applications and experimental analysis of five types of ordered decision tables show that the three new measures appear to be well suited for evaluating the decision performance of a decision-rule set extracted from each of these five types of decision tables and the results are much better than those of the two extended measures.

1. Introduction

Rough set theory proposed by Pawlak in [41,42] is a relatively new soft computing mechanism for the analysis of a vague description of an object, and has become a popular mathematical framework for such areas as pattern recognition, image processing, feature selection, neuro computing, conflict analysis, decision support, data mining and knowledge discovery process from large data sets [1,20,39,40,43–48,72,73]. The indiscernibility relation constitutes a mathematical basis of rough set theory. It induces a partition of the universe into blocks of indiscernible objects, called elementary sets, which can be used to build knowledge about a real or abstract world [37,42,52,55,65,66,68–71,75].

The original rough set theory does not consider attributes with preference-ordered domains, that is, criteria [68–71]. However, in many real situations, we often face problems in which the ordering of properties of the considered attributes plays a crucial role. One such type of problems is the ordering of objects. For this reason, Greco et al. [11,12] proposed an extension of rough set theory, called the dominance-based rough set approach (DRSA) to take into account the ordering properties of criteria. This innovation is mainly based on a substitution of the indiscernibility relation by a dominance relation. In DRSA, condition attributes are criteria, classes are preference ordered, the knowledge (approximated) is a collection of upward and downward unions of classes, and the granules of knowledge are sets of objects defined by using a dominance relation. In recent years, many studies in DRSA have been made [6,7,60,61,64]. DRSA starts from a so-called ordered decision table, which is used to extract a decision-rule set in practical decision problems.

For decision problems in rough set theory, by various kinds of reduction techniques, a set of decision rules is generated from a decision table for classification and prediction using information granules [5,18,26,31,62]. In the past two decades, many kinds of reduction techniques for information systems and decision tables have been proposed in rough set theory [4,23,27,34–38,42,43,59,63,67,74,75]. For our further developments, as follows, we briefly review some methods for attribute reduction from decision tables. β-reduct proposed by Ziarko provides a kind of attribute reduction methods for the variable precision rough set model [74]. α-reduct and α-relative reduct that allow the occurrence of additional inconsistency were proposed in [38] for information systems and decision tables, respectively. An attribute reduction method that preserves the class membership distribution of all objects in information systems was proposed by Slezak in [63,64]. Five kinds of attribute reducts and their relationships in inconsistent systems were investigated by Kryszkiewicz [23],

* Corresponding author. Tel./fax: +86 0351 7018176.
E-mail addresses: jinchengqyh@126.com (Y. Qian), ljy@sxu.edu.cn (J. Liang), songpeng@sxu.edu.cn (P. Song), mecdang@cityu.edu.hk (C. Dang), weiwei@sxu.edu.cn (W. Wei).

0950-7051/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2012.05.006

Li et al. [28] and Mi et al. [36], respectively. By eliminating some rigorous conditions required by the distribution reduct, a maximum distribution reduct was introduced by Mi et al. in [36]. Unlike the possible reduct [28], the maximum distribution reduct can derive decision rules that are compatible with the original system. Shao and Zhang proposed a kind of attribute reduction technique to reduce the number of criteria in an incomplete ordered information system and an incomplete ordered decision table [61].

Generally speaking, a set of decision rules can be generated from a decision table by adopting any kind of rule extracting methods. In recent years, the method of evaluating the decision performance of a decision rule has become a very important issue in rough set theory [17,19,29,31]. In [9], based on information entropy, Düntsch suggested some uncertainty measures of a decision rule and proposed three criteria for model selection. In [13], Greco et al. applied some well-known confirmation measures in the rough set approach to discover relationships in data in terms of decision rules. For a decision rule set consisting of every decision rule induced from a decision table, three parameters are traditionally associated: the strength, the certainty factor and the coverage factor of the rule [13]. In many practical decision problems, we always adopt several rule-extracting methods for the same decision table. In this case, it is very important to check whether or not each of the rule-extracting approaches adopted is suitable for a given decision table. In other words, it is desirable to evaluate the decision performance of the decision-rule set extracted by each of the rule-extracting approaches. This strategy can help a decision maker to determine which rule-extracting method should be adopted for a given decision table. However, all of the above measures are only defined for a single decision rule and are not suitable for evaluating the decision performance of a decision-rule set. There are two more kinds of measures in the literature [42,45], namely approximation accuracy for decision classification and consistency degree for a decision table. Although these two measures, in some sense, could be regarded as measures for evaluating the decision performance of all decision rules generated from a complete decision table, 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. We know that when the approximation accuracy or consistency degree is equal to zero, it only implies that there is no decision rule with the certainty of one in the decision table. This shows that the approximation accuracy and consistency degree of a decision table cannot be used to well measure the certainty and consistency of a rule set. To overcome the shortcomings of the existing measures, Qian et al. proposed four new evaluation measures for evaluating the decision performance of a set of decision rules extracted from a complete/incomplete decision table, which are certainty measure (α), consistency measure (β), support measure (γ) and covering measure (ϑ) [51,57].

Like that in the case of complete/incomplete decision tables, it is also very important to check whether or not each of the rule-extracting approaches adopted is suitable for a given ordered decision table. To date, however, no method for assessing the decision performance of a decision-rule set extracted from an ordered decision table has been reported. As mentioned by Greco et al. in [11], an ordered decision table can be interpreted as a set of ordered decision rules. In this study, we still read an ordered decision table as ordered decision rules. Like those for the existing measures, the certainty, consistency, support and covering of a decision-rule set extracted from an ordered decision table will also be analyzed in order to assess their decision performances. These measures are based on ordered decision rules instead of rough approximations for the dominance-based rough set approach. We know that each object can induce its corresponding dominance class and generate its corresponding ordered decision rules. Under this consideration, the support measure of each decision rule is easily determined by those objects that support the decision rule. With a view to having simplicity, we will not deal with the support measure γ in this paper.

In what follows, we explain the meaning of the certainty, consistency and covering measures from the viewpoint of a set of ordered decision rules from an ordered decision table, respectively.

- The certainty measure characterizes the entire certainty of all ordered decision rules from an ordered decision table. In other words, in some sense, this measure is to assess the average certainty of all extracted ordered decision rules. The greater the coefficient, the better the decision performance of these ordered decision rules.
- The consistency measure denotes the entire consistency degree of all ordered decision rules from an ordered decision table. If the certainty degree of each of ordered decision rules induced by a given condition class is equal to $1/2$, then the decision rules are the worst from the viewpoint of decision performance. In this situation, a decision maker does not know which ordered decision rule should be adopted. Using a fuzzy measure for evaluating this uncertainty, we can characterize the consistency of an ordered decision table by taking into consideration the fuzziness of each condition class. Like the certainty, the greater the coefficient, the better the decision performance of these ordered decision rules.
- The covering measure is also an important index for evaluating the decision performance of all ordered decision rules from an ordered decision table, which is used to measure the level of granulation determined by the condition classes of this decision table.

In fact, the approximation accuracy and consistency degree can be extended to evaluate the decision performance of the ordered decision rules from an ordered decision table. Nevertheless, these two extensions have the same limitations as the original measures and still cannot give elaborate depictions of the certainty and consistency of a decision-rule set extracted from an ordered decision table. If the approximation accuracy (or consistency degree) of one ordered decision table is the same as that of another ordered decision table, it does not imply that these two ordered decision tables have the same certainty/consistency, because that the measure cannot really reveal the certainty of an ordered decision table from the viewpoint of ordered decision rules. One should take into account the certainty of every ordered decision rule in evaluating the decision performance of an ordered decision table. It is worth pointing out that the existing four measures (α, β, γ and ϑ) are very disappointing at evaluating the decision performance of an ordered decision table in which the classes for constructing decision rules are not equivalence classes or tolerance classes, but dominance classes, and decision rules extracted are also not classical decision rules, but dominance rules. In particular, decision classes in ordered decision tables are a series of upward unions or downward unions, but not an equivalence partition. Hence, it is necessary to define several new measures for evaluating the decision performance of an ordered decision table. For this purpose, this paper introduces three new measures for evaluating the decision performance of a set of decision rules extracted from an ordered decision table, namely certainty measure (α), consistency measure (β) and covering measure (G).

The rest of this paper is organized as follows. Some preliminary concepts such as ordered information systems, ordered decision tables, dominance relation and decision rules are briefly reviewed in Section 2. In Section 3, we introduce some new concepts, reveal the limitations of the two extended measures, and propose three new measures (α, β and G) for evaluating the decision performance of a set of rules extracted from an ordered decision table. It is analyzed how each of these three measures depends on the condition granulation and decision granulation of an ordered decision table. In Section 4, applications and experimental analysis of each of the measures (α, β and G) are performed on five types of practical ordered decision tables. Finally, Section 5 concludes this paper with some remarks and discussion.

2. Preliminaries

In this section, we review some basic concepts of ordered information systems, ordered decision tables, dominance relation and ordered decision rules.

An information system (IS) is a quadruple $S = (U, AT, V, f)$, where U is a finite nonempty set of objects and AT is a finite nonempty set of attributes, $V = \bigcup_{a \in AT} V_a$ with V_a being a domain of attribute a, and $f : U \times AT \to V$ is a total function such that $f(x, a) \in V_a$ for every $a \in AT$ and $x \in U$, called an information function. A decision table is a special case of an information system in which, among all the attributes, we distinguish one (called a decision attribute) from the others (called condition attributes). Therefore, $S = (U, C \cup \{d\}, V, f)$ and $C \cap \{d\} = \emptyset$, where set C contains so-called condition attributes and d, the decision attribute.

If the domain (scale) of a condition attribute is ordered according to a decreasing or increasing preference, then the attribute is a criterion.

Definition 1 11. An information system is called an ordered information system (OIS) if all condition attributes are criteria.

It is assumed that the domain of a criterion $a \in AT$ is completely pre-ordered by an outranking relation $\succeq_a$; $x \succeq_a y$ means that x is at least as good as (outranks) y with respect to criterion a. In the following, without any loss of generality, we consider a condition criterion having a numerical domain, that is, $V_a \subseteq \mathbf{R}$ ($\mathbf{R}$ denotes the set of real numbers) and being of type gain, that is, $x \succeq_a y \Longleftrightarrow f(x, a) \geqslant f(y, a)$ (according to the increasing preference) or $x \succeq_a y \Longleftrightarrow f(x, a) \leqslant f(y, a)$ (according to decreasing preference), where $a \in AT, x, \ y \in U$. For a subset of attributes $B \subseteq C$, we say $x \succeq_B y$ if, for all $a \in B, \ f(x, a) \geqslant f(y, a)$. In other words, x is at least as good as y with respect to all attributes in B. In general, the domain of a condition criterion may be also discrete, but the preference order between its values has to be provided.

