Knowledge-Based Systems xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Decision-theoretic rough sets under dynamic granulation

Yanli Sang [a], Jiye Liang [b,*], Yuhua Qian [a,b]

[a] School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006 Shanxi, China

ARTICLE INFO

Article history:
Received 23 January 2015
Received in revised form 31 July 2015
Accepted 1 August 2015
Available online xxxx

Keywords:
Decision-theoretic rough set theory
Dynamic granulation
Monotonicity
Bayesian decision theory

ABSTRACT

Decision-theoretic rough set theory is quickly becoming a research direction in rough set theory, which is a general and typical probabilistic rough set model with respect to its threshold semantics and decision features. However, unlike the Pawlak rough set, the positive region, the boundary region and the negative region of a decision-theoretic rough set are not monotonic as the number of attributes increases, which may lead to overlapping and inefficiency of attribute reduction with it. This may be caused by the introduction of a probabilistic threshold. To address this issue, based on the local rough set and the dynamic granulation principle proposed by Qian et al., this study will develop a new decision-theoretic rough set model satisfying the monotonicity of positive regions, in which the two parameters α and β need to dynamically update for each granulation. In addition to the semantic interpretation of its thresholds itself, the new model not only ensures the monotonicity of the positive region of a target concept (or decision), but also minimizes the local risk under each granulation. These advantages constitute important improvements of the decision-theoretic rough set model for its better and wider applications.

1. Introduction

Rough set theory proposed by Pawlak in 1982 [23] has become an important tool for dealing with uncertainty management and uncertainty reasoning. Because of no prior knowledge, the rough set theory has a wide variety of applications including pattern recognition, data mining, machine learning, knowledge discovery, and so on [3,6,7,10,12,13,11,16,29,34,52]. As we know, the lower approximation of a set in rough set theory is defined by a strict inclusion relation, which may lead to its sensitivity to noisy data for attribute reduction and classification tasks. For this observation, through incorporating probabilistic approaches to rough set theory, several probabilistic generalizations of rough sets have been proposed [37,42,46,60], in which threshold values are aforehand given. In recent years, based on different threshold arrangements, different versions of probabilistic rough set approaches were proposed one after another, such as the 0.5-probabilistic rough set [24], the decision-theoretic rough set model [43,44,47], the variable precision rough set (VPRS) model [59], membership functions [26], parameterized rough set models [4], Bayesian rough set model [35], game-theoretic rough set [5], and so on.

Within the family of probabilistic rough sets, the semantic interpretation of the required threshold parameters is the most fundamental difficulty with the probabilistic approximations. In the literature [43,44], we saw the first report to solve this difficulty for probabilistic rough set approximations in a decision-theoretic framework. In the framework of the decision theory, Bayesian decision theory was firstly introduced to minimize the decision costs, which provides a scientific method for determining and interpreting threshold values through taking costs and risks into account. From this viewpoint, we can say that the decision-theoretic rough set has a threshold semantic interpretation. It deserves to point out that the decision-theoretic rough set model can be regarded as a generalization of probabilistic rough set models [46] because it can derive various existing rough set models through setting different thresholds. Based on this framework, Yao [47] then presented a new decision-making method, called a three-way decision method, in which positive region, boundary region and negative region are respectively seen as three actions. In the literature [48], the author further emphasized the superiority of three-way decisions in probabilistic rough set models. More recently, Zhang et al. [53] introduced a new recommender system to consult the user for the choice by combining three-way decisions and random forests. Yu et al. [50] proposed a tree-based incremental overlapping clustering method using three-way decision theory. To date, the theoretical framework have been largely enriched since

* Corresponding author. Tel./fax: +86 0351 7018176.
E-mail addresses: sangyl@sxu.edu.cn (Y. Sang), ljy@sxu.edu.cn (J. Liang), jinchengqyh@126.com (Y. Qian).

http://dx.doi.org/10.1016/j.knosys.2015.08.001
0950-7051/© 2015 Published by Elsevier B.V.

the decision-theoretic rough sets were proposed [8,9,32,38,57]. The decision-theoretic rough set model, in recent years, has also been used in many applications, such as decision-making [38], clustering analysis [49,50], spam filtering [58], investment decisions [21], multi-view decision models [57] and multiple-category classification [56].

It is well known that, in the Pawlak rough set model [25], the lower approximation of a given target concept with respect to an equivalence relation R is much smaller than the corresponding lower approximation with respect to an equivalence relation $R' \prec R$. This property is called monotonicity. Naturally, given a target decision, its positive region, boundary region and negative region are all monotonic in the framework of the Pawlak rough set as well. However, in probabilistic approximations, because of the introduction of probabilistic thresholds, the conditional probability of an object x classified into a target concept may increase or decrease as the number of attributes becomes bigger. In other words, the monotonicity of lower approximations of a target concept may not hold in probabilistic approximation models. Accordingly, the positive region, boundary region and negative region of a given target decision have the same observation in terms of probabilistic approximations.

In what follows, we analyze the importance of the monotonicity of a lower approximation in the decision-theoretic rough set (DTRS). As we know, attribute reduction is one key issue in rough set theory, based on which one can extract decision rules for prediction from an information system with class labels. Attribute reduction of a target decision aims at finding a subset of attributes such that it is at least as good as the original attribute set from the viewpoint of decision ability. If the lower approximation of a target concept is not monotonic, a found attribute reduct may be overlapping because of the strict definition of attribute reduction. Except for this shortcoming, the process of attribute reduction is also computationally time-consuming. To overcome these two issues, it is very desirable to develop a new decision-theoretic rough set satisfying the monotonicity of a target concept, which is the main motivation of this study.

In fact, several studies about the monotonicity of attribute reduction using DTRS have been reported [8,21,22,45,55]. Yao and Zhao [45] presented various criteria including the decision-monotonicity criterion, the generality criterion and the cost criterion for attribute reduction of probabilistic rough set models. From the viewpoint of information theory, Ma et al. [22] proposed three new monotonic measure functions by considering variants of conditional information entropy for obtaining a monotonic attribute reduction process. Li et al. [15] developed a so-called positive region expanding reduct. Blaszczyński [1] considered three types of monotonicity properties and proposed several new measures with monotonicity such that the corresponding lower approximation satisfies monotonicity. Although these studies have provided several alternative solutions, how to solve the non-monotonicity of lower approximations keeping the conditional probability form unchanged is still an open problem in the decision-theoretic rough set.

