Sequential Three-way Decisions with Probabilistic Rough Sets

Yiyu Yao* and Xiaofei Deng
Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2
{yyao,deng200x}@cs.uregina.ca

Abstract

When approximating a concept, probabilistic rough set models use probabilistic positive, boundary and negative regions. Rules obtained from the three regions are recently interpreted as making three-way decisions, consisting of acceptance, deferment, and rejection. A particular decision is made by minimizing the cost of correct and incorrect classifications. This framework is further extended into sequential three-way decision-making, in which the cost of obtaining required evidence or information is also considered.

1. Introduction

Three-way decisions, consisting of acceptance, rejection and deferment (or further investigation), are a common problem solving strategy used in many disciplines. They have been used in editorial peer review process [15], medical decision-making [8, 9], statistical inference, statistical decision-making and applications [2, 3, 16], sequential hypothesis testing [14], information retrieval and intelligent agents [4, 5], email spam filtering [29, 30], investment management [7], cluster analysis and data analysis [6, 13], and others. One may broadly classify existing studies on three-way decision-making into two classes, namely, a single one-step three-way decision-making and a sequential, multi-step three-way decision-making. The former may be viewed as one step or a special case of the latter.

There are different motivations and justifications for making three-way decisions with an added option of indecision. Recently, an interpretation of three-way decision making is proposed within the framework of decision-theoretic rough set models [20, 21, 22]. It is based on the Bayesian decision theory that minimizing the risk of various decisions. The result is similar to hypothesis testing in statistics. A hypothesis is accepted if there is a strong evidence supporting it, is rejected if there is a strong evidence refuting it, and is neither accepted nor rejected, but needs to be further tested, if there is no strong evidence supporting or refuting it. The interpretation justifies three-way decision-making based on the risk or cost of different decisions. It lacks a consideration of the cost of obtaining and using evidence.

Decision-making is normally based on available evidence. For example, a piece of evidence may be the result of a test or an observation. Obtaining different pieces of evidence is associated with different costs. A cost-effective decision-making method aims at achieving a required level of accuracy with a minimal cost in obtaining evidence. This offers another interpretation of three-way decisions called sequential three-way decision-making. A decision-making process consists of a sequence of steps. In each step, one makes a decision of acceptance or rejection if the available evidence is sufficiently strong, otherwise, we move to the next step in which a rule of deferment in last step is refined into a set of rules of acceptance, rejection or deferment by adding new evidence. The process continues till a certain condition is met, for example, all possible tests are used or it is too costly to continue.

Sequential three-way decision-making implements an idea of granular computing called progressive computing [23]. Suppose that a problem is described with multiple levels of granularity, abstraction or details, i.e., a problem has multiple representations. Decision-making at a higher level is typically associated with lower costs than decision-making at a lower level. A decision of acceptance and rejection can be made with tolerable degrees of error at a higher level with less cost, and a decision of deferment may be further investigated at a lower level by observing more details.

The main idea of sequential three-way decision may be further illustrated by an example. Considering

*This work is partially supported by a discovery grant from NSERC Canada.

the problem of selecting a set of relevant papers from a set of papers. At the very first step or level, by reading only paper titles, we can accepting some papers as being relevant, reject some papers as being non-relevant, or defer a definite decision. Quick decisions are made based on the informativeness of the paper titles with much less cost of reading time. When a paper title is less informative, at the second step or level, we need to read section and subsection headings to determine the relevance of the paper, which requires more time. The results are again three-way decisions. At the third step or level, we need to read the introduction, conclusion and some paragraphs to determine the relevance of the paper. Even more time is required. As demonstrated by this example, sequential three-way decision making suggests that if a decision of acceptance or rejection with certain tolerable levels of errors can be made at a higher level, it is not necessary to move to a lower level.

The intuitive notion of sequential three-way decision has been explored and used in many areas. The main objective of this paper is to provide a formal description of this method within the framework of probabilistic rough sets.

2. Multiple Representations of Objects in an Information Table

In rough set analysis [10], one uses an information table to describe a finite universe of objects U by a finite set of attributes At. Formally, an information table is the following tuple:

$$S = (U, At, \{V_a \mid a \in At\}, \{I_a \mid a \in At\}), \qquad (1)$$

where

U is a finite nonempty set of objects,
At is a finite nonempty set of attributes,
V_a is a nonempty set of values for an attribute $a \in At$,
$I_a : U \longrightarrow V_a$ is a complete information function that maps an object of U to exactly one value in V_a.