In the following, we review the dominance relation that identifies granules of knowledge. In a given OIS, we say that x dominates y with respect to $B \subseteq C$ if $x \succeq_B y$, and denote it by $xR_B^{\geqslant} y$ [11]. That is

$$R_B^{\geqslant} = \{(y, x) \in U \times U | y \succeq_B x\}.$$

Obviously, if $(y, x) \in R_B^{\geqslant}$, then y dominates x with respect to B.

Let B_1 be an attribute set according to an increasing preference and B_2 an attribute set according to a decreasing preference. Then, $B = B_1 \cup B_2$. The granules of knowledge induced by the dominance relation $R_B^{\geqslant}$ are the set of objects dominating x,

$$[x]_B^{\geqslant} = \{y \in U | f(y, a_1) \geqslant f(x, a_1) (\forall a_1 \in B_1),$$
$$f(y, a_2) \leqslant f(x, a_2) (\forall a_2 \in B_2)\} = \{y \in U | (y, x) \in R_B^{\geqslant}\},$$

and the set of objects dominated by x,

$$[x]_B^{\leqslant} = \{y \in U | f(y, a_1) \leqslant f(x, a_1) (\forall a_1 \in B_1),$$
$$f(y, a_2) \geqslant f(x, a_2) (\forall a_2 \in B_2)\} = \{y \in U | (x, y) \in R_B^{\geqslant}\},$$

which are called a B-dominating set and a B-dominated set with respect to $x \in U$, respectively. For simplicity, without any loss of generality, we only consider in the following the condition attributes with an increasing preference.

An ordered decision table (ODT) is an ordered information system $S = (U, C \cup \{d\}, V, f)$, where d is an overall preference called the decision, and all the elements of C are criteria. Furthermore, assume that the decision attribute d induces a partition of U into a finite number of classes; let $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ be an ordered set of these classes, that is, for all $i, j \leqslant r$, if $i \geqslant j$, then the objects from D_i are preferred to the objects from D_j. The sets to be approximated are an upward union and a downward union of classes [11], which are defined as follows:

$$D_i^{\geqslant} = \bigcup_{j \geqslant i} D_j, \quad D_i^{\leqslant} = \bigcup_{j \leqslant i} D_j, \quad (i, j \leqslant r).$$

The statement $x \in D_i^{\geqslant}$ means "x belongs to at least class D_i", whereas $x \in D_i^{\leqslant}$ means "x belongs to at most class D_i".

Definition 2 (11,12). Let $S = (U, C \cup \{d\}, V, f)$ be an ODT, $A \subseteq C$ and $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ the decision induced by d. Then, the lower and upper approximations of $D_i^{\geqslant}$ $(i \leqslant r)$ with respect to the dominance relation $R_A^{\geqslant}$ are defined by

$$\underline{R_A^{\geqslant}}(D_i^{\geqslant}) = \{x \in U | [x]_A^{\geqslant} \subseteq D_i^{\geqslant}\}, \quad \overline{R_A^{\geqslant}}(D_i^{\geqslant}) = \bigcup_{x \in D_i^{\geqslant}} [x]_A^{\geqslant}.$$

Similarly, one can define the lower and upper approximations of $D_i^{\leqslant} (i \leqslant r)$ with respect to the dominance relation $R_A^{\geqslant}$ in an ODT.

Definition 3 (11,12). Let $S = (U, C \cup \{d\}, V, f)$ be an ODT, $A \subseteq C$ and $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ the decision induced by d. Then, the lower and upper approximations of $D_i^{\leqslant}$ $(i \leqslant r)$ with respect to the dominance relation $R_A^{\geqslant}$ are defined by

$$\underline{R_A^{\geqslant}}(D_i^{\leqslant}) = \{x \in U | [x]_A^{\leqslant} \subseteq D_i^{\leqslant}\}, \quad \overline{R_A^{\geqslant}}(D_i^{\leqslant}) = \bigcup_{x \in D_i^{\leqslant}} [x]_A^{\leqslant}.$$

Naturally, the A-boundaries of $D_i^{\geqslant}$ $(i \leqslant r)$ and $D_i^{\leqslant}$ $(i \leqslant r)$ can be defined by

$$Bn_A(D_i^{\geqslant}) = \overline{R_A^{\geqslant}}(D_i^{\geqslant}) - \underline{R_A^{\geqslant}}(D_i^{\geqslant}), \quad Bn_A(D_i^{\leqslant}) = \overline{R_A^{\geqslant}}(D_i^{\leqslant}) - \underline{R_A^{\geqslant}}(D_i^{\leqslant}).$$

The lower approximations $\underline{R_A^{\geqslant}}(D_i^{\geqslant})$ and $\underline{R_A^{\geqslant}}(D_i^{\leqslant})$ can be used to extract certain decision rules, while the boundaries $Bn_A(D_i^{\geqslant})$ and $Bn_A(D_i^{\leqslant})$ can be used to induce possible decision rules from an ordered decision table.

In [60], an atomic expression over a single attribute a is defined as either $(a, \geqslant)$ (according to increasing preference) or $(a, \leqslant)$ (according to decreasing preference) in an ordered information system. For any $A \subseteq AT$, an expression over A in an ordered information system is defined by $\bigwedge_{a \in A} e(a)$, where $e(a)$ is an atomic expression over a. The set of all expressions over A in an OIS is denoted by $E(A)$. For instance, in Table 1, $AT = \{a_1, a_2, a_3\}$, the set of $E(AT)$ is as follows:

$$E(\{a_1, a_2, a_3\}) = \{(a_1, \geqslant) \wedge (a_2, \geqslant) \wedge (a_3, \geqslant), (a_1, \geqslant) \wedge (a_2, \geqslant) \wedge (a_3, \leqslant), \ldots, (a_1, \leqslant) \wedge (a_2, \leqslant) \wedge (a_3, \leqslant)\}.$$

In an OIS, for $a \in AT$ and $v_1 \in V_a$, an atomic formula over a single attribute a is defined as either $(a, \geqslant, v_1)$ (according to increasing preference) or $(a, \leqslant, v_1)$ (according to decreasing preference). For any $A \subseteq AT$, a formula over A in an OIS is defined by $\bigwedge_{a \in A} m(a)$, where $m(a)$ is an atomic formula over a. The set of all formulas over A in an OIS is denoted by $M(A)$. Let the formula $\phi \in M(A)$, and $\|\phi\|$ denotes the set of objects satisfying formula ϕ. For example, if $(a, \geqslant, v_1)$ and $(a, \leqslant, v_1)$ are atomic formulas, then

$$\|(a, \geqslant, v_1)\| = \{x \in U | f(x, a) \geqslant v_1\},$$
$$\|(a, \leqslant, v_1)\| = \{x \in U | f(x, a) \leqslant v_1\}.$$

Now we consider an ODT $S=(U,C\cup\{d\},V,f)$ and $A\subseteq C$. For two formulas $\phi\in M(A)$ and $\varphi\in M(d)$, a decision rule, denoted by $\phi\to\varphi$, is read as "if ϕ then φ." The formula ϕ is called the rule's *antecedent*, and the formula φ is called the rule's *consequent*. We say that an object supports a decision rule if it matches both the condition and the decision parts of the rule. On the other hand, an object is covered by a decision rule if it matches the condition part of the rule. A decision rule states how "evaluation of objects on attributes A is at least as good as a given level" or "evaluation of objects on attributes A is at most as good as a given level" determines "objects belong (or possibly belong) to at least a given class" or "objects belong (or possibly belong) to at most a given class." As follows, there are four types of decision rules to be considered [11,12]:

(1) certain $\geqslant$-decision rules with the following syntax: if $(f(x,a_1)\geqslant v_{a_1})\wedge(f(x,a_2)\geqslant v_{a_2})\wedge\cdots\wedge(f(x,a_k)\geqslant v_{a_k})\wedge(f(x,a_{k+1})\leqslant v_{a_{k+1}})\wedge\cdots\wedge(f(x,a_p)\leqslant v_{a_p})$, then $x\in D_i^{\geqslant}$;
(2) possible $\geqslant$-decision rules with the following syntax: if $(f(x,a_1)\geqslant v_{a_1})\wedge(f(x,a_2)\geqslant v_{a_2})\wedge\cdots\wedge(f(x,a_k)\geqslant v_{a_k})\wedge(f(x,a_{k+1})\leqslant v_{a_{k+1}})\wedge\cdots\wedge(f(x,a_p)\leqslant v_{a_p})$, then x could belong to $D_i^{\geqslant}$;
(3) certain $\leqslant$-decision rules with the following syntax: if $(f(x,a_1)\leqslant v_{a_1})\wedge(f(x,a_2)\leqslant v_{a_2})\wedge\cdots\wedge(f(x,a_k)\leqslant v_{a_k})\wedge(f(x,a_{k+1})\geqslant v_{a_{k+1}})\wedge\cdots\wedge(f(x,a_p)\geqslant v_{a_p})$, then $x\in D_i^{\leqslant}$;
(4) possible $\leqslant$-decision rules with the following syntax: if $(f(x,a_1)\leqslant v_{a_1})\wedge(f(x,a_2)\leqslant v_{a_2})\wedge\cdots\wedge(f(x,a_k)\leqslant v_{a_k})\wedge(f(x,a_{k+1})\geqslant v_{a_{k+1}})\wedge\cdots\wedge(f(x,a_p)\geqslant v_{a_p})$, then x could belong to $D_i^{\leqslant}$; where $O_1=\{a_1,a_2,\ldots,a_k\}\subseteq C, O_2=\{a_{k+1},a_{k+2},\ldots,a_p\}\subseteq C, C=O_1\cup O_2, O_1$ with increasing preference and O_2 with decreasing preference, $(v_{a_1},v_{a_2},\ldots,v_{a_p})\in V_{a_1}\times V_{a_2}\times\cdots\times V_{a_p}, i\leqslant r$.

Therefore, in an ODT, for a given upward or downward union $D_i^{\geqslant}$ or $D_j^{\leqslant},\ i,j\leqslant r$, the rules induced under a hypothesis that objects belonging to $\underline{R_A^{\geqslant}}(D_i^{\geqslant})$ or to $\underline{R_A^{\leqslant}}\left(D_j^{\leqslant}\right)$ are positive and all the others negative suggest the assignment of an object to "at least class D_i" or to "at most class D_j", respectively. Similarly, the rules induced under a hypothesis that objects belonging to $\overline{R_A^{\geqslant}}(D_i^{\geqslant})$ or to $\overline{R_A^{\leqslant}}\left(D_j^{\leqslant}\right)$ are positive and all the others negative suggest the assignment of an object could belong to "at least class D_i" or to "at most class D_j", respectively.