To address the above problem, from the viewpoint of granular computing [19,20,41,51], this paper develops a new probabilistic rough set framework under dynamic granulation, called the decision-theoretic rough set under dynamic granulation (DG-DTRS). There are two main improvements in the proposed model. For the first improvement, given a target concept, we only judge whether each of objects within it is included in its lower approximation or not, rather than the entire universe. For the second improvement, we need to dynamically update the threshold parameters α and β when granular structures for approximating a target concept/decision are changed. Therefore, besides the semantic interpretation of its thresholds, the proposed model not only ensures the monotonicity of the positive region of a target concept (or decision), but also minimizes the local risk under each granulation. Hence, the DG-DTRS with these advantages can be seen as an important improvement of the existing decision-theoretic rough set model.

The study is organized as follows. Some basic concepts in Pawlak rough sets and decision-theoretic rough sets are briefly reviewed in Section 2. In Section 3, a new probabilistic set-approximation approach is constructed in the context of dynamic granulation world, and some of its nice properties are explored. Furthermore, based on Bayesian decision procedure, we also give a method for updating the required threshold parameters in the proposed model. Finally, Section 4 concludes this paper by bringing some remarks and discussions.

2. Preliminary knowledge on decision-theoretic rough sets

In this section, we briefly review some basic concepts of decision-theoretic rough set model.

2.1. Pawlak's rough set

A decision table is a tuple $S=(U, AT=C\cup D, V_a|a\in At, I_a|a\in At)$, where U is a finite non-empty set of objects, called a universe, C is a non-empty finite set of conditional attributes, D is a finite set of decision attributes, V_a ($a\in AT$) is the domain of attribute a, and $I_a : U \to V_a$ is an information function that maps an object in U to exactly one value in V_a. A decision table is simply denoted by $S=(U, At=C\cup D)$ [25].

An attribute subset $A\subseteq At$ determines an equivalence relation E_A (or simply E). That is,

$$E_A=\{(x,y)\in U\times U|\forall a\in A,\ I_a(x)=I_a(y)\}. \quad (181)$$

Two objects in U are equivalent to each other if and only if they have the same values on all attributes in A. An equivalence relation is reflexive, symmetric and transitive.

The pair $apr=\langle U, E_A\rangle$ is called an approximation space defined by the attribute set A [25]. The equivalence relation E_A induces a partition of U, denoted by U/E_A or U/A. An object $x\in U$ is described by its equivalence class of U/E_A : $[x]_{E_A}=[x]_A=\{y\in U|(x,y)\in E_A\}$. Each equivalence class $[x]_A$ may be viewed as an information granule consisting of indistinguishable elements. The granular structure induced by an equivalence relation is a partition of the entire universe.

Given an approximation space $\langle U, E_A\rangle$. For an arbitrary subset $X\subseteq U$, one can construct its lower and upper approximations with information granules of the universe induced by the partition U/A via the following definition:

$$\underline{apr}_A(X)=\cup\{[x]_A\subseteq X|x\in U\},$$
$$\overline{apr}_A(X)=\cup\{[x]_A\cap X\neq\emptyset|x\in U\}. \quad (199)$$

The pair $\langle\underline{apr}_A(X), \overline{apr}_A(X)\rangle$ is called a rough set of X with respect to the equivalence relation E_A. Equivalently, they can also be rewritten as

$$\underline{apr}_A(X)=\{x|P(X|[x]_A)=1|x\in U\},$$
$$\overline{apr}_A(X)=\{x|P(X|[x]_A)>0|x\in U\}, \quad (205)$$

where $P(X|[x]_A)$ denotes the conditional probability that the object x belongs to a target concept X.

Through using the rough set approximations of X defined by A, the universe U is divided into three disjoint regions: the positive region $POS_A(X)$, the boundary region $BND_A(X)$ and the negative region $NEG_A(X)$ of X:

$$POS_A(X) = \underline{apr}_A(X),$$
$$BND_A(X) = \overline{apr}_A(X) - \underline{apr}_A(X),$$
$$NEG_A(X) = U - (POS_A(X) \cup BND_A(X)) = U - \overline{apr}_A(X).$$

These three regions are often used to predict the class label of an unseen object in rough set theory.

2.2. Decision-theoretic rough sets

A decision-theoretic rough set model is a typical probabilistic rough set model, in which Bayesian decision procedure is introduced to minimize the decision costs. The rough set model provides a systematic method to set the required threshold parameters from the viewpoint of loss functions. In this subsection, we review some basic concepts in the decision-theoretic rough set model [43].

In the Bayesian decision procedure, a finite set of states can be written as $\Omega = \{\omega_1, \ldots, \omega_s\}$, and a finite set of m possible actions can be denoted by $\mathcal{A} = \{a_1, \ldots, a_m\}$. Let $P(\omega_j|\mathbf{x})$ be the conditional probability of an object x being in state ω_j given that the object is described by $\mathbf{x}$. Let $\lambda(a_i|\omega_j)$ denote the loss, or cost, for taking action a_i when the state is ω_j. Suppose taking action a_i when the state is ω_j, then the expected loss associated with taking action a_i can be given by:

$$R(a_i|\mathbf{x}) = \sum_{j=1}^{s} \lambda(a_i|\omega_j)P(\omega_j|\mathbf{x})$$

In the decision-theoretic rough set theory, given an approximation space $apr = \langle U, E_A \rangle$ and an arbitrary subset $X \subseteq U$, the approximation operators partition the universe into three disjoint classes: the positive region $POS_A(X)$, the boundary region $BND_A(X)$ and the negative region $NEG_A(X)$ [47]. The classification of objects according to approximation operators can be easily fitted into the Bayesian decision-theoretic framework. The set of states is given by $\Omega = \{X, X^c\}$ indicating that an object is in a decision class X and not in X, respectively. Based on the three regions, the set of actions is given by $\mathcal{A} = \{a_1, a_2, a_3\}$, where a_1, a_2 and a_3 represent the three actions in classifying an object x, deciding $POS_A(X)$, deciding $NEG_A(X)$, and deciding $BND_A(X)$, respectively. Through using the conditional probability $P(X|[x]_A)$, the Bayesian decision procedure can decide how to assign x into these three disjoint regions [50,52]. Let $\lambda(a_i|X)$ denote the loss incurred for taking action a_i when an object belongs to X, and let $\lambda(a_i|X^c)$ denote the loss incurred for taking the same action when the object does not belong to X.