That is, an object is determined only by its values on a set of attributes.

In an information table, a subset of attributes $A \subseteq At$ defines an equivalence relation $E_A \subseteq U \times U$ on U:

$$xE_Ay \Longleftrightarrow \forall a \in A(I_a(x) = I_a(y)), \qquad (2)$$

where $I_a(x)$ is the value of x on attribute a. That is, x and y are equivalent if and only if they have the same values on all attributes in A. The equivalence relation E_A induces a partition of U, $U/E_A = \{[x]_A \mid x \in U\}$, where $[x]_A = [x]_{E_A} = \{y \mid xE_Ay\}$ is the equivalence class containing x.

The equivalence class $[x]_A$ may be interpreted as a granule and the partition U/E_A as a granulation of U. One immediately obtains a partition based framework of granular computing using an information table [18].

Let $\mathrm{Des}_A(x)$ denote the description of x with respect to a subset of attributes $A \subseteq At$, namely, $\mathrm{Des}_A(x) = \bigwedge_{a\in A}(a = I_a(x))$. A condition (*attribute_name* = *attribute_value*) is called an atomic formula. An object is therefore equivalently described by a conjunction of atomic formulas with respect to the set of attributes A. All objects in the equivalence class $[x]_A$ have the same description as x and $\mathrm{Des}_A(x)$ is in fact the description of $[x]_A$.

Let $\preceq$ denote the refinement-coarseing relation between partitions, namely, $\pi_2 \preceq \pi_1$ holds between two partitions if and only if every block of π_1 is the union of some blocks of π_2. Consider two sets of attributes $A_1 \subset A_2 \subseteq At$, we have $[x]_{A_2} \subseteq [x]_{A_1}$ and $U/A_2 \preceq U/A_1$. That is, a larger set of attributes produces a finer partition of the universe and a smaller set of attributes produces a coarser partition. An object x is described by two different representations, namely, $\mathrm{Des}_{A_1}(x)$ and $\mathrm{Des}_{A_2}(x)$. For a sequence of sets of attributes

$$A_1 \subset A_2 \subset \ldots \subset A_k \subseteq At, \qquad (3)$$

we have

$$[x]_{A_k} \subseteq \ldots \subseteq [x]_{A_2} \subseteq [x]_{A_1}, \qquad (4)$$

and

$$U/A_k \preceq \ldots U/A_2 \preceq U/A_1. \qquad (5)$$

The sequence of partitions provides a multilevel granular structures of the universe. With respect to a sequence of sets of attributes, we have a sequence of different descriptions of an object,

$$\mathrm{Des}_{A_1}(x),\ \mathrm{Des}_{A_2}(x),\ \ldots,\ \mathrm{Des}_{A_k}(x), \qquad (6)$$

with increasing levels of details. Namely, more atomic formulas are added in the sequence of descriptions.

In general, we can consider a meet semi-lattice $(\{U/A \mid A \in 2^{At}\}, \sqcap)$ defined by the power set 2^{At} of the set of attributes At, where the meet operator $\sqcap$ is defined by $U/A \sqcap U/B = U/(A \cup B)$ for $A, B \in 2^{At}$, namely, $[x]_A \cap [x]_B = [x]_{A\cup B}$. The meet semi-lattice gives a more general multilevel granular structure of the universe.

Based on these concepts and notations, three-way decisions can be formulated with respect to a set of attributes and a sequence of nested sets of attributes, respectively.

3. Three-way Decision Making with a Set of Attributes

Decision-theoretic rough set models (DTRSM) [17, 19, 24, 25] are quantitative probabilistic extensions of the qualitative classical rough set model [10]. A pair of thresholds on conditional probability is used to obtain probabilistic positive, boundary and negative regions.