From the definitions of $D_i^{\geqslant}$ and $D_i^{\leqslant}$, it is easy to see that there is a complement relation between $D_i^{\geqslant}$ and $D_{i-1}^{\leqslant}$. Therefore, in this paper, we only investigate the former two types of decision rules, i.e., the decision rules induced by $D_i^{\geqslant}$. Let $S=(U,C\cup\{d\},V,f)$ be an ODT, $A\subseteq C$ and $\mathbf{D}=\{D_1,D_2,\ldots,D_r\}$ the decision induced by d. For our further development, we denote a decision rule by

$$Z_{ij}:\ des([x_i]_A^{\geqslant})\to\left(x_i\in D_j^{\geqslant}\right),\ i\leqslant|U|,\ j\leqslant r,$$

where $des([x_i]_A^{\geqslant})$ denotes the description (i.e., the condition part of each of the above four kinds of decision rules) of the dominance class $[x_i]_A^{\geqslant}$ in S.

3. Three measures for evaluating the decision performance of an ordered decision table

In this section, by introducing a partial relation in an ordered information system and an ordered decision table, three measures are proposed for evaluating the decision performance of an ordered decision table, which are certainty measure (α), consistency measure (β) and covering measure (G). Furthermore, it is analyzed how each of these three measures depends on the condition granulation and the decision granulation of an ordered decision table as well.

In the first part of this section, we introduce several new concepts and notations, which will be applied in what follows.

Let $S=(U,AT,V,f)$ be an ordered information system, $P,Q\subseteq AT$, $U/R_P^{\geqslant}=\{[x_1]_P^{\geqslant},[x_2]_P^{\geqslant},\ldots,[x_{|U|}]_P^{\geqslant}\}$ and $U/R_Q^{\geqslant}=\left\{[x_1]_Q^{\geqslant},[x_2]_Q^{\geqslant},\ldots,[x_{|U|}]_Q^{\geqslant}\right\}$. We define a partial relation $\preceq$ as follows: $P\preceq Q\iff[x_i]_P^{\geqslant}\subseteq[x_i]_Q^{\geqslant}$ for any $x_i\in U$, where $[x_i]_P^{\geqslant}\in U/R_P^{\geqslant}$ and $[x_i]_Q^{\geqslant}\in U/R_Q^{\geqslant}$. If $P\preceq Q$, we say that Q is coarser than P (or P is finer than Q).

Let $S=(U,C\cup\{d\},V,f)$ be an ordered decision table, $U/R_C^{\geqslant}=\{[x_1]_C^{\geqslant},[x_2]_C^{\geqslant},\ldots,[x_{|U|}]_C^{\geqslant}\}$ and $U/R_d^{\geqslant}=\{[x_1]_d^{\geqslant},[x_2]_d^{\geqslant},\ldots,[x_{|U|}]_d^{\geqslant}\}$. If $C\preceq\{d\}$, then S is said to be a consistent ordered decision table; otherwise, S is said to be inconsistent.

In general, knowledge granulation is employed to measure the discernibility ability of knowledge in rough set theory. The smaller granulation of knowledge, the stronger its discernibility ability [50,53,56,58]. Liang et al. introduced a knowledge granulation $G(A)$ to measure the discernibility ability of knowledge in an information system [32,33]. In [52], Qian and Liang proposed another kind of knowledge granulations, called combination granulations, in complete and incomplete information systems. In [30], Liang and Qian established an axiomatic approach of knowledge granulation in information systems. Accordingly, we introduce a new knowledge granulation to measure the discernibility ability of knowledge in an ordered information system, which is given in the following definition.

Definition 4. Let $S=(U,AT,V,f)$ be an ordered information system and $U/R_{AT}^{\geqslant}=\{[x_1]_{AT}^{\geqslant},[x_2]_{AT}^{\geqslant},\ldots,[x_{|U|}]_{AT}^{\geqslant}\}$. Knowledge granulation of AT is defined as

$$G(AT)=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_{AT}^{\geqslant}|}{|U|}. \quad (1)$$

Following this definition, for a given ordered decision table $S=(U,C\cup\{d\},V,f)$, we call $G(C)$, $G(d)$ and $G(C\cup d)$ condition granulation, decision granulation and granulation of S, respectively.

As a result of the above discussion, we come to the following theorem.

Theorem 1. *Let $S=(U,AT,V,f)$ be an ordered information system and $P,Q\subseteq AT$ with $P\preceq Q$. Then, $G(P)\leqslant G(Q)$.*

In rough set theory, several measures for a decision rule $Z_{ij}:des(X_i)\to des(Y_j)$ have been introduced in [42], such as certainty measure $\mu(X_i,Y_j)=|X_i\cap Y_j|/|X_i|$, support measure $s(X_i,Y_j)=|X_i\cap Y_j|/|U|$ and coverage measure $\tau(X_i,Y_j)=|X_i\cap Y_j|/|Y_j|$. Naturally, the extensions of these measures are also suitable for evaluating the decision performance of a decision rule extracted from an ordered decision table. However, because $\mu(X_i,Y_j),\ s(X_i,Y_j)$ and $\tau(X_i,Y_j)$ are only defined for a single decision rule, they are not suitable for evaluating the decision performance of a decision-rule set extracted from an ordered decision table.

In [42], approximation accuracy of a classification is introduced by Pawlak. Let $F=\{Y_1,Y_2,\ldots,Y_n\}$ be a classification or decision of the universe U (it can be regarded as a partition induced by decision attribute set D in a decision table, i.e., $F=U/D$) and C a condition attribute set. $\underline{C}F=\{\underline{C}Y_1,\underline{C}Y_2,\ldots,\underline{C}Y_n\}$ and $\overline{C}F=\{\overline{C}Y_1,\overline{C}Y_2,\ldots,\overline{C}Y_n\}$ are called *C-lower* and *C-upper* approximations of F, respectively, where $\underline{C}Y_i=\bigcup\{x\in U|[x]_C\subseteq Y_i\in F\}\quad(1\leqslant i\leqslant n)$ and $\overline{C}Y_i=\bigcup\{x\in U|[x]_C\cap Y_i\neq\emptyset,Y_i\in F\}$ $(1\leqslant i\leqslant n)$. The approximation accuracy of F by C is defined as

$$a_C(F)=\frac{\sum_{Y_i\in U/D}|\underline{C}Y_i|}{\sum_{Y_i\in U/D}|\overline{C}Y_i|}. \quad (2)$$

Table 1
An ordered decision table.

U	a_1	a_2	d
x_1	1	2	1
x_2	3	2	2
x_3	1	1	1
x_4	2	1	2
x_5	3	3	1
x_6	3	2	2

It is the percentage of possible correct decisions when classifying objects by employing the attribute set C.

In an ordered decision table, similar to formula (2), the approximation accuracy of $\mathbf{D}$ by C can be defined as

$$a_C(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}\left|\overline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}. \tag{3}$$

According to Pawlak's viewpoint, $a_C(\mathbf{D})$ can be used to measure the certainty of an ordered decision table. However, it has some limitations, one of which is illustrated in the following example.

Example 1. An ODT is presented in Table 1, where $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ and $C = \{a_1, a_2\}$.

In this table, from the definition of dominance classes, one can obtain that the dominance classes determined by $\{a_1\}$ and C are

$$[x_1]_{\{a_1\}}^{\geqslant} = [x_3]_{\{a_1\}}^{\geqslant} = \{x_1, x_2, x_3, x_4, x_5, x_6\}, \quad [x_2]_{\{a_1\}}^{\geqslant} = [x_5]_{\{a_1\}}^{\geqslant} = [x_6]_{\{a_1\}}^{\geqslant} = \{x_2, x_5, x_6\}, \quad [x_4]_{\{a_1\}}^{\geqslant} = \{x_2, x_4, x_5, x_6\};$$

$$[x_1]_C^{\geqslant} = \{x_1, x_2, x_5, x_6\}, \quad [x_2]_C^{\geqslant} = [x_6]_C^{\geqslant} = \{x_2, x_5, x_6\}, \quad [x_3]_C^{\geqslant} = \{x_1, x_2, x_3, x_4, x_5, x_6\}, \quad [x_4]_C^{\geqslant} = \{x_2, x_4, x_5, x_6\}, \quad [x_5]_C^{\geqslant} = \{x_5\};$$

and the ordered classes determined by d are

$$D_1^{\geqslant} = \{x_1, x_2, x_3, x_4, x_5, x_6\} \quad \text{and} \quad D_2^{\geqslant} = \{x_2, x_4, x_6\}.$$

Therefore,

$$a_{\{a_1\}}(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}\left|\overline{R_C^{\geqslant}}(D_i^{\geqslant})\right|} = \frac{6+0}{6+4} = 0.6 \quad \text{and } a_C(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}\left|\overline{R_C^{\geqslant}}(D_i^{\geqslant})\right|} = \frac{6+0}{6+4} = 0.6.$$

That is to say $a_{\{a_1\}}(\mathbf{D}) = a_C(\mathbf{D})$. This implies that the relation $C < \{a_1\}$ ($<$ denotes finer) is not revealed by the extended approximation accuracy.

In fact, the shortcoming is mainly caused by the construction of the coefficient. The measure cannot really reveal the certainty of the decision rule set from an ordered decision table. To overcome this deficiency, one should take into account the certainty of every ordered decision rule for evaluating the entire certainty. For a consistent ordered decision table, the certainty of each ordered decision rule is equal to one. On the other side, in an inconsistent ordered decision table, there exists at least one dominance class in the condition part that cannot be included in the lower approximation of the target decision. This dominance class can induce some uncertain ordered decision rules. Hence, one can draw the conclusion that the extension of the approximation accuracy can not be employed to effectively evaluate the decision performance of an ordered decision table. To overcome this drawback of the extended measures, any new measure should take into account the certainty of each ordered decision rule in evaluating the decision performance of the decision rule set from an ordered decision table. Therefore, a more comprehensive and effective measure for evaluating the certainty of the decision rule set from an ordered decision table is desired.

The consistency degree of a complete decision table $S = (U, C \cup D)$, another important measure proposed in [42], is defined as

$$c_C(D) = \frac{\sum_{i=1}^{n}|\underline{C}Y_i|}{|U|}. \tag{4}$$

It is the percentage of objects which can be correctly classified to decision classes of U/D by a condition attribute set C. In some situations, $c_C(D)$ can be employed to evaluate the consistency of a decision table.

The consistency degree of an ordered decision table is defined as

$$c_C(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}|D_i^{\geqslant}|}. \tag{5}$$

For an ordered decision table, one can also extend the consistency degree for measuring the consistency of a decision-rule set. However, similar to formula (3), the extended consistency degree cannot well characterize the consistency of an ordered decision table because it only considers the lower approximation of a target decision. This is revealed in the following example.