The expected loss $R(a_i|[x]_A)$ associated with taking the individual actions can be expressed as:

$$R_1 = R(a_1|[x]_A) = \lambda_{11}P(X|[x]_A) + \lambda_{12}P(X^c|[x]_A),$$
$$R_2 = R(a_2|[x]_A) = \lambda_{21}P(X|[x]_A) + \lambda_{22}P(X^c|[x]_A),$$
$$R_3 = R(a_3|[x]_A) = \lambda_{31}P(X|[x]_A) + \lambda_{32}P(X^c|[x]_A),$$

where $\lambda_{i1} = \lambda(a_i|X)$, $\lambda_{i2} = \lambda(a_i|X^c), i = 1, 2, 3$. The Bayesian decision procedure leads to the following minimum-risk decision rules:

(P) if $R_1 \leqslant R_2$ and $R_1 \leqslant R_3$, decide $x \in POS_A(X)$;
(N) if $R_2 \leqslant R_1$ and $R_2 \leqslant R_3$, decide $x \in NEG_A(X)$;
(B) if $R_3 \leqslant R_1$ and $R_3 \leqslant R_2$, decide $x \in BND_A(X)$.

Consider a special kind of loss functions with $\lambda_{11} \leqslant \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \leqslant \lambda_{32} < \lambda_{12}$, that is, the cost of classifying an object x belonging to X into the positive region $POS(X)$ is less than or equal to the cost of classifying x into the boundary region $BND(X)$, and both of these costs are strictly less than the cost of classifying x into the negative region $NEG(X)$. The reverse order of cost is used for classifying an object not in X. This assumption implies that $\alpha \in (0, 1], \gamma \in (0, 1)$, and $\beta \in [0, 1)$. In this case, the minimum-risk decision rules can be re-expressed as:

(P) if $P(X|[x]_A) \geqslant \alpha$ and $P(X|[x]_A) \geqslant \gamma$, decide $x \in POS_A(X)$,
(N) if $P(X|[x]_A) \leqslant \beta$ and $P(X|[x]_A) \leqslant \gamma$, decide $x \in NEG_A(X)$,
(B) if $\beta \leqslant P(X|[x]_A) \leqslant \alpha$, decide $x \in BND_A(X)$,

where

$$\alpha = \frac{\lambda_{12} - \lambda_{32}}{(\lambda_{31} - \lambda_{32}) - (\lambda_{11} - \lambda_{12})},$$
$$\gamma = \frac{\lambda_{12} - \lambda_{22}}{(\lambda_{21} - \lambda_{22}) - (\lambda_{11} - \lambda_{12})},$$
$$\beta = \frac{\lambda_{32} - \lambda_{22}}{(\lambda_{21} - \lambda_{22}) - (\lambda_{31} - \lambda_{32})}.$$

If a loss function with $\lambda_{11} \leqslant \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \leqslant \lambda_{32} < \lambda_{12}$, it further satisfies the condition: $(\lambda_{12} - \lambda_{32})(\lambda_{21} - \lambda_{31}) \geqslant (\lambda_{31} - \lambda_{11})(\lambda_{32} - \lambda_{22})$, then $1 \geqslant \alpha > \gamma > \beta \geqslant 0$. In this case, after tie-breaking, the following simplified decision rules are obtained:

(P1) if $P(X|[x]_A) \geqslant \alpha$, decide $x \in POS_A(X)$;
(N1) if $P(X|[x]_A) \leqslant \beta$, decide $x \in NEG_A(X)$;
(B1) if $\beta < P(X|[x]_A) < \alpha$, decide $x \in BND_A(X)$.

After computing the two parameters α and β from the loss functions, using the above decision rules, we get the probabilistic approximations as follows:

$$\underline{apr}_A^{(\alpha,\beta)}(X) = \{x \in U | P(X|[x]_A) \geqslant \alpha\},$$
$$\overline{apr}_A^{(\alpha,\beta)}(X) = \{x \in U | P(X|[x]_A) > \beta\}.$$

The combination of these two approximations is called a decision-theoretic rough set (DTRS). In DTRS, three kinds of probabilistic regions (positive, boundary and negative regions) of concept X are defined as follows:

$$POS_A^{(\alpha,\beta)}(X) = \underline{apr}_A^{(\alpha,\beta)}(X),$$
$$BND_A^{(\alpha,\beta)}(X) = \overline{apr}_A^{(\alpha,\beta)}(X) - \underline{apr}_A^{(\alpha,\beta)}(X),$$
$$NEG_A^{(\alpha,\beta)}(X) = U - POS_A^{(\alpha,\beta)}(X) \cup BND_A^{(\alpha,\beta)}(X).$$

In the framework of decision-theoretic rough sets, many existing models such as the Pawlak rough set model, variable precision rough set model and Bayesian rough set model, can be explicitly derived by considering various classes of loss functions. Therefore, we have regarded it as a general and fundamental probabilistic rough set model.

3. Decision-theoretic rough set models under dynamic granulation

Multigranulation rough set theory was proposed by Qian [27] in 2006, in which lower and upper approximations are approximated by granular structures induced by multiple binary relations instead of single binary relation. In a sense, the multigranulation rough set is a kind of information fusion strategies through fusing multiple granular structures. Qian et al. [30,31] have proposed optimistic and pessimistic multigranulation rough sets which are based on optimistic and pessimistic strategies, respectively. In recent years, many extended multigranulation rough set models have also been proposed and studied [14,17,29,36,39,40,54]. Another multigranulation rough set is characterized by dynamic granular structures

Please cite this article in press as: Y. Sang et al., Decision-theoretic rough sets under dynamic granulation, Knowl. Based Syst. (2015), http://dx.doi.org/10.1016/j.knosys.2015.08.001

[28]. For example, the positive approximation can be seen as one representative of them, in which a rough set is constructed by a dynamic granulation order with hierarchical structure [28]. The positive approximation is constructed by a sequence of granulation worlds stretching from coarse to fine granulation, which can be used to accelerate a heuristic process of attribute reduction.