3.1. Three-way decision rules

Let C denote a concept, or be the name of the concept, whose instances are the subset of objects $C \subseteq U$. Suppose that $Pr(C|[x]_A)$ denotes the conditional probability of an object in C given that the object is in $[x]_A$. It may be roughly estimated based on the cardinality of set as follows:

$$Pr(C|[x]_A) = \frac{|C \cap [x]_A|}{|[x]_A|}, \tag{7}$$

where $|\cdot|$ denotes the cardinality of a set. With conditional probability, decision-theoretic rough models introduce probabilistic positive, boundary and negative regions: for $0 \leq \beta < \alpha \leq 1$,

$$\begin{aligned} \mathrm{POS}_{(\alpha,\beta)}(C) &= \{x \in U \mid Pr(C|[x]_A) \geq \alpha\}, \\ \mathrm{BND}_{(\alpha,\beta)}(C) &= \{x \in U \mid \beta < Pr(C|[x]_A) < \alpha\}, \\ \mathrm{NEG}_{(\alpha,\beta)}(C) &= \{x \in U \mid Pr(C|[x]_A) \leq \beta\}. \end{aligned} \tag{8}$$

Pawlak model may be viewed as a special of probabilistic models in which $\beta = 0$ and $\alpha = 1$.

From the three probabilistic regions, we can obtain three types of quantitative probabilistic rules, i.e., rules of acceptance, rules of deferment and rules of rejection:

$$\begin{aligned} &\text{rule of acceptance} : [x]_A \subseteq \mathrm{POS}_{(\alpha,\beta)}(C), \\ &\quad \mathrm{Des}_A(x) \rightarrow \text{accept } x \in C; \\ &\text{rule of deferment} : [x]_A \subseteq \mathrm{BND}_{(\alpha,\beta)}(C), \\ &\quad \mathrm{Des}_A(x) \rightarrow \text{neither accept nor reject } x \in C; \\ &\text{rule of rejection} : [x]_A \subseteq \mathrm{NEG}_{(\alpha,\beta)}(C), \\ &\quad \mathrm{Des}_A(x) \rightarrow \text{reject } x \in C. \end{aligned}$$

Probabilistic three-way decisions are of a quantitative nature. By definition of probabilistic regions, decisions of acceptance and rejection are associated with some errors. Three-way decisions may be related to statistical inferences and hypothesis testing. A hypothesis is given as $x \in C$ and a piece of evidence is given as $\mathrm{Des}_A(x)$, that is, based on the description of x we want to infer its membership in C. If the evidence strongly supports the hypothesis (i.e., $Pr(C|[x]_A) \geq \alpha$), we accept $x \in C$. If the evidence weakly supports the hypothesis (i.e., $Pr(C|[x]_A) \leq \beta$), or strongly refutes the hypothesis (i.e., $Pr(C^c|[x]_A) \geq 1-\beta$), we reject $x \in C$. If the evidence is not sufficiently strong, we defer such a definite acceptance or rejection decision.

In the Pawlak model, $x \in C$ is true for the positive region, false for the negative region, and neither true nor false for the boundary region. The same interpretation is no longer valid for probabilistic regions. Thus, the naming of three types of rule reflects the quantitative characteristics of probabilistic regions and is more meaningful.

3.2. Computation of thresholds

For practical applications of probabilistic rough sets, it is necessary to interpret the meaning of the pair of thresholds, (α, β), and to provide a method for estimating or computing them. The formulation of decision-theoretic rough set models in fact solves these problems based on the Bayesian decision theory [1]; a particular decision of choosing the positive region, boundary region or negative region is made with minimum risk.

With respect to a concept $C \subseteq U$, we have a set of two states $\Omega = \{C, C^c\}$ indicating that an object is in C and not in C, respectively. Corresponding to three-way decisions, we have a set of three actions $\mathscr{A} = \{a_P, a_B, a_N\}$, where a_P, a_B, and a_N represent the three actions in classifying an object x, namely, deciding $x \in \mathrm{POS}(C)$, deciding $x \in \mathrm{BND}(C)$, and deciding $x \in \mathrm{NEG}(C)$, respectively. The loss function regarding the risk or cost of actions in different states is given by a 3×2 matrix:

	C (P)	C^c (N)
a_P (P)	λ_{PP}	λ_{PN}
a_B (B)	λ_{BP}	λ_{BN}
a_N (N)	λ_{NP}	λ_{NN}

In the matrix, P, B and N inside the parentheses are used as subscripts to label various loss. In particular, λ_{PP}, λ_{BP} and λ_{NP} denote the losses incurred for taking actions a_P, a_B and a_N, respectively, when an object belongs to C, and λ_{PN}, λ_{BN} and λ_{NN} denote the losses incurred for taking the same actions when the object does not belong to C.