Example 2 (*Continued from Example 1*). Computing the consistency degree, we have that

$$c_{\{a_1\}}(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}|D_i^{\geqslant}|} = \frac{6+0}{6+3} = 0.6667 \quad \text{and}$$

$$c_C(\mathbf{D}) = \frac{\sum_{i=1}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right|}{\sum_{i=1}^{r}|D_i^{\geqslant}|} = \frac{6+0}{6+3} = 0.6667.$$

Obviously, $c_{\{a_1\}}(\mathbf{D}) = c_C(\mathbf{D})$. In other words, one can draw the conclusion that the extension of the consistency degree cannot be employed to effectively evaluate the consistency of an ordered decision table.

In [11] Greco et al. extended the quality of approximation to ordered decision tables, which is defined by the following form

$$\Upsilon_C(\mathbf{D}) = \frac{\left|U - \left(\left(\bigcup_{i\leqslant r} Bn_C(D_i^{\geqslant})\right) \cup \left(\bigcup_{i\leqslant r} Bn_C(D_i^{\leqslant})\right)\right)\right|}{|U|}.$$

In addition, Dembczynski et al. [8] proposed another form of the quality of approximation, which is equivalent to the quality of approximation

$$\Upsilon_C(\mathbf{D}) = \frac{\sum_{i=2}^{r}\left|\underline{R_C^{\geqslant}}(D_i^{\geqslant})\right| + \sum_{i=1}^{r-1}\left|\underline{R_C^{\geqslant}}(D_i^{\leqslant})\right|}{\sum_{i=2}^{r}|D_i^{\geqslant}| + \sum_{i=1}^{r-1}|D_i^{\leqslant}|}$$

defined by Düntsch and Gediga [10]. These two measures are both used to characterize the average relative width of C-generalized decisions of reference objects [8]. However they also cannot well characterize the decision performance of an ordered decision table from the viewpoint of ordered decision rules. These measures have the same limitations as the approximation accuracy and consistency degree, which are also based on the lower/upper approximations in the dominance-based rough set approach. Thus, to depict the decision performance of an ordered decision rule set, a more comprehensive and effective measure is desired for evaluating the consistency of the decision rules set from an incomplete ordered decision table.

In order to evaluate the decision performance of a decision-rule set extracted from a complete/incomplete decision table, one must take into consideration three important factors, that is, the certainty, consistency and support of the decision-rule set [49,51,57]. For decision problems in ordered decision tables, these three factors also play important roles. Furthermore, the degree of the covering induced by the dominance classes in the condition part can affect the decision performance of a decision-rule set extracted from an ordered decision table. However, since the support measure of each decision rule from a given ordered decision table is one,[1] this measure will be ignored in this paper.

In the next part, we deal with how to evaluate the decision performance of the decision rule set from an ordered decision table. Firstly, we investigate the certainty of an ordered decision rule set.

Definition 5. Let $S=(U,C\cup\{d\},V,f)$ be an ordered decision table, $A\subseteq C$, $U/R_A^{\geqslant}=\{[x_1]_A^{\geqslant},[x_2]_A^{\geqslant},\ldots,[x_{|U|}]_A^{\geqslant}\}$, $\mathbf{D}=\{D_1,D_2,\ldots,D_r\}$ and $RULE=\left\{Z_{ij}|Z_{ij}:\ des([x_i]_A^{\geqslant})\rightarrow\left(x\in D_j^{\geqslant}\right),\ i\leqslant|U|,\ j\leqslant r\right\}$. Certainty measure α of *RULE* is defined as

$$\alpha(S)=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{\left|[x_i]_A^{\geqslant}\cap D_j^{\geqslant}\right|}{\left|[x_i]_A^{\geqslant}\right|},\tag{6}$$

where N_i is the number of ordered decision classes with nonempty intersection with the dominance class $[x_i]_A^{\geqslant}$ in the ordered decision table.

The mechanism of this definition is illustrated by the following example.

Example 3 (*Continued from Example 1*). Let S_1 be the ordered decision table induced by $\{a_1\}$ and S_2 the ordered decision table induced by C. Computing the certainty measure, we have that

$$\begin{aligned}\alpha(S_1)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{\left|[x_i]_C^{\geqslant}\cap D_j^{\geqslant}\right|}{\left|[x_i]_C^{\geqslant}\right|}\\&=\frac{1}{6}\left[\frac{1}{2}\left(1+\frac{1}{2}\right)\times2+\frac{1}{2}\left(1+\frac{2}{3}\right)\times3+\frac{1}{2}\left(1+\frac{3}{4}\right)\right]=0.8125\end{aligned}$$

and

$$\begin{aligned}\alpha(S_2)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{\left|[x_i]_C^{\geqslant}\cap D_j^{\geqslant}\right|}{\left|[x_i]_C^{\geqslant}\right|}\\&=\frac{1}{6}\left[\frac{1}{2}\left(1+\frac{1}{2}\right)\times2+\frac{1}{2}\left(1+\frac{2}{3}\right)\times2+\frac{1}{2}\left(1+\frac{3}{4}\right)+1\right]\\&=0.8403.\end{aligned}$$

That is $\alpha(S_2)>\alpha(S_1)$. Thus, the measure α is much better than the extended approximation accuracy for measuring the certainty of the decision rule set from an inconsistent ordered decision table.

In what follows, we discuss the monotonicity of measure α in an ordered decision table.

Theorem 2. *Let $S_1=(U,C_1\cup\{d_1\},V_1,f_1)$ and $S_2=(U,C_2\cup\{d_2\},V_2,f_2)$ be two ordered decision tables. If $U/R_{C_1}^{\geqslant}=U/R_{C_2}^{\geqslant}$ and $d_1\preceq d_2$, then $\alpha(S_1)\leqslant\alpha(S_2)$.*

[1] From the definition of an ordered decision rule, we know that the ordered decision rule induced by an object is only supported by itself in an ordered decision table, and its support measure is equal to one.

Proof. Let $\mathbf{D}_1=\{D_1,D_2,\ldots,D_r\}$ and $\mathbf{D}_2=\{K_1,K_2,\ldots,K_s\}$ be the ordered decisions of S_1 and S_2, respectively. From $d_1\preceq d_2$, it follows that $r\geqslant s$, and there exists some partition $T=\{T_1,T_1,\ldots,T_s\}$ of $\{1,2,\ldots,r\}$ such that $K_t=\bigcup_{k\in T_t}D_k,\ t=1,2,\ldots,s$. Hence, for any $D_j\in\mathbf{D}_1$, there exists some $K_t\in\mathbf{D}_2$ such that $D_j\subseteq K_t$. Thus, one has that $D_j^{\geqslant}\subseteq K_t^{\geqslant}$ and $[x_i]_C^{\geqslant}\cap D_j^{\geqslant}\subseteq[x_i]_C^{\geqslant}\cap K_t^{\geqslant}$. Let $N_i(S_1)$ and $N_i(S_2)$ denote the number of ordered decision classes induced by the dominance classes $[x_i]_{C_1}^{\geqslant}$ and that induced by $[x_i]_{C_2}^{\geqslant}$, respectively. So, it follows from $d_1\preceq d_2$ that $N_i(S_1)\geqslant N_i(S_2)$. And since $U/R_{C_1}^{\geqslant}=U/R_{C_2}^{\geqslant}$, one has that $[x_i]_{C_1}^{\geqslant}=[x_i]_{C_2}^{\geqslant},\ i\leqslant|U|$. Therefore,

$$\begin{aligned}\alpha(S_1)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{N_i(S_1)}\sum_{j=1}^{N_i(S_1)}\frac{\left|[x_i]_{C_1}^{\geqslant}\cap D_j^{\geqslant}\right|}{\left|[x_i]_{C_1}^{\geqslant}\right|}\\&\leqslant\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{N_i(S_2)}\sum_{t=1}^{N_i(S_2)}\frac{\left|[x_i]_{C_2}^{\geqslant}\cap K_t^{\geqslant}\right|}{\left|[x_i]_{C_2}^{\geqslant}\right|}=\alpha(S_2).\end{aligned}$$

This completes the proof. □

Theorem 2 states that the certainty measure α of all decision rules from an ordered decision table decreases as its ordered decision classes becomes finer.

Next, we discuss the consistency of the decision rule set from an ordered decision table.

Definition 6. Let $S=(U,C\cup\{d\},V,f)$ be an ordered decision table, $A\subseteq C$, $U/R_A^{\geqslant}=\{[x_1]_A^{\geqslant},[x_2]_A^{\geqslant},\ldots,[x_{|U|}]_A^{\geqslant}\}$, $\mathbf{D}=\{D_1,D_2,\ldots,D_r\}$ and $RULE=\left\{Z_{ij}|Z_{ij}:\ des([x_i]_A^{\geqslant})\rightarrow\left(x\in D_j^{\geqslant}\right),\ i\leqslant|U|,\ j\leqslant r\right\}$. Consistency measure β of *RULE* is defined as

$$\beta(S)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{r}\sum_{j=1}^{r}\mu(Z_{ij})(1-\mu(Z_{ij}))\right],\tag{7}$$

where $\mu(Z_{ij})=\frac{\left|[x_i]_A^{\geqslant}\cap D_j^{\geqslant}\right|}{\left|[x_i]_A^{\geqslant}\right|}$ is the certainty degree of the decision rule Z_{ij}. The following example will be helpful for understanding the meaning of this definition.

Example 4 (*Continued from Example 3*). Computing the measure β, we have that

$$\begin{aligned}\beta(S_1)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{r}\sum_{j=1}^{r}\mu(Z_{ij})(1-\mu(Z_{ij}))\right]\\&=\frac{1}{6}\left[\left(1-\frac{1}{2}\right)\times2+\left(1-\frac{4}{9}\right)\times3+\left(1-\frac{3}{8}\right)\right]=0.5486\end{aligned}$$

and

$$\begin{aligned}\beta(S_2)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{r}\sum_{j=1}^{r}\mu(Z_{ij})(1-\mu(Z_{ij}))\right]\\&=\frac{1}{6}\left[\left(1-\frac{1}{2}\right)\times2+\left(1-\frac{4}{9}\right)\times2+\left(1-\frac{3}{8}\right)+(1-0)\right]=0.6227.\end{aligned}$$

That is $\beta(S_2)>\beta(S_1)$. It can be interpreted in the sense that ordered decision table S_2 has much bigger consistency and much smaller fuzziness than S_1. Unlike the extended consistency degree, the measure β can be used to evaluate the consistency of an ordered decision table.

In the following, we investigate the monotonicity of the measure β in an ordered decision table.