In the view of granular computing [51], in existing decision-theoretic rough set models, a target concept described by a set is always characterized with upper and lower approximations under a single granulation. Qian et al. [32] proposed multigranulation decision-theoretic rough sets (MG-DTRS) for extending its wider applications such as multi-source data analysis, knowledge discovery from data with high dimensions and distributive information systems. However, unlike the Pawlak rough set, the positive region, the boundary region and the negative region of a decision-theoretic rough set is not monotonic as the number of attributes increases, which may lead to overlapping and inefficiency of attribute reduction with it.

To address this issue, without loss of generality, in this section we first investigate the monotonicity of positive regions through comparing the Pawlak rough set model with the decision-theoretic rough set model, develop a new decision-theoretic rough set under dynamic granulation from the viewpoint of granular computing (called decision-theoretic rough sets under dynamic granulation), and investigate some of its important properties.

3.1. Non-monotonicity of probabilistic positive regions in DTRS

Given a decision table $S = (U, At = C \cup D)$ with $P, Q \subseteq C$. A partial relation $\preccurlyeq$ on 2^C can be defined as follows [2,18,28]:

$$P \preccurlyeq Q \iff \forall x \in U,\ [x]_P \subseteq [x]_Q.$$

That is, if $P \preccurlyeq Q$, then Q is said to be coarser than P (or P is finer than Q). If $P \preccurlyeq Q$ and $U/P \neq U/Q$, Q is said to be strictly coarser than P or P is strictly finer than Q, denoted by $P \prec Q$.

Given a decision table $S = (U, At = C \cup D)$, for an arbitrary subset $X \subseteq U$, from the definition of lower/upper approximation in the Pawlak rough set, we can immediately obtain the monotonicity of the positive region of X as follows:

$$P \preccurlyeq Q \Rightarrow POS_P(X) \supseteq POS_Q(X).$$

That is, a thinner partition induces a larger positive region.

In the following, we can extend the monotonic property of a single set to a decision partition $U/D = \{D_1, D_2, \ldots, D_m\}$ of the universe as follows:

$$P \preccurlyeq Q \Rightarrow \forall D_i \in U/D, \quad POS_P(D_i) \supseteq POS_Q(D_i),$$

and thus

$$P \preccurlyeq Q \Rightarrow POS_P(D) \supseteq POS_Q(D).$$

From the above properties, we can see that the positive regions of a decision partition induced by the decision attributes also satisfy the monotonicity in the context of the Pawlak rough set model. Naturally, given a target decision, its negative region and boundary region have the same monotonicity in the framework of the Pawlak rough set model [25].

However, in the decision-theoretic rough set, we cannot obtain the monotonicity of probabilistic positive regions of a target (or decision). In the decision-theoretic rough set, if one object is included in the lower approximation of a target concept X, then all objects coming from its equivalence class are putted into this lower approximation. This means that the lower approximation of a probabilistic rough set may overflow the range of a target concept. In addition, in the process of a heuristic attribute reduction, the probabilistic positive region of a target decision may not monotonically increase as the number of attributes becomes larger, which is caused by the fact that the conditional probability function is not a monotonic function with respect to the equivalence class $[x]$. This is illustrated by the following example.

Example 1. Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be a universe, $U/P, U/Q$ two partitions on U, where

$$U/P = \{\{x_1, x_2, x_3\}, \{x_4\}, \{x_5, x_6, x_7, x_8\}, \{x_9, x_{10}\}\},$$
$$U/Q = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6\}, \{x_7, x_8\}, \{x_9, x_{10}\}\}.$$

Here, we suppose two parameters $(\alpha, \beta) = (0.6,\ 0.2)$. Then, from the definition of the partial relation, it is obvious that $Q \preccurlyeq P$ holds.

Take a target concept $X = \{x_1, x_3, x_5, x_6, x_7, x_9\}$. Based on the definition of probabilistic lower approximation in DTRS, through computing the condition probability of $x \in U$, we have that

$$\underline{apr}_P^{(0.6,0.2)}(X) = \{x_1, x_2, x_3, x_5, x_6, x_7, x_8\},$$
$$\underline{apr}_Q^{(0.6,0.2)}(X) = \{x_3, x_5, x_6\}.$$

That is, for $Q \preccurlyeq P$, we obtain $POS_Q^{(\alpha,\beta)}(X) \subseteq POS_P^{(\alpha,\beta)}(X)$. This means that probabilistic positive region of a target concept with the number of attributes decreasing may enlarge, which indicates that the monotonicity of positive regions does not hold in the DTRS model.

In addition, the probabilistic lower approximation defined in DTRS may overflow the range of a target concept, which would seriously affect the implementation of the monotonicity.

In order to facilitate this study, we will adopt the form of local rough set approximations proposed by Qian et al. [33] to modify the original decision-theoretic rough set. Based on this idea, we first give its definition as follows.

Definition 1. Let $K = \langle U, E_A \rangle$ be an approximation space and an arbitrary subset $X \subseteq U$. Then the $L - (\alpha, \beta)$ lower and upper approximations are defined by

$$\underline{apr}_A^{(\alpha,\beta)}(X) = \{x | P(X|[x]_A) \geqslant \alpha, x \in X\},$$
$$\overline{apr}_A^{(\alpha,\beta)}(X) = \cup\{[x]_A | P(X|[x]_A) > \beta, x \in X\}.$$

The pair $\langle \underline{apr_A}^{(\alpha,\beta)}(X), \overline{apr_A}^{(\alpha,\beta)}(X) \rangle$ is called a local decision-theoretic rough set (L-DTRS).