Given the loss matrix, the expected losses associated with taking different actions for objects in $[x]_A$ can be expressed as:

$$\begin{aligned} R(a_P|[x]_A) &= \lambda_{PP} Pr(C|[x]_A) + \lambda_{PN} Pr(C^c|[x]_A), \\ R(a_B|[x]_A) &= \lambda_{BP} Pr(C|[x]_A) + \lambda_{BN} Pr(C^c|[x]_A), \\ R(a_N|[x]_A) &= \lambda_{NP} Pr(C|[x]_A) + \lambda_{NN} Pr(C^c|[x]_A). \end{aligned} \tag{9}$$

The Bayesian decision procedure suggests the following minimum-risk decision rules:

(P) If $R(a_P|[x]_A) \leq R(a_N|[x]_A)$ and $R(a_P|[x]_A) \leq R(a_B|[x]_A)$, then decide $x \in \text{POS}(C)$;

(B) If $R(a_B|[x]_A) \leq R(a_P|[x]_A)$ and $R(a_B|[x]_A) \leq R(a_N|[x]_A)$, then decide $x \in \text{BND}(C)$;

(N) If $R(a_N|[x]_A) \leq R(a_P|[x]_A)$ and $R(a_N|[x]_A) \leq R(a_B|[x]_A)$, then decide $x \in \text{NEG}(C)$.

Tie-breaking criteria are added so that each object is put into only one region. Since $Pr(C|[x]_A) + Pr(C^c|[x]_A) = 1$, we can simplify the rules based only on the probabilities $Pr(C|[x]_A)$ and the loss function λ.

Consider a special kind of loss functions with:

$$
\text{(c0)} \qquad \begin{aligned} &\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}, \\ &\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}. \end{aligned} \tag{10}
$$

That is, the loss of classifying an object x belonging to C into the positive region $\text{POS}(C)$ is less than or equal to the loss of classifying x into the boundary region $\text{BND}(C)$, and both of these losses are strictly less than the loss of classifying x into the negative region $\text{NEG}(C)$. The reverse order of losses is used for classifying an object not in C. Under condition (c0), the decision rules can be re-expressed as:

(P) If $Pr(C|[x]_A) \geq \alpha$ and $Pr(C|[x]_A) \geq \gamma$, then decide $x \in \text{POS}(C)$;

(B) If $Pr(C|[x]_A) \leq \alpha$ and $Pr(C|[x]_A) \geq \beta$, then decide $x \in \text{BND}(C)$;

(N) If $Pr(C|[x]_A) \leq \beta$ and $Pr(C|[x]_A) \leq \gamma$, then decide $x \in \text{NEG}(C)$;

where the parameters α, β, and γ are are computed from the loss function as:

$$
\begin{aligned}
\alpha &= \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}, \\
\beta &= \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}, \\
\gamma &= \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}.
\end{aligned} \tag{11}
$$

In other words, thresholds are practically interpret and systematically determined by the loss function.

Consider now an additional condition on the loss function:

$$
\text{(c1)} \qquad \begin{aligned} &(\lambda_{PN} - \lambda_{BN})(\lambda_{NP} - \lambda_{BP}) > \\ &(\lambda_{BN} - \lambda_{NN})(\lambda_{BP} - \lambda_{PP}). \end{aligned} \tag{12}
$$

Conditions (c0) and (c1) imply that $1 \geq \alpha > \gamma > \beta \geq 0$. After tie-breaking, the following simplified rules are obtained:

(P) If $Pr(C|[x]_A) \geq \alpha$, then decide $x \in \text{POS}(C)$;

(B) If $\beta < Pr(C|[x]_A) < \alpha$, then decide $x \in \text{BND}(C)$;

(N) If $Pr(C|[x]_A) \leq \beta$, then decide $x \in \text{NEG}(C)$.