Theorem 3. *Let $S_1=(U,C_1\cup\{d_1\},V_1,f_1)$ and $S_2=(U,C_2\cup\{d_2\},V_2,f_2)$ be two ordered decision tables. If $U/R_{C_1}^{\geqslant}=U/R_{C_2}^{\geqslant}$ and $d_1\preceq d_2$, then $\beta(S_1)\leqslant\beta(S_2)$ for $\forall\mu(Z_{ij})\geqslant\frac{1}{2}$.*

Proof. From Definition 6, it follows that

$$\beta(S)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{r}\sum_{j=1}^{r}\mu(Z_{ij})(1-\mu(Z_{ij}))\right]=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{4}{r}\sum_{j=1}^{r}\left(\mu(Z_{ij})-\frac{1}{2}\right)^2.$$

Let $\mathbf{D}_1=\{D_1,D_2,\ldots,D_r\}$ and $\mathbf{D}_2=\{K_1,K_2,\ldots,K_s\}$ be the ordered decisions of S_1 and S_2, respectively. From $d_1\preceq d_2$, it follows that $r\geqslant s$, and there exists some partition $T=\{T_1,T_1,\ldots,T_s\}$ of $\{1,2,\ldots,r\}$ such that $K_t=\bigcup_{k\in T_t}D_k,\ t=1,2,\ldots,s$. Hence, for any $D_j\in\mathbf{D}_1$, there exists some $K_t\in\mathbf{D}_2$ such that $D_j\subseteq K_t$. Thus, one has that $D_j^{\geqslant}\subseteq K_t^{\geqslant}$ and $[x_i]_C^{\geqslant}\cap D_j^{\geqslant}\subseteq[x_i]_C^{\geqslant}\cap K_t^{\geqslant}$. And since $U/R_{C_1}^{\geqslant}=U/R_{C_2}^{\geqslant}$, one has that $[x_i]_{C_1}^{\geqslant}=[x_i]_{C_2}^{\geqslant},\ i\leqslant|U|$. So, it follows that $\mu(Z_{ij})\leqslant\mu(Z_{it})$. Therefore, when $\forall\mu(Z_{ij})\geqslant\frac{1}{2}$, we have that

$$\begin{aligned}\beta(S_1)&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{r}\sum_{j=1}^{r}\mu(Z_{ij})(1-\mu(Z_{ij}))\right]\\&=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{4}{r}\sum_{j=1}^{r}\left(\mu(Z_{ij})-\frac{1}{2}\right)^2\leqslant\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{4}{s}\sum_{k=1}^{s}\left(\mu(Z_{ik})-\frac{1}{2}\right)^2\\&=\frac{1}{|U|}\sum_{i=1}^{|U|}\left[1-\frac{4}{s}\sum_{k=1}^{s}\mu(Z_{ik})(1-\mu(Z_{ik}))\right]=\beta(S_2).\end{aligned}$$

That is $\beta(S_1)\leqslant\beta(S_2)$. This completes the proof. □

Theorem 3 shows that the consistency measure β of all decision rules from an ordered decision table decreases with decision classes becoming finer for $\forall\mu(Z_{ij})\geqslant\frac{1}{2}$.

From these two definitions and their properties, it can be seen that their successes are because that the two measures are constructed through considering certainty/consistency of each ordered decision rule from a given ordered decision table. From this idea, these two proposed measures can characterize the entire decision performance of an ordered decision rule set, and the old ones can not do.

It is worth pointing out that the values of the two new measures (α and β), in some sense, are dependent on the situation of the covering induced by the dominance classes in the condition part of an ordered decision table. In the following, we investigate how to measure the degree of the covering in the condition part of an ordered decision table. In fact, from the viewpoint of granular computing, the degree of the covering is also seen as the level of granulation of objects. Knowledge granulation in Definition 4 can be used to characterize the degree of the covering. In order to characterize the covering in ordered decision tables, we call it the knowledge granulation covering measure, still denoted by G.

4. Evaluations on the performance of an ordered decision table

In this section, we will apply the three measures (α,β and ϑ) proposed in this paper to five types of ordered decision tables and demonstrate through experimental analysis the validity and effectiveness of each of them for evaluating the decision performance of each of these five types of ordered decision tables through experimental analysis. The five types of ordered decision tables are single-valued ordered decision tables, incomplete ordered decision tables, interval ordered decision tables, disjunctive set-valued ordered decision tables and conjunctive set-valued ordered decision tables.

4.1. Five types of ordered decision tables

4.1.1. Single-valued ordered decision tables

A *single-valued ordered decision table* is an ordered information system $S=(U,C\cup\{d\},V,f)$, where d ($d\notin C$ and $f(x,a)$, $f(x,d)$ $(x\in U,a\in C)$ are all single-valued) is an overall preference called the decision and all the elements of C are criteria. Furthermore, assume that the decision attribute d induces a partition of U into a finite number of classes; let $\mathbf{D}=\{D_1,D_2,\ldots,D_r\}$ be an ordered set of these classes, that is, for all $i,j\leqslant r$, if $i\geqslant j$, then the objects from D_i are preferred to the objects from D_j. In fact, the type of ordered decision tables discussed in Section 2 are single-valued ordered decision tables.

4.1.2. Incomplete ordered decision tables

An *incomplete ordered decision table* (IODT) is an incomplete ordered information system $S=(U,C\cup\{d\},V,f)$, where d ($d\notin C$ and $f(x,d)$ $(x\in U)$ is single-valued) is an overall preference called the decision and all the elements of C are criteria. In [11], Greco et al. proposed a general framework for incomplete ordered decision tables. Let $R_A^{*\geqslant}$ with $A\subseteq AT$ denote a dominance relation between objects that are possibly dominant in terms of values of attributes set A, in which "$*$" denotes a missing value [2,3,15,16,21–25]. The dominance relation is defined by

$$\begin{aligned}R_A^{*\geqslant}&=\{(y,x)\in U\times U|\forall a\in A,\ f(y,a)\geqslant f(x,a)\quad\text{or } f(x,a)\\&\quad=*\ \text{or } f(y,a)=*\}.\end{aligned}$$

By the definition of $R_A^{*\geqslant}$, it can be observed that if a pair of objects (y,x) from $U\times U$ is in $R_A^{*\geqslant}$, then they are perceived as y dominates x; in other words, y may have a better property than x with respect to A in reality. Defined by

$$[x]_A^{*\geqslant}=\{y\in U|(y,x)\in R_A^{*\geqslant}\},$$

$[x]_A^{*\geqslant}$ describes objects that may dominate x in terms of A. Let $U/R_A^{*\geqslant}$ denote classification, which is the family set $\{[x]_A^{*\geqslant}|x\in U\}$. Any element from $U/R_A^{*\geqslant}$ will be called a dominance class. The lower and upper approximations of $D_i^{\geqslant}$ with respect to the dominance relation $R_A^{*\geqslant}$ are defined in [11,61] as

$$\underline{R_A^{*\geqslant}}(D_i^{\geqslant})=\{x\in U|[x]_A^{*\geqslant}\subseteq D_i^{\geqslant}\},\quad\overline{R_A^{*\geqslant}}(D_i^{\geqslant})=\bigcup_{x\in D_i^{\geqslant}}[x]_A^{*\geqslant}.$$

4.1.3. Interval ordered decision tables

Interval information systems are an important type of data tables, and generalized models of single-valued information systems. An *interval information system* (IIS) is a quadruple $S=(U,AT,V,f)$, where U is a finite non-empty set of objects and AT is a finite non-empty set of attributes, $V=\bigcup_{a\in AT}V_a$ and V_a is a domain of attribute $a,f:U\times AT\to V$ is a total function such that $f(x,a)\in V_a$ for every $a\in AT,\ x\in U$, called an information function, where V_a is a set of interval numbers. Denoted by

$$f(x,a)=[a^L(x),a^U(x)]=\{p|a^L(x)\leqslant p\leqslant a^U(x),a^L(x),a^U(x)\in\mathbf{R}\},$$

we call it the interval number of x under the attribute a. In particular, $f(x,a)$ would degenerate into a real number if $a^L(x)=a^U(x)$. Under this consideration, we regard a single-valued information system as a special form of interval information systems.

Given $A\subseteq AT$ with increasing preference, we define a dominance relation $R_A^{\sqsupseteq}$ in interval ordered information systems as follows:

$$R_A^{\sqsupseteq}=\{(y,x)\in U\times U|a^L(y)\geqslant a^L(x),\ a^U(y)\geqslant a^U(x),\ \forall a\in A).$$

The dominance classes induced by the dominance relation $R_A^{\sqsupseteq}$ are the set of objects dominating x, that is,

$$[x]_A^{\sqsupseteq}=\{y\in U|a^L(y)\geqslant a^L(x),\ a^U(y)\geqslant a^U(x),\ \forall a\in A).$$

An *interval ordered decision table* (IODT) is an interval ordered information system $S=(U,C\cup d,V,f)$, where d ($d\notin C$ and $f(x,d)(x\in U)$ is single-valued) is an overall preference called the decision and all the elements of C are criteria. Furthermore, assume that the decision attribute d induces a partition of U into a finite number of classes; let $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ be a set of these classes that are ordered, that is, for all $i, j \leqslant r$, if $i \geqslant j$, then the objects from D_i are preferred to the objects from D_j. Table 2 gives an interval ordered decision table.

Let $S = (U, C \cup d, V, f)$ be an IODT, $A \subseteq C$ and $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ be the decision induced by d. Lower and upper approximations of $D_i^{\geqslant}$ $(i \leqslant r)$ with respect to the dominance relation $R_A^{\sqsupseteq}$ are defined as [54] 1 mm

$$\underline{R_A^{\sqsupseteq}}(D_i^{\geqslant}) = \{x \in U | [x]_A^{\sqsupseteq} \subseteq D_i^{\geqslant}\}, \quad \overline{R_A^{\sqsupseteq}}(D_i^{\geqslant}) = \bigcup_{x \in D_i^{\geqslant}} [x]_A^{\sqsupseteq}.$$

4.1.4. Conjunctive set-valued ordered decision tables

Set-valued information systems are another important type of data tables, and generalized models of single-valued information systems. Let U be a finite set of objects, called the universe of discourse, and AT be a finite set of attributes. With every attribute $a \in AT$, a set of its values V_a is associated. $f : U \times AT \to V$ is a total function such that $f(x, a) \subseteq V_a$ for every $a \in AT$, $x \in U$. If each attribute has a unique attribute value, then (U, AT, V, f) with $V = \cup_{a \in AT} V_a$ is called a single-valued information system; if a system is not a single-valued information system, it is called a set-valued (multi-valued) information system. A set-valued decision table is always denoted by $S = (U, C \cup \{d\}, V, f)$, where C is a finite set of condition attributes and d is a decision attribute with $C \cap d = \emptyset$.

There are many ways to give a semantic interpretation of the set-valued information systems. Here we summarize them as two types [14]: conjunctive set-valued information systems and disjunctive set-valued information systems. In this section, through introduction of a dominance relation to a conjunctive set-valued information system, we investigate conjunctive set-valued ordered decision tables and dominance decision rules extracted from this type of decision tables, and apply the three measures (α, β and ϑ) for evaluating the decision performance of an conjunctive set-valued ordered decision table.