It can be seen from the above definition, compared with the classical probabilistic set-approximations, that we change the range of the objects in the lower approximation of a concept. That is to say, in $L - (\alpha, \beta)$ approximations, we only judge whether the objects coming from a target concept belong to its lower/upper approximations or not, while in the existing decision-theoretic rough set, we need to consider all objects in the entire universe. It deserves to point out that the computation of its lower/upper approximation is only based on the information granules determined by objects within a target concept, rather than the given universe.

Obviously, the above $L - (\alpha, \beta)$ lower approximation satisfies the following property

$$\underline{apr}_A^{(\alpha,\beta)}(X) \subseteq X.$$

However, for a classical decision-theoretic rough set, this property may not hold.

In the following studies, in order to overcome the non-monotonicity of positive regions in the DTRS model, we will introduce a new probabilistic rough set approximation approach through combining the local decision-theoretic rough set and the idea of dynamic granulation, in which a target concept is approximated by the dynamic granular structures.

A partition induced by an equivalence relation provides a granulation world for describing a target concept. Thus, a sequence of granulation worlds can be determined by a sequence of attribute sets in the power set of attributes, which is called a dynamic granulation order [28]. For the sake of the monotonicity study, in this paper, we only discuss that the dynamic granulation order is a sequence of granulation worlds stretching from coarse to fine granulation which can be determined by a sequence of attribute sets with granulations from coarse to fine in the power set of attributes.

Generally, we introduce the description of dynamic granulation worlds as follows [28]:

Given a decision table $S = (U, At = C \cup D)$, $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l \leqslant n$, we can define a dynamic granulation order denoted by $P_l = \{A_1, A_2, \ldots, A_l\}, l \leqslant n$. In practice, a granulation order on an attribute set can be appointed by users or experts constructed according to the significance of each attribute. Based on this viewpoint, we can redefine the probabilistic approximation under dynamic granulation worlds by using local decision-theoretic rough set approximations.

3.2. Thresholds computing under dynamic granulation worlds

The DTRS model is a typical probabilistic rough set model in which Bayesian decision theory is introduced to minimize the decision costs, and it provides a scientific method to calculate threshold values based on loss functions using more familiar notions of costs (or risks) [46]. To modify the classical decision-theoretic rough set, in this subsection, we firstly need to give a method for updating the required threshold parameters α and β. The Bayesian decision procedure is still employed for achieving this task.

In the following, we give an approach to calculate the required threshold parameters in the new model, which needs to continually perform a Bayesian decision procedure on the gradually reduced universe for obtaining a sequence of threshold parameters under a given dynamic granulation order. The approach of updating threshold parameters is to select a series of actions for which the classification risk is as small as possible.

Let $G = \{\langle U_1, E_{A_1}\rangle, \ldots, \langle U_n, E_{A_n}\rangle\}$ be a group of approximation spaces. Let $U_k \subseteq U, (k = 1, 2, \ldots, n)$ denote a gradually reduced universe satisfied with $U_1 = U, U_{k+1} = U_k - \underline{apr}_{A_k}^{(\alpha_k, \beta_k)}(X_k)$, where $\underline{apr}_{A_k}^{(\alpha_k, \beta_k)}(X_k) = \{x | P(X_k | [x]_{A_k}) \geqslant \alpha_k, x \in X_k\}$ (see Definition 1 for details) and $P = \{A_1, A_2, \ldots, A_n\}$ is a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_k \in 2^C$, $k = 1, 2, \ldots, n$. Then, we present a brief description of the updating parameters process with the Bayesian decision theory for the kth approximation space.

Given the kth approximation space $\langle U_k, E_{A_k}\rangle \in G$ $(k \leqslant n)$. On the universe U_k, the equivalence relation E_{A_k} induces a partition U_k / E_{A_k} and the subset $X_k \subseteq U_k$ is updated with $X_{k+1} = X_k - \underline{apr}_{A_k}^{(\alpha_k, \beta_k)}(X_k)$. $P(X_k | [x]_{A_k})$ and $P(X_k^c | [x]_{A_k})$ are the conditional probabilities of an object in the equivalence class $[x]_{A_k}$ within X_k and X_k^c, respectively. Given the loss function matrix under the kth granular space, the expected loss $R(a_i | [x]_{A_k})$ associated with taking action a_i under the kth granular space can be expressed as:

$$
\begin{aligned}
R(a_1 | [x]_{A_k}) &= \lambda_{11}^k P(X_k | [x]_{A_k}) + \lambda_{12}^k P(X_k^c | [x]_{A_k}), \\
R(a_2 | [x]_{A_k}) &= \lambda_{21}^k P(X_k | [x]_{A_k}) + \lambda_{22}^k P(X_k^c | [x]_{A_k}), \\
R(a_3 | [x]_{A_k}) &= \lambda_{31}^k P(X_k | [x]_{A_k}) + \lambda_{32}^k P(X_k^c | [x]_{A_k}),
\end{aligned}
$$

where λ_{ij}^k denotes the loss function for taking action a_i when state is ω_j by the kth granular space, and $\lambda_{ij}^r \neq \lambda_{ij}^s$ $(r, s \in \{1, 2, \ldots, n\}, r \neq s)$.

In practical applications, in our opinion, according to various requirements under the change of granular space, the loss functions regarding the risk or cost of actions are also updated correspondingly. Thus, one assumes that the values of $\lambda_{ij}^k (k \leqslant n)$ in each granular space could not be equivalent to each other. In other words, each granular space should have its independent loss (or cost) functions itself.

Like the decision-theoretic rough set, briefly, we also assume that the loss function satisfies the conditions:

(i) $\lambda_{11}^k \leqslant \lambda_{31}^k < \lambda_{21}^k$,
(ii) $\lambda_{22}^k \leqslant \lambda_{32}^k < \lambda_{12}^k$,
(iii) $(\lambda_{12}^k - \lambda_{32}^k)(\lambda_{21}^k - \lambda_{31}^k) \geqslant (\lambda_{31}^k - \lambda_{11}^k)(\lambda_{32}^k - \lambda_{22}^k)$.