The parameter γ is no longer needed. Thus, (α, β)-probabilistic positive, negative and boundary regions are given, respectively, by:

$$
\begin{aligned}
\text{POS}_{(\alpha,\beta)}(C) &= \{x \in U \mid Pr(C|[x]_A) \geq \alpha\}, \\
\text{BND}_{(\alpha,\beta)}(C) &= \{x \in U \mid \beta < Pr(C|[x]_A) < \alpha\}, \\
\text{NEG}_{(\alpha,\beta)}(C) &= \{x \in U \mid Pr(C|[x]_A) \leq \beta\}.
\end{aligned} \tag{13}
$$

The derivation shows that decision-theoretic rough set models have a solid theoretical basis and a systematical practical way of computing the required thresholds.

4. Sequential Three-Way Decision-making with a Sequence of Sets of Attributes

For simplicity and clarity, we assume that a sequence of nested sets of attributes, $A_1 \subset A_2 \subset \ldots \subset A_k \subseteq At$, is given so that we can concentrate on a sequential three-way decision-making process. It suffices to say that such a sequence may be obtained by adding one or more attributes each time. Attributes may be added based on their importance given by a user [27, 28] or the cost of obtaining their values. Furthermore, the first set A_1 may be the set of core attributes [10] and sequence can be obtained by an addition strategy used in reduct construction [26].

4.1. Non-monotonicity of sequential decision-making

Consider two sets of attributes $A_1 \subset A_2 \subseteq At$. Values on additional attributes in $A_2 - A_1$ may be viewed as new evidence. In other words, by using additional attributes we change a coarser description $\bigwedge_{a \in A_1}(a = I_a(x))$ of an object x into a finer description $(\bigwedge_{a \in A_1}(a = I_a(x))) \wedge (\bigwedge_{b \in A_2 - A_1}(b = I_b(x)))$. In terms of equivalence classes, we have $[x]_{A_2} \subseteq [x]_{A_1}$. On the other hand, such a monotonicity does not hold for conditional probability. One of the following three scenarios may happen:

$$
\begin{aligned}
&Pr(C|[x]_{A_2}) > Pr(C|[x]_{A_1}),\\
&Pr(C|[x]_{A_2}) = Pr(C|[x]_{A_1}),\\
&Pr(C|[x]_{A_2}) < Pr(C|[x]_{A_1}). \qquad (14)
\end{aligned}
$$

They suggest that the new evidence supports, is neutral, and refutes C, respectively. A decision of acceptance, deferment, and rejection made at a higher level may incorrect and need to be revised at a lower level.

The non-monotonicity brings a difficulty to sequential decision-making. It requires a reconsideration of decisions at different levels. On the other hand, we want to make a decision at a higher level without moving into a lower levels. Recall that three-way decision with probabilistic rough sets is based on an acceptance of some tolerable levels of classification errors. In a similar way, we need to accept incorrect decisions made in sequential three-way decision-making. At the same time, we need to keep chance of revision of decisions at different levels as small as possible.

In order to avoid decision revision, a user may be more conservative in expressing loss functions at higher levels, where a major probability revision is more likely. In term of the thresholds, they satisfy the following conditions:

$$
\begin{aligned}
&0 \le \beta_i < \alpha_i \le 1, \quad 1 \le i \le k,\\
&\beta_1 \le \beta_2 \le \ldots \le \beta_k < \alpha_k \le \ldots \alpha_2 \le \alpha_1. \qquad (15)
\end{aligned}
$$

It suggests that at higher levels a user is more biased towards a deferment decision so that an acceptance or a rejection decision can be made when more evidence and details is available at lower levels.

4.2. Sequential three-way decisions

In forming sequential three-way decision rules, we start with the first set of attribute A_1 in a sequence $A_1 \subset A_2 \subset \ldots \subset A_k \subseteq At$. The formulation of three-way decisions is basically discussed in Section 3. In the next step, we will keep rules of acceptance and rules of rejection, and only need to revise deferment rules in the previous step. In this step, only objects in the boundary region $\mathrm{BND}_{(\alpha_1,\beta_1)}(C)$ are considered. The same method is used in all subsequent steps.