For $x \in U$ and $c \in C$, $c(x)$ is interpreted conjunctively. For example, if c is the attribute "speaking a language", then $c(x) = \{\text{German}, \text{Polish}, \text{French}\}$ can be interpreted as: x speaks German, Polish, and French. When considering the attribute "feeding habits" of animals, if we denote the attribute value of herbivore as "0" and carnivore as "1", then animals possessing attribute value $\{0, 1\}$ are considered as possessing both herbivorous and carnivorous nature. Let us take blood origin for another example. If we denote the three types of pure blood as "0", "1" and "2", then we can denote the mixed-blood as $\{0, 1\}$ or $\{1, 2\}$, etc. Under this interpretation, we say it is a "$\bigwedge$" set-valued information system in this paper.

In what follows, we define a dominance relation $R_A^{\wedge\geqslant}$ in a "$\bigwedge$" set-valued information system as [53]

$$R_A^{\wedge\geqslant} = \{(y, x) \in U \times U | f(y, a) \supseteq f(x, a), \ \forall a \in A)\}.$$

By the definition of the dominance relation $R_A^{\wedge\geqslant}$, it can be observed that if a pair of objects (y, x) from $U \times U$ lies in $R_A^{\wedge\geqslant}$, then they are perceived as y dominates x; in other words, y may have a better property than x with respect to A in reality. Furthermore, denoted by

$$[x]_A^{\wedge\geqslant} = \{y \in U | (y, x) \in R_A^{\wedge\geqslant}\},$$

where the dominance class $[x]_A^{\wedge\geqslant}$ describes objects that may dominate x in terms of A in a "$\bigwedge$" set-valued ordered information system.

A "$\bigwedge$" *set-valued ordered decision table* (ODT) is a "$\bigwedge$" set-valued ordered information system $S = (U, C \cup d, V, f)$, where d ($d \notin C$ and $f(x, d)(x \in U)$ is single-valued) is an overall preference called the decision, and all the elements of C are criterions, and $f : U \times C \to 2^V$ is a set-valued mapping. For example, Table 3 shows a conjunctive set-valued ordered decision table.

Let $S = (U, C \cup d, V, f)$ be a "$\bigwedge$" set-valued ODT, $A \subseteq C$, and $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ be the decision induced by d. The lower and upper approximations of $D_i^{\geqslant}$ $(i \leqslant r)$ with respect to the dominance relation $R_A^{\wedge\geqslant}$ are defined as [53]

$$\underline{R_A^{\wedge\geqslant}}(D_i^{\geqslant}) = \{x \in U | [x]_A^{\wedge\geqslant} \subseteq D_i^{\geqslant}\}, \quad \overline{R_A^{\wedge\geqslant}}(D_i^{\geqslant}) = \bigcup_{x \in D_i^{\geqslant}} [x]_A^{\wedge\geqslant}.$$

4.1.5. Disjunctive set-valued ordered decision tables

For a "$\bigvee$" set-valued information system $S = (U, AT, V, f)$, the relationships among any set $f(x, a)$, $x \in U$, $a \in AT$ are disjunctive. For convenience, let $R_A^{\vee\geqslant}$, $A \subseteq AT$, denote a dominance relation between objects that are possibly dominant in terms of values of attributes set A. Under this consideration, we call S a "$\bigvee$" *set-valued ordered information system*. Let us define the dominance relation more precisely as follows:

$$R_A^{\vee\geqslant} = \{(y, x) \in U \times U | \forall a \in A, \exists u_y \in f(y, a), \exists v_x \in f(x, a) \text{ such that } u_y \geqslant v_x\}.$$

By the definition of the dominance relation $R_A^{\vee\geqslant}$, it can be observed that if a pair of objects (y, x) from $U \times U$ lies in $R_A^{\vee\geqslant}$, then they are perceived as y dominates x; in other words, y may have a better property than x with respect to A in reality. In fact, this dominance relation is equivalent to the representation below

$$R_A^{\vee\geqslant} = \{(y, x) \in U \times U | \forall a \in A, \max f(y, a) \geqslant \min f(x, a)\}.$$

A "$\bigvee$" *set-valued ordered decision table* (ODT) is a "$\bigvee$" set-valued ordered information system $S = (U, C \cup d, V, f)$, where d ($d \notin C$ and $f(x, d)(x \in U)$ is single-valued) is an overall preference called the decision, and all the elements of C are criterions, and $f : U \times C \to 2^V$ is a set-valued mapping. A disjunctive set-valued is shown in Table 4.

Let $S = (U, C \cup d, V, f)$ be a "$\bigvee$" set-valued ODT, $A \subseteq C$, $\mathbf{D} = \{D_1, D_2, \ldots, D_r\}$ is the decision induced by d, the lower and upper approximations of $D_i^{\geqslant}$ $(i \leqslant r)$ with respect to the dominance relation $R_A^{\vee\geqslant}$ are defined as [53]

Table 2
An interval ordered decision table.

U	a_1	a_2	a_3	a_4	a_5	d
x_1	1	[0, 1]	2	1	[1, 2]	1
x_2	[0, 1]	0	[1, 2]	0	1	1
x_3	[0, 1]	0	[1, 2]	1	1	1
x_4	0	0	1	0	1	1
x_5	2	[1, 2]	3	[1, 2]	[2, 3]	2
x_6	[0, 2]	[1, 2]	[1, 3]	[1, 2]	[2, 3]	1
x_7	1	1	2	1	2	2
x_8	[1, 2]	[1, 2]	[2, 3]	2	[2, 3]	2
x_9	[1, 2]	2	[2, 3]	[0, 2]	3	2
x_{10}	2	2	3	[0, 1]	3	2

Table 3
A "$\bigwedge$" set-valued ordered decision table about language ability.

U	Audition	Spoken language	Reading	Writing	d
x_1	$\{E\}$	$\{E\}$	$\{F, G\}$	$\{F, G\}$	Poor
x_2	$\{E, F, G\}$	$\{E, F, G\}$	$\{F, G\}$	$\{E, F, G\}$	Good
x_3	$\{E, G\}$	$\{E, F\}$	$\{F, G\}$	$\{F, G\}$	Good
x_4	$\{E, F\}$	$\{E, G\}$	$\{F, G\}$	$\{F\}$	Poor
x_5	$\{F, G\}$	$\{F, G\}$	$\{F, G\}$	$\{F\}$	Poor
x_6	$\{F\}$	$\{F\}$	$\{E, F\}$	$\{E, F\}$	Poor
x_7	$\{E, F, G\}$	$\{E, F, G\}$	$\{E, G\}$	$\{E, F, G\}$	Good
x_8	$\{E, F\}$	$\{F, G\}$	$\{E, F, G\}$	$\{E, G\}$	Good
x_9	$\{F, G\}$	$\{G\}$	$\{F, G\}$	$\{F, G\}$	Poor
x_{10}	$\{E, F\}$	$\{E, G\}$	$\{F, G\}$	$\{E, F\}$	Good

Table 4
A "$\bigvee$" set-valued information system.

U	a_1	a_2	a_3	a_4	a_5	d
x_1	{1}	{0, 1}	{0}	{1, 2}	{2}	2
x_2	{0, 1}	{2}	{1, 2}	{0}	{0}	1
x_3	{0}	{1, 2}	{1}	{0, 1}	{0}	1
x_4	{0}	{1}	{1}	{1}	{0, 2}	1
x_5	{2}	{1}	{0, 1}	{0}	{1}	2
x_6	{0, 2}	{1}	{0, 1}	{0}	{1}	1
x_7	{1}	{0, 2}	{0, 1}	{1}	{2}	2
x_8	{0}	{2}	{1}	{0}	{0, 1}	1
x_9	{1}	{0, 1}	{0, 2}	{1}	{2}	2
x_{10}	{1}	{1}	{2}	{0, 1}	{2}	2

$$\underline{R_A^{\vee\geqslant}}(D_i^{\geqslant}) = \{x \in U \mid [x]_A^{\vee\geqslant} \subseteq D_i^{\geqslant}\}, \quad \overline{R_A^{\vee\geqslant}}(D_i^{\geqslant}) = \bigcup_{x \in D_i^{\geqslant}} [x]_A^{\vee\geqslant}.$$