It follows that $1 \geqslant \alpha \geqslant \gamma \geqslant \beta \geqslant 0$. By decision rules (P1)-(B1), we can obtain the corresponding positive region, the boundary region and the negative region under the kth granular space as follows:

$$
\begin{aligned}
POS_{A_k}^{(\alpha,\beta)}(X_k) &= \{x | P(X_k | [x]_{A_k}) \geqslant \alpha_k, x \in U_k\}, \\
BND_{A_k}^{(\alpha,\beta)}(X_k) &= \{x | \beta_k < P(X_k | [x]_{A_k}) < \alpha_k, x \in U_k\}, \\
NEG_{A_k}^{(\alpha,\beta)}(X_k) &= \{x | P(X_k | [x]_{A_k}) \leqslant \beta_k, x \in U_k\},
\end{aligned}
$$

where

$$
\alpha_k = \frac{\lambda_{12}^k - \lambda_{32}^k}{(\lambda_{31}^k - \lambda_{32}^k) - (\lambda_{11}^k - \lambda_{12}^k)},
$$

$$
\beta_k = \frac{\lambda_{32}^k - \lambda_{22}^k}{(\lambda_{21}^k - \lambda_{22}^k) - (\lambda_{31}^k - \lambda_{32}^k)}.
$$

Hence, according to the calculation procedure of threshold parameters in the approximation space $\langle U_k, E_{A_k}\rangle$ above, given a dynamic granulation order P_l $(l \leqslant n)$, we can obtain a sequence of the threshold parameters $(\alpha, \beta)_l = \{(\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_l, \beta_l)\}$, which means the procedure of dynamically updating the required threshold parameters with the various costs or risks by every granular space. The threshold parameters sequence will be used in the definition of the probabilistic approximations that will be proposed in next subsection. It deserves to point out that when the loss function in different granular spaces satisfies with the condition: $\lambda_{12}^k = \lambda_{21}^k = 1$, $\lambda_{11}^k = \lambda_{22}^k = \lambda_{31}^k = \lambda_{32}^k = 0$, $k \leqslant l$, from the above equation, we have $(\alpha_k, \beta_k) = (1, 0)$, which can be regraded as a special case.

3.3. Decision-theoretic rough sets under dynamic granulation

In this subsection, we introduce a new decision-theoretic rough set under dynamic granulation orders and investigate some of its important properties.

Firstly, we give the definition of the new decision-theoretic rough set as follows.

Definition 2. Let $S = (U, At = C \cup D)$ be a decision table, $X \subseteq U$ and $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l = 1, 2, \ldots, n$. Given a dynamic granulation order $P_l = \{A_1, A_2, \ldots, A_l\}$ $(l \leqslant n)$, we define $P_l^{(\alpha,\beta)}$-lower approximation $\underline{P_l}^{(\alpha,\beta)}(X)$ and $P_l^{(\alpha,\beta)}$-upper approximation $\overline{P_l}^{(\alpha,\beta)}(X)$ of X under the dynamic granulation order as

$$
\underline{P_l}^{(\alpha,\beta)_l}(X) = \{x | P(X_k | [x]_{A_k}) \geqslant \alpha_k, x \in X_k, k = 1, 2 \ldots, l\}, \tag{1}
$$

$$
\overline{P_l}^{(\alpha,\beta)_l}(X) = \cup\{[x]_{A_k} | P(X_k | [x]_{A_k}) > \beta_k, x \in X_k, k = 1, 2 \ldots, l\}, \tag{2}
$$

Please cite this article in press as: Y. Sang et al., Decision-theoretic rough sets under dynamic granulation, Knowl. Based Syst. (2015), http://dx.doi.org/10.1016/j.knosys.2015.08.001

where $X_1 = X, X_{k+1} = X - \bigcup_{j=1}^{k} \underline{apr}_{A_j}^{(\alpha_j,\beta_j)}(X_j)$, $(\alpha,\beta)_l = \{(\alpha_1,\beta_1),(\alpha_2,\beta_2),\ldots,(\alpha_l,\beta_l)\}$ indicates the dynamic threshold parameter sequence under the current granulation order P_l, and $[x]_{A_k}$ represents the equivalence class including x in the partition U_k/A_k in which $U_k = U_{k-1} - \underline{apr}_{A_{k-1}}^{(\alpha_{k-1},\beta_{k-1})}(X_{k-1})$ is the gradually reduced universe.

It can be seen from the above definition that the target concept can be gradually approximated by using dynamic granulations stretching from coarse to fine on the gradually reduced universe.

In addition, we can find that the computation of its lower/upper approximation is only based on the information granules determined by objects within a target concept X, rather than the universe U. Obviously, we have the property

$$\underline{P_l}^{(\alpha,\beta)}(X) \subseteq X.$$

In order to further characterize the structure of probabilistic approximations in the DG-DTRS, we can use local probabilistic approximations in a single granulation world to redefine $P_l^{(\alpha,\beta)}$-set approximations of a target concept X, which can be regarded as an equivalent form of the above definition. That is

$$\underline{P_l}^{(\alpha,\beta)_l}(X) = \bigcup_{k=1}^{l} \underline{apr}_{A_k}^{(\alpha_k,\beta_k)}(X_k), \tag{3}$$

$$\overline{P_l}^{(\alpha,\beta)_l}(X) = \bigcup_{k=1}^{l} \overline{apr}_{A_k}^{(\alpha_k,\beta_k)}(X_k), \tag{4}$$

where $X_1 = X, X_{k+1} = X_k - \underline{apr}_{A_k}^{(\alpha_k,\beta_k)}(X_k)$. The above definition form can reflect the structure feature of probabilistic approximations in DG-DTRS.

Fig. 1 visualizes the hierarchical construction of lower approximation of a target concept in the DG-DTRS model.