The main ideas of sequential three-way decision rules construction can be described as follows:

Step 1: Let $U_1 = U$ and $C_1 = C$. Compute three probabilistic regions $\mathrm{POS}_{(\alpha_1,\beta_1)}(C_1)$, $\mathrm{BND}_{(\alpha_1,\beta_1)}(C_1)$, and $\mathrm{NEG}_{(\alpha_1,\beta_1)}(C_1)$, respectively. Construct rules of acceptance, rules of rejection, and rules of deferment based on the positive, negative and boundary regions, respectively.

Step i $(1 < i \le k)$**:** Let

$$
\begin{aligned}
U_i &= \mathrm{BND}_{(\alpha_{i-1},\beta_{i-1})}(C_{i-1}),\\
C_i &= C \cap \mathrm{BND}_{(\alpha_{i-1},\beta_{i-1})}(C_{i-1}). \qquad (16)
\end{aligned}
$$

Compute, respectively, three probabilistic regions $\mathrm{POS}_{(\alpha_i,\beta_i)}(C_i)$, $\mathrm{BND}_{(\alpha_i,\beta_i)}(C_i)$, and $\mathrm{NEG}_{(\alpha_i,\beta_i)}(C_i)$. Construct rules of acceptance, rules of rejection, and rules of deferment based on the positive, negative and boundary regions, respectively.

Step i $(i > 1)$ refines rules of deferment constructed from the boundary region of step $i-1$, in which U_i is considered as the new universe and C_i as the new concept. For the construction of probabilistic regions, we can simply use U and C, due to the use of a sequence of nested sets of attributes. However, we explicitly construct them to show and emphasize that in each step only a fraction of the original universe of objects is used.

In step i $(i > 1)$, a rule of deferment of step $i-1$ is replaced by a set of rules of acceptance, rejection or deferment, based on additional attributes in $A_i - A_{i-1}$. In sequential decision-making, it is typically the case that at step i the values of attributes in $A_i - A_{i-1}$ are not known. One must perform some test or observation to obtain these values. Thus, sequential decision-making has an advantage that one only needs to collect new evidence or perform new test whenever an acceptance or a rejection decision cannot be made based on available evidence or information. It is crucial that the sequence of subsets of attributes should be constructed such that attributes in $A_i - A_{i-1}$ are most informative.

The proposed method of constructing sequential three-way decision rules draws ideas from decision tree construction in machine learning [12] and a fast algorithm for computing rough set approximations [11]. Like a decision tree, a multilevel organization of rules immediately leads to a sequential three-way decision-making process.

Additional comments on the proposed method are given for its possible generalizations. We assume that a sequence of sets of attributes is given. In fact, one can build the sequence inside the k steps for constructing three-way decision rules. Recall that in general we may form a meet semi-lattice by using a family of sets of attributes. The argument based on a sequence of sets of attributes may be easily generalized to a meet semi-lattice of sets of attributes. In this way, we have a lattice based three-way decision-making process.

5. Conclusion

We generalize three-way decisions with probabilistic rough sets into a sequential three-way framework.

This enables us to consider both the cost of various mis-classifications and the cost of obtaining the necessary evidence for making a classification decision. Although the latter is very important in decision-making, it has not received sufficient attention. This paper reports some preliminary results on the topic.

The exploration of multiple representations of objects for decision-making is a useful direction in granular computing. Typically, decisions made at a higher level of granularity or abstraction may be less accurate or reliable but with a lower cost of resources (i.e., time for obtaining evidence or performing a test). As future work, a formal analysis of such cost-accuracy tradeoff is needed to further justifying sequential three-way decision-making.

References

[1] Duda, R.O. and Hart, P.E., Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[2] Forster, M.R., Key concepts in model selection: performance and generalizability, Journal of Mathematical Psycholog 44, 205-231, 2000.

[3] Goudey, R., Do statistical inferences allowing three alternative decision give better feedback for environmentally precautionary decision-making, Journal of Environmental Management 85, 338-344, 2007.

[4] Li, Y., Zhang, C. and Swanb, J.R., Rough set based model in information retrieval and filtering, Proceeding of the 5th International Conference on Information Systems Analysis and Synthesis, 398-403, 1999.

[5] Li, Y., Zhang, C. and Swanb, J.R., An information flter-ing model on the Web and its application in JobAgent. Knowledge-Based Systems 13, 285-296, 2000.