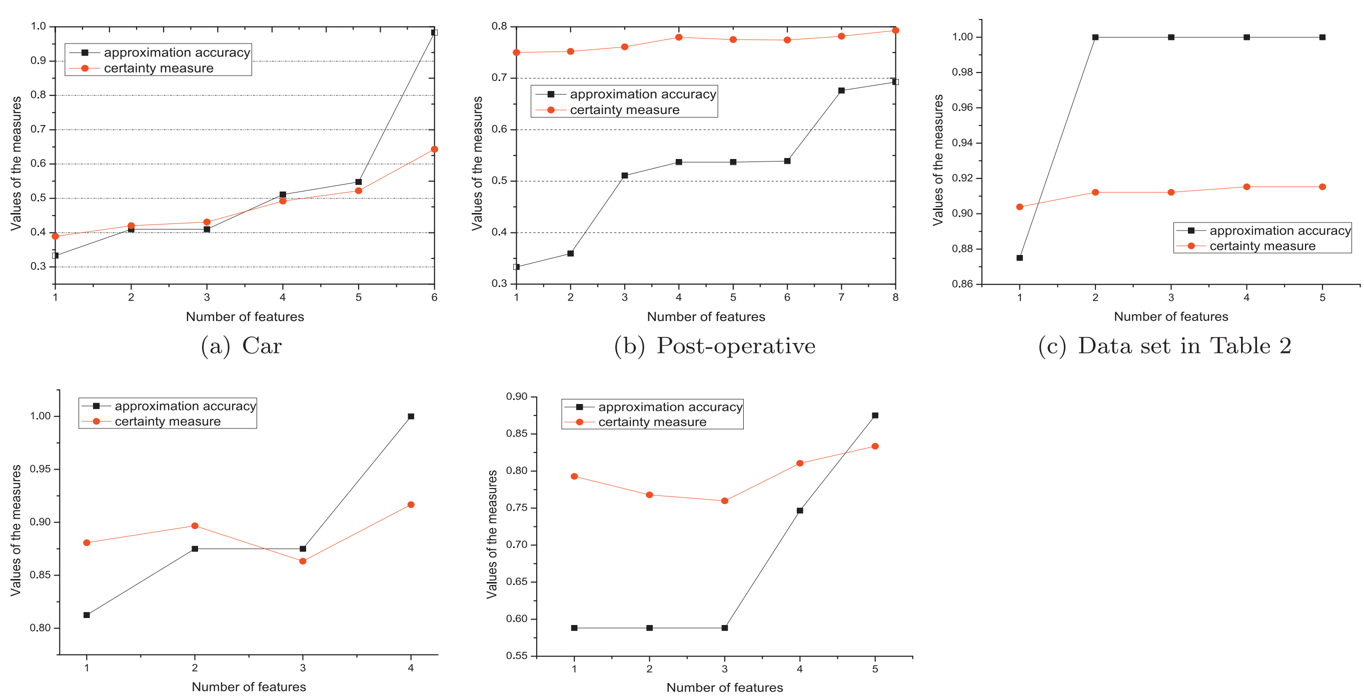
4.2. Experimental analysis

In order to verify the effectiveness of the measure α over the extended measure $a_C(\mathbf{D})$, we first compare the certainty measure α with the measure $a_C(\mathbf{D})$ through the evaluation of the certainty of each of five types of ordered decision tables. For this task, we have downloaded the public data sets Car (a single-valued ordered decision table) and Post-operative (an incomplete ordered decision table) from UCI Repository of machine learning databases [76], and have employed Table 2 (an interval ordered decision table), Table 3 (a conjunctive set-valued ordered decision table) and Table 4 (a disjunctive set-valued ordered decision table). In the data set Car, there are six condition attributes and one decision attribute. Their orders within the value sets of attributes are $low \rightarrow mid \rightarrow high \rightarrow v-high$ (buying), $low \rightarrow mid \rightarrow high \rightarrow v-high$ (maint), $5-more \rightarrow 4 \rightarrow 3 \rightarrow 2$ (doors), $more \rightarrow 4 \rightarrow 2$ (persons), $big \rightarrow mid \rightarrow small$ (lug boot), $high \rightarrow mid \rightarrow low$ (safety), and $v-good \rightarrow good \rightarrow acc \rightarrow unacc$ (decision attribute). In the data set Post-operative, there are eight condition attributes and one decision attribute. Their orders within the value sets of attributes are $low \rightarrow mid \rightarrow high$ (L-CORE), $low \rightarrow mid \rightarrow high$ (L-SURF), $excellent \rightarrow good \rightarrow fair \rightarrow poor$ (L-O2), $low \rightarrow mid \rightarrow high$ (L-BP), $stable \rightarrow mod-stable \rightarrow unstable$ (SURF-STBL), $stable \rightarrow mod-stable \rightarrow unstable$ (CORE-STBL), $stable \rightarrow mod-stable \rightarrow unstable$ (BP-STBL), $20 \rightarrow 19 \rightarrow 18 \cdots \rightarrow 1 \rightarrow 0$ (COMFORT), and $S \rightarrow A \rightarrow I$ (decision attribute). The comparisons of values of two measures with the numbers of features are shown in Figs. 1–5.

It can be seen from sub-figure (a) in Fig. 1 that the values of the extended approximation accuracy are unchanged when the number of features falls in between 2 and 3. In this situation, one lower/upper approximation of the target decision is the same as another lower/upper approximation of the target decision in the single-valued ordered decision table. But, for the same situation, as the number of features varies from 2 to 3, the value of the certainty measure α changes from 0.420 to 0.431. By adding a new attribute to existing attributes, the condition classes may become much finer, which can induce more ordered decision rules with bigger certainty accordingly. The proposed certainty measure α does characterize the character of ordered decision rules, while the extended approximation accuracy is not competent for the objective. From other sub-figures, one can see the same situation. Thus, the measure α is much better than the extended approximation accuracy for the single-valued ordered decision table. In other words, when the value of $a_C(D)$ is kept unchanged, the measure α may be still valid for evaluating the certainty of the set of decision rules obtained by using these selected features. Therefore, the measure α may be better than the extended approximation accuracy for evaluating the certainty of a single-valued ordered decision table.

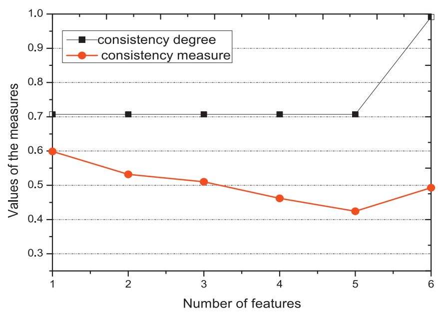
Now, we show the effectiveness of the measure β proposed in this paper and compare the consistency measure β with the measure $c_C(\mathbf{D})$ through evaluation of the consistency of each of the five types of ordered decision tables. Comparisons of values of two measures with the numbers of features are shown in Fig. 2.

From sub-figure (a) in Fig. 2, it is easy to see that the values of the consistency degree equal 0.707 when the number of features falls in between 1 and 5. In this situation, the lower approxima-

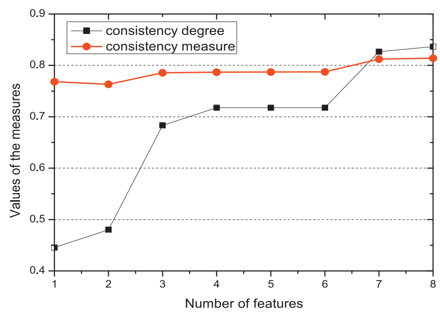


(a) Car (b) Post-operative (c) Data set in Table 2
(d) Data set in Table 3 (e) Data set in Table 4

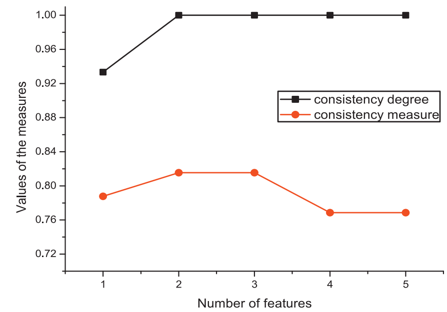
Fig. 1. Variation of certainty measure α and the approximation accuracy with the number of features.



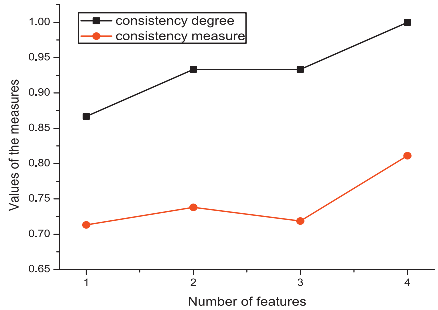
(a) Car



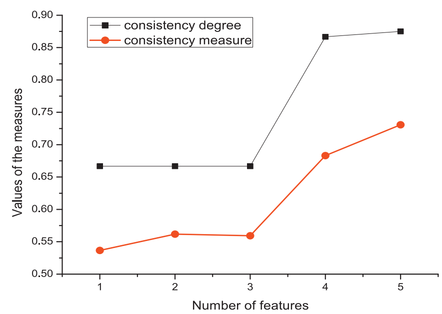
(b) Post-operative



(c) Data set in Table 2

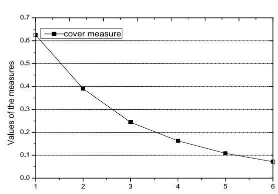


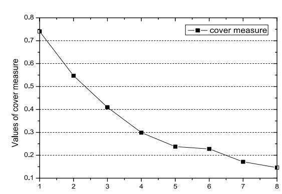
(d) Data set in Table 3

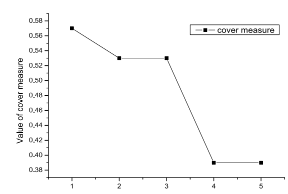


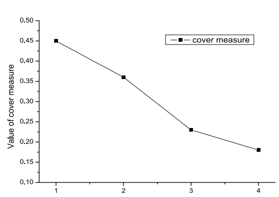
(e) Data set in Table 4

Fig. 2. Variation of consistency measure β and the quality of approximation with the number of features.


(a) AC on Sonar


(b) PR on Sonar


(c) RE on Sonar


(d) AC on Wine

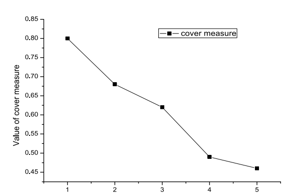

(e) PR on Wine

Fig. 3. Variation of the covering measure ϑ with the number of features.

tions of the target decision retain in the single-valued ordered decision table Car. However, through adding new features, those condition classes in lower approximations may gradually become much smaller, which will change the entire consistency of ordered decision rules. Because the extension of consistency degree only depends on lower approximations, it hence cannot be used to effectively characterize the consistency of the single-valued ordered decision table when the value of the consistency degree is invariable. However, for the same situation, as the number of features varies from 1 to 5, the value of the consistency measure β changes within the interval [0.424, 0.599]. It shows that unlike the extended consistency degree, the consistency measure β is still valid for evaluating the consistency of the single-valued ordered decision table when the lower approximation of the target decision keeps unchanged. Sub-figures (b)–(e) support the same conclusion. Therefore, the measure β is much better than the extended consistency degree for evaluating the decision performance based on the idea of reading the ordered decision table a set of ordered decision rules.

Finally, we investigate the variation of the values of the covering measure G with the numbers of features in ordered decision tables. The values of the measure with the number of features in ordered decision tables are shown in Fig. 3.

From Fig. 3, one can see that the value of the covering measure G decreases with the number of condition features becoming bigger in the same data set. Note that one may extract more decision rules through adding the number of condition features in general. In fact, the greater the number of decision rules, the smaller the value of the covering measure in the same data set. Therefore, the measure G is able to effectively evaluate the covering degree of all dominance classes in a given ordered decision table.

5. Conclusions and discussion

In rough set theory, several classical measures for evaluating a decision rule or a decision table, such as the certainty, support and coverage measures of a decision rule and the approximation accuracy and consistency degree (quality of approximation) of a decision table, can be extended for evaluating the decision performance of a decision rule (set) extracted from an ordered decision table. However, these extensions are not effective for evaluating the decision performance of a set of ordered decision rules. In this paper, the limitations of these extensions have been analyzed on ordered decision tables. To overcome these limitations, three new and more effective measures (α, β and G) have been introduced for evaluating the certainty, consistency and covering of a decision-rule set extracted from an ordered decision table, respectively. It has been analyzed how each of these three new measures depends on the condition granulation and decision granulation of ordered decision tables.

In order to apply the three new measures for evaluating the decision performance of a decision-rule set in practical decision problems, the experimental analysis on five types of ordered decision tables have been performed, which are single-valued ordered decision tables, incomplete ordered decision tables, interval ordered decision tables, conjunctive set-valued ordered decision tables and disjunctive set-valued ordered decision tables. Experimental results show that the three new measures (α, β, G) are adequate for evaluating the decision performance of a decision-rule set extracted from any type of ordered decision tables. The three measures may be helpful for determining which of rule extracting approaches is preferred for a practical decision problem in the context of ordered decision tables.

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments on this study. This work was supported by National Natural Science Fund of China(Nos. 60903110 and 71031006), MOE Project of Humanities and Social Sciences (No. 12YJC630174), Natural Science Fund of Shanxi Province, China (Nos. 2009021017-1 and 2008011038), and Shanxi Scholarship Council of China (No. 201008).