In Fig. 1, let $P_1 = \{A_1\}$ and $P_2 = \{A_1, A_2\}$ with $A_1 \succcurlyeq A_2$ be two granulation orders. $\underline{apr}_{A_1}^{(\alpha_1,\beta_1)}(X_1)$ is the L-lower approximation of X_1 obtained by the equivalence relation E_{A_1} on the universe U_1, where the parameter $(\alpha_1,\beta_1) = (0.8, 0.2)$; $\underline{apr}_{A_2}^{(\alpha_2,\beta_2)}(X_2)$ is the L-lower approximation of X_2 obtained by the equivalence relation E_{A_2} on the universe U_2, where the parameter $(\alpha_2,\beta_2) = (0.6, 0.2)$. Hence, $\underline{P_2}^{(\alpha,\beta)_2} = \underline{apr}_{A_1}^{(\alpha_1,\beta_1)}(X_1) \cup \underline{apr}_{A_2}^{(\alpha_2,\beta_2)}(X_2)$. The mechanism illustrates the hierarchical structure of probabilistic approximations in the DG-DTRS, which can be used to gradually compute the lower approximation of a target concept.

From the above definition and Fig. 1, we have the following theorem.

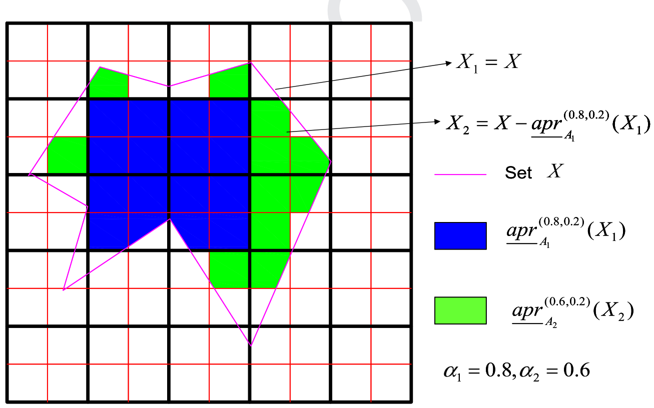

Fig. 1. Dynamic granular structures of the lower approximation in DG-DTRS.

Theorem 1 (*Lower approximation monotonicity*). *Let $S = (U, At = C \cup D)$ be a decision table, $X \subseteq U$ and $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l = 1, 2\ldots, n$. Given $P_l = \{A_1, A_2, \ldots, A_l\}$, then for any P_l, we have*

$$\underline{P_1}^{(\alpha,\beta)_1}(X) \subseteq \underline{P_2}^{(\alpha,\beta)_2}(X) \subseteq \cdots \subseteq \underline{P_l}^{(\alpha,\beta)_l}(X),$$

where $(\alpha,\beta)_l$ indicates the sequence of probabilistic threshold parameters under the granulation order P_l.

This theorem shows that the monotonicity property of $P_l^{(\alpha,\beta)}$-lower approximation of a given target concept X under dynamic granulation orders holds in the DG-DTRS model. It is illustrated by the following example.

Example 2. Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be a universe, $U/A_1, U/A_2$ two partitions on U, where

$$U/A_1 = \{\{x_1,x_2\},\{x_3,x_4\},\{x_5,x_6,x_7,x_8\},\{x_9,x_{10}\}\},$$
$$U/A_2 = \{\{x_1\},\{x_2\},\{x_3,x_4\},\{x_5,x_6\},\{x_7,x_8\},\{x_9\},\{x_{10}\}\}.$$

Obviously, $A_1 \succcurlyeq A_2$ holds. Thus, we can construct two dynamic granulation orders $P_1 = \{A_1\}$ and $P_2 = \{A_1, A_2\}$.

Given a target concept $X = \{x_2, x_3, x_5, x_6, x_8, x_{10}\}$, assume $(\alpha,\beta)_2 = \{(0.7, 0.2), (0.8, 0.2)\}$. From Definition 2, by computing the lower and upper approximations of X under these two granulation orders, one easily obtains that

$$\underline{P_1}^{(\alpha,\beta)_1}(X) = \{x_5, x_6, x_8\},$$
$$\overline{P_1}^{(\alpha,\beta)_1}(X) = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\},$$
$$\underline{P_2}^{(\alpha,\beta)_2}(X) = \{x_2, x_5, x_6, x_8, x_{10}\},$$
$$\overline{P_2}^{(\alpha,\beta)_2}(X) = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}.$$

That is to say, the target concept X can be approximated by using the two granulation orders P_1 and P_2 in DG-DTRS. Moreover, $\underline{P_1}^{(\alpha,\beta)_1}(X) \subseteq \underline{P_2}^{(\alpha,\beta)_2}(X)$ holds.

Based on Eqs. (3) and (4), the corresponding probabilistic positive region, boundary region and negative region of a target concept X are respectively defined by

$$POS_{P_l}^{(\alpha,\beta)_l}(X) = \underline{P_l}^{(\alpha,\beta)_l}(X),$$
$$BND_{P_l}^{(\alpha,\beta)_l}(X) = \overline{P_l}^{(\alpha,\beta)_l}(X) - \underline{P_l}^{(\alpha,\beta)_l}(X),$$
$$NEG_{P_l}^{(\alpha,\beta)_l}(X) = U - \underline{P_l}^{(\alpha,\beta)_l}(X).$$

In order to describe the recursive relation between two dynamic granulation orders P_l and P_{l+1}, the following principle is given.

Theorem 2. *Let $S = (U, At = C \cup D)$ be a decision table, $X \subseteq U$, and $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l = 1, 2, \ldots, n$. Then, for a given $P_l = \{A_1, A_2, \ldots, A_l\}$, we have*

$$POS_{P_{l+1}}^{U_{(\alpha,\beta)_{l+1}}}(X) = POS_{P_l}^{U_{(\alpha,\beta)_l}}(X) \cup POS_{A_{l+1}}^{U_{l+1(\alpha_{l+1},\beta_{l+1})}}(X_{l+1}), \tag{5}$$

where $X_1 = X, U_{l+1} = U - POS_{P_l}^{(\alpha,\beta)}(X)$ and $X_{l+1} = X - POS_{P_l}^{(\alpha,\beta)}(X)$.

Here, $POS_{P_l}^{U_{(\alpha,\beta)_l}}(X)$ indicates the positive region of X on the universe U under the dynamic granulation P_l, $POS_{A_{l+1}}^{U_{l+1(\alpha_{l+1},\beta_{l+1})}}(X_{l+1})$ denotes the positive region of X_{l+1} on the universe U_{l+1} with respect to the equivalence relation A_{l+1}.