[6] Lingras, P., Chen, M., and Miao, D.Q., Rough cluster quality index based on decision theory, IEEE Transactions on Knowledge and Data Engineering 21, 1014-1026, 2009.

[7] Liu, D., Yao, Y.Y., and Li, T.R., Three-way investment decisions with decision-theoretic rough sets, International Journal of Computational Intelligence Systems 4, 66-74, 2011.

[8] Lurie, J.D. and Sox, H.C., Principles of medical decision making, Spine 24, 493-498, 1999.

[9] Pauker, S.G. and Kassirer, J.P., The threshold approach to clinical decision making, The New England Journal of Medicine 302, 1109-1117, 1980.

[10] Pawlak, Z., Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Boston, 1991.

[11] Qian, Y.H., Liang, J.Y., Pedrycz, W. and Dang, C.Y., Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174, 597-618, 2010.

[12] Quinlin, J.R., Induction of decision tree, Machine Learning 1, 81-106, 1986.

[13] Ślęzak, D., Wróblewski, J., Eastwood, V. and Synak, P., Brighthouse: an analytic data warehouse for ad-hoc queries, Proceedings of the VLDB Endowment 1, 1337-1345, 2008.

[14] Wald, A., Sequential tests of statistical hypotheses, The Annals of Mathematical Statistics 16, 117-186, 1945.

[15] Weller, A.C., Editorial Peer Review: Its Strengths and Weaknesses, Information Today, Inc., Medford, NJ, 2001.

[16] Woodward, P.W. and Naylor, J.C., An application of Bayesian methods in SPC. The Statistician 42, 461-469, 1993.

[17] Yao, Y.Y., Probabilistic approaches to rough sets, Expert Systems 20, 287-297, 2003.

[18] Yao, Y.Y., A partition model of granular computing, LNCS Transactions on Rough Sets I, LNCS 3100, 232-253, 2004.

[19] Yao, Y.Y., Probabilistic rough set approximations, International Journal of Approximation Reasoning 49, 255-271, 2008.

[20] Yao, Y.Y., Three-way decision: an interpretation of rules in rough set theory. Proceedings of RSKT'09, LNCS (LNAI) 5589, 642-649, 2009.

[21] Yao, Y.Y., Three-way decisions with probabilistic rough sets, Information Sciences 180, 341-353, 2010.

[22] Yao, Y.Y., The superiority of three-way decisions in probabilistic rough set models. Information Sciences 181, 1080-1096, 2011.

[23] Yao, Y.Y. and Luo, J.G., Top-down progressive computing, manuscript, 2011.

[24] Yao, Y.Y. and Wong, S.K.M., A decision theoretic framework for approximating concepts, International Journal of Man-machine Studies 37, 793-809, 1992.

[25] Yao, Y.Y., Wong, S.K.M. and Lingras, P., A decision-theoretic rough set model, in: Ras, Z.W., Zemankova, M. and Emrich, M.L. (Eds.). Methodologies for Intelligent Systems 5, North-Holland, New York, pp. 17-24, 1990.

[26] Yao, Y.Y., Zhao, Y., and Wang, J., On reduct construction algorithms, LNCS Transactions on Computational Science, Vol. II, LNCS 5150, 100-117, 2008.

[27] Yao, Y.Y., Zhao, Y., Wang, J. and Han, S.Q., A model of user-oriented reduct construction for machine learning, LNCS Transactions on Rough Sets, Vol. VIII, LNCS 5084, 341-360, 2008.

[28] Zhao, K. and Wang, J., A reduction algorithm meeting users requirements, Journal of Computer Science and Technology 17, 578593, 2002.

[29] Zhao, W.Q. and Zhu, Y.L., An email classification scheme based on decision-theoretic rough set theory and analysis of email security, Proceeding of 2005 IEEE Region 10 TENCON, 2005,
doi: 10.1109/TENCON.2005.301121

[30] Zhou, B., Yao, Y.Y., and Luo, J.G.: A three-way decision approach to email spam filtering, Proceedings of the 23th Canadian Conference on Artificial Intelligence (AI 2010), LNCS (LNAI) 6085, 28-39, 2010.