References

[1] J. Bazan, J.F. Peters, A. Skowron, H.S. Nguyen, M. Szczuka, Rough set approach to pattern extraction from classifiers, Electronic Notes in Theoretical Computer Science 82 (4) (2003) 1–10.
[2] T. Beaubouef, Rules in incomplete information systems, Information Sciences 113 (1999) 271–292.
[3] T. Beaubouef, F.E. Perty, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1998) 185–195.
[4] M. Beynon, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research 134 (3) (2001) 592–605.
[5] J. Blaszczynski, S. Greco, R. Slowinski, M. Szelag, Monotonic variable consistency rough set approaches, International Journal of Approximate Reasoning 50 (2009) 979–999.
[6] K. Dembczynski, R. Pindur, R. Susmaga, Generation of exhaustive set of rules within dominance-based rough set approach, Electronic Notes in Theoretical Computer Science 82 (4) (2003).
[7] K. Dembczynski, R. Pindur, R. Susmaga, Dominance-based rough set classifier without induction of decision rules, Electronic Notes in Theoretical Computer Science 82 (4) (2003).
[8] K. Dembczynski, S. Greco, W. Kotlowski, R. Slowinski, Quality of rough approximation in multi-criteria classification problems, Lecture Notes in Computer Science 4259 (2006) 318–327.
[9] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1998) 109–137.
[10] I. Düntsch, G. Gediga, Approximation quality for sorting rules, Computational Statistics & Data Analysis 403 (2002) 499–526.
[11] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (2001) 1–47.
[12] S. Greco, B. Matarazzo, R. Slowinski, Rough sets methodology for sorting problems in presence of multiple attributes and criteria, European Journal of Operational Research 138 (2002) 247–259.
[13] S. Greco, Z. Pawlak, R. Slowinski, Can Bayesian confirmation measures be useful for rough set decision rules?, Engineering Applications of Artificial Intelligence 17 (2004) 345–361
[14] Y.Y. Guan, H.K. Wang, Set-valued information systems, Information Science 176 (2006) 2507–2525.
[15] J.W. Grzymala-Busse, Characteristic relations for incomplete data: a generalization of the indiscernibility relation, Transactions on Rough Sets 4 (2005) 58–68.
[16] J.W. Grzymala-Busse, W. Grzymala-Busse, An experimental comparison of three rough set approaches to missing attribute value, Transactions on Rough Sets 6 (2007) 31–50.
[17] J. Huysmans, B. Baesens, J. Vanthienen, A new approach for measuring rule set consistency, Data and Knowledge Engineering 63 (1) (2007) 167–182.
[18] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: a tutorial, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hubridization: A New Trend in Decision Making, Springer, Singapore, 1999, pp. 3–98.
[19] I. Kononenko, I. Bratko, E. Roskar, Experiments in automatic learning of medical diagnostic rules, Techinical Report, Jozef Stefan Institute, Ljubljana, Yugoslavia, 1984.
[20] M. Kryszkiewicz, Rough set approach to incomplete information systems, in: Proceedings of Second Annual Joint Conference on Information Sciences, Fuzzy Logic, Neural Computing, Pattern Recognition, Computer Vision, Evolutionary Computing, Information Theory, Computational Intelligence, 1995, pp. 194–197.
[21] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1998) 39–49.
[22] M. Kryszkiewicz, Rule in incomplete information systems, Information Sciences 113 (1999) 271–292.
[23] M. Kryszkiewicz, Comparative study of alternative type of knowledge reduction in inconsistent systems, International Journal of Intelligent Systems 16 (2001) 105–120.
[24] R. Latkowski, On decomposition for incomplete data, Fundamenta Informaticae 54 (1) (2003) 1–16.
[25] R. Latkowski, M. Mikolajczyk, Data decomposition and decision rule joining for classification of data with missing values, Transactions on Rough Sets 3 (2004) 299–320.
[26] R. Latkowski, Flexible indiscernibility relations for missing attribute values, Fundamenta Informaticae 67 (1) (2005) 131–147.
[27] Y. Leung, D.Y. Li, Maximal consistent block technique for rule acquisition in incomplete information systems, Information Sciences 153 (2003) 85–106.
[28] D.Y. Li, B. Zhang, Y. Leung, On knowledge reduction in inconsistent decision information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (5) (2004) 651–672.
[29] J.Y. Liang, D.Y. Li, Uncertainty and Knowledge Acquisition in Information Systems, Science Press, Beijing, China, 2005.
[30] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Science in China (Series F) 51 (9) (2008) 1–18.
[31] J.Y. Liang, K.S. Qu, Information measures of roughness of knowledge and rough sets in incomplete information systems, Journal of System Science and System Engineering 24 (5) (2001) 544–547.
[32] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (1) (2004) 37–46.
[33] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information system, International Journal of General Systems 35 (6) (2006) 641–654.
[34] J.Y. Liang, Z.B. Xu, The algorithm on knowledge reduction in incomplete information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 24 (1) (2002) 95–103.
[35] E. Marczewski, A general scheme of independence in mathematics, Bulletin de Academie Polonaise des Sciences Serie des Sciences Mathematiques Astronomiques et Physique 6 (1958) 331–362.
[36] J.S. Mi, W.Z. Wu, W.X. Zhang, Comparative studies of knowledge reductions in inconsistent systems, Fuzzy Systems and Mathematics 17 (3) (2003) 54–60.
[37] J.S. Mi, W.Z. Wu, W.X. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (2004) 255–272.
[38] H.S. Nguyen, D. Slezak, Approximation reducts and association rules correspondence and complexity results, Proceeding of RSFDGrC′ 99 (1999) 137–145.
[39] S.K. Pal, P. Mitra, Case generation using rough sets with fuzzy representation, IEEE Transactions on Knowledge and Data Engineering 16 (3) (2004) 292–300.
[40] S.K. Pal, W. Pedrycz, A. Skowron, R. Swiniarski, Presenting the special issue on rough-neuro computing, Neurocomputing 36 (2001) 1–3.
[41] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (1982) 341–356.
[42] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.

[43] Z. Pawlak, Rough set theory and its applications in data analysis, Cybernetics and Systems 29 (1998) 661–688.
[44] Z. Pawlak, Some remarks on conflict analysis, Information Sciences 166 (2005) 649–654.
[45] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
[46] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (2007) 28–40.
[47] Z. Pawlak, A. Skowron, Rough sets and Boolean reasoning, Information Sciences 177 (2007) 41–73.
[48] Z. Pawlak, J.W. Grzymala-Busse, R. Slowiski, W. Ziarko, Rough sets, Communications on ACM 38 (11) (1995) 89–95.
[49] Y.H. Qian, C.Y. Dang, J.Y. Liang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets and Systems 159 (2008) 2353–2377.
[50] Y.H. Qian, J.Y. Liang, C.Y. Dang, Incomplete multigranulation rough set, IEEE Transactions on Systems, Man and Cybernetics – Part A 40 (2) (2010) 420–431.
[51] Y.H. Qian, C.Y. Dang, J.Y. Liang, H.Y. Zhang, J.M. Ma, On the evaluation of the decision performance of an incomplete decision table, Data & Knowledge Engineering 65 (3) (2008) 373–400.
[52] Y.H. Qian, J.Y. Liang, Combination entropy and combination granulation in rough set theory, International, Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16 (2) (2008) 179–193.
[53] Y.H. Qian, J.Y. Liang, C.Y. Dang, Set-valued ordered information systems, Information Sciences 179 (2009) 2809–2832.
[54] Y.H. Qian, J.Y. Liang, C.Y. Dang, Interval ordered information systems, Computers and Mathematics with Applications 56 (2008) 1994–2009.
[55] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: a multigranulation rough set, Information Sciences 180 (2010) 949–970.
[56] Y.H. Qian, J.Y. Liang, Witold Pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174 (2010) 597–618.
[57] Y.H. Qian, J.Y. Liang, D.Y. Li, H.Y. Zhang, C.Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences 178 (2008) 181–202.
[58] Y.H. Qian, J.Y. Liang, C.Y. Dang, Knowledge structure, knowledge granulation and knowledge distance in a knowledge base, International Journal of Approximate Reasoning 50 (2009) 174–188.
[59] M. Quafatou, α-RST: a generalization of rough set theory, Information Sciences 124 (2000) 301–316.
[60] Y. Sai, Y.Y. Yao, N. Zhong, Data analysis and mining in ordered information tables, in: Proceedings of 2001 IEEE International Conference on Data Mining, 2001, pp. 497–504.
[61] M.W. Shao, W.X. Zhang, Dominance relation and rules in an incomplete ordered information system, International Journal of Intelligent Systems 20 (2005) 13–27.
[62] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: R. Slowinski (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic, Dordrecht, 1992, pp. 331–362.
[63] D. Slezak, Approximate reducts in decision tables, Proceedings of IPMU′ 96 (1996) 1159–1164.
[64] R. Slowinski, S. Greco, B. Matarazzo, Rough sets in decision making, in: R.A. Meyers (Ed.), Encyclopedia of Complexity and Systems Science, Springer, New York, 2009, pp. 7753–7786.
[65] R. Slowinski, D. Vanderpooten, A generalized definition of rough approximations based on similarity, IEEE Transactions on Knowledge and Data Engineering 12 (2) (2000) 331–336.
[66] R. Slowinski, J. Stefanowski, Rough-set reasoning about uncertain data, Foundamenta Informaticae 27 (2–3) (1996) 229–244.
[67] W.Z. Wu, M. Zhang, H.Z. Li, J.S. Mi, Knowledge reduction in random information systems via Dempster–Shafer theory of evidence, Information Sciences 174 (2005) 143–164.
[68] Z. Xiao, X.L. Yang, Y. Pang, X. Dang, The prediction for listed companies financial distress by using multiple prediction methods with rough set and Dempster–Shafer evidence theory, Knowledge-Based Systems 26 (2012) 196–206.
[69] W.H. Xu, Y. Li, X.W. Liao, Approaches to attribute reductions based on rough set and matrix computation in inconsistent ordered information systems, Knowledge-Based Systems 27 (2012) 78–91.
[70] X.B. Yang, M. Zhang, H.L. Dou, J.Y. Yang, Neighborhood systems-based rough sets in incomplete information system, Knowledge-Based Systems 24 (2011) 858–867.
[71] H.L. Yang, S.G. Li, Z.L. Guo, C.H. Ma, Transformation of bipolar fuzzy rough set models, Knowledge-Based Systems 27 (2012) 60–68.
[72] Y.Y. Yao, A comparative study of fuzzy sets and rough sets, Information Sciences 109 (1998) 227–242.
[73] Y.Y. Yao, Neighborhood systems and approximate retrieval, Information Sciences 174 (2007) 143–164.
[74] W. Ziarko, Variable precision rough set model, Journal of Computer System Science 46 (1993) 39–59.
[75] W.X. Zhang, W.Z. Wu, J.Y. Liang, D.Y. Li, Theory and Method of Rough Sets, Science Press, Beijing, China, 2001.
[76] The UCI Machine Learning Respository. <http://mlearn.ics.uci.edu/MLRepository.html>.