This theorem can be used to dynamically compute the positive region of a target concept (or decision) in the decision-theoretic rough set, which can largely save computational time. The recursive relation can be understood by the following example.

Example 3. Continued by Example 2. We can obtain $POS_{P_1}^{U_{(\alpha,\beta)_1}}(X) = \{x_5, x_6, x_8\}$. Let $U_1 = U$ and $X_1 = X$. Then, the universe is updated as

$$U_2 = U - POS_{P_1}^{U_{(\alpha,\beta)_1}}(X) = \{x_1, x_2, x_3, x_4, x_7, x_9, x_{10}\},$$

and the target concept X is updated as

$$X_2 = X - POS_{P_1}^{U_{(\alpha,\beta)_1}}(X) = \{x_2, x_3, x_{10}\}.$$

Through computing, one has that

$$POS_{A_2}^{U_{2(\alpha_2,\beta_2)}}(X_2) = \{x_2, x_{10}\},$$

and

$$POS_{P_2}^{U_{(\alpha,\beta)_2}}(X) = \{x_2, x_5, x_6, x_8, x_{10}\} = POS_{P_1}^{U_{(\alpha,\beta)_1}}(X) \cup POS_{A_2}^{U_{2(\alpha_2,\beta_2)}}(X_2).$$

That is to say, the positive regions of the target concept X under the dynamic granulation orders satisfy the above recursive principle.

3.4. Computing approximation of a target concept under dynamic granulation orders

In this part, we construct a computing lower approximation algorithm under a dynamic granulation order in DG-DTRS. Furthermore, we extend the proposed set-approximation approach to a decision partition.

The detailed algorithm for computing a lower approximation of a target concept in DG-DTRS is formally described as follows.

Algorithm 1. Computing the lower approximation of a target concept under a dynamic granulation order (DGLAC).

Input: A decision table $S = (U, AT = C \cup D)$, a target concept set $X \subseteq U$, and a family of attribute sets $P = \{A_1, A_2, \ldots, A_n\}$ with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n (A_l \in 2^C, l \leqslant n)$. Given a dynamic granulation order $P_l = \{A_1, A_2, \ldots, A_l\}$, and the loss function $\lambda_{ij}^k (k = l, 2, \ldots, l)$ with respect to P_l.

Output: The $P_l^{(\alpha,\beta)_l}$-lower approximation L of X.

```
k ← 1, X_1 ← X, U_1 ← U, L ← φ and P_1 = {A_1}
while k ⩽ l and X_k ≠ φ do
   Compute (α_k, β_k) with respect to λ_ij^k
   {compute threshold parameters for each granulation}
   for all x ∈ X_{A_k} do
      compute [x]_{A_k} of x
      {compute equivalence class of x on universe U_k}
      if P(X_k | [x]_{A_k}) ⩾ α_k then
         L ← L ∪ x
         i ← i + 1
      end if
   end for
   X_{k+1} = X_k − POS_{A_k}^{(α_k,β_k)}(X_k), U_{k+1} = U_k − POS_{A_k}^{(α_k,β_k)}(X_k)
   k = k + 1
   P_k ← {A_1, A_2, ..., A_k}
end while
return L
```

The algorithm shows the process of computing a lower approximation under a given dynamic granulation order. In fact, under dynamic granulation worlds, a target concept or decision can be gradually approximated by a dynamic granulation order from coarse to fine. This means that a suitable dynamic granulation order can be chosen for a target concept approximation according to the practical requirements, instead of strictly satisfying the stopping criterion in algorithm. Here, we consider $\eta_{(\alpha,\beta)}(P, X) = \frac{|POS_P^{(\alpha,\beta)}|}{|X|}$ as the precision of the positive region of $X \subseteq U$ with respect to the granulation order P, which describes the ability of granulation orders for dynamically approximating the target concept (or decision). Therefore, in the above algorithm, we also can set a threshold parameter to control the stop of the algorithm.

The algorithm is easily illustrated by the following example.

Example 4. Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be a universe, U/A_1, U/A_2 two partitions on U, where $U/A_1 = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7, x_8\}, \{x_9, x_{10}\}\}$ and $U/A_2 = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}, \{x_9\}, \{x_{10}\}\}$.

Obviously, $A_1 \succcurlyeq A_2$ holds. Thus, we can construct two granulation orders $P_1 = \{A_1\}$ and $P_2 = \{A_1, A_2\}$.

Given a target concept $X = \{x_1, x_3, x_5, x_6, x_7, x_9\}$. For simplicity, we suppose $(\alpha, \beta)_2 = \{(0.7, 0.2), (0.8, 0.2)\}$ by a dynamic granulation order P_2. The family of threshold parameters can be computed from the various loss functions.

According to Algorithm 1, we compute the lower approximation of X by the granulation orders.

1. Let $U_1 = U$, $X_1 = X$, $P_1 = \{A_1\}$. For each $x \in X_1$, by computing $P(X|[x]_{A_1})$ of $x, (\alpha_1, \beta_1) = (0.7, 0.2)$, we can easily obtain $\underline{P_1}^{(0.7,0.2)}(X) = \{x_5, x_6, x_7\}$.
2. Updating the universe $U_2 = U_1 - \underline{A_1}^{(0.7,0.2)}(X) = \{x_1, x_2, x_3, x_4, x_8, x_9, x_{10}\}$, $X_2 = X_1 - \underline{P_1}^{(0.7,0.2)}(X) = \{x_1, x_3, x_9\}$, $P_2 = \{A_1, A_2\}$ and $(\alpha_2, \beta_2) = (0.8, 0.2)$. For each $x \in X_2$, by computing $[x]_{A_2}$ of x in universe U_2, we have
$$[x_1]_{A_2} = \{x_1\}, \quad [x_3]_{A_2} = \{x_3, x_4\}, \quad [x_9]_{A_2} = \{x_9\}.$$
Then, by computing $P(X_2|[x]_{A_2})$ of x, we can easily obtain
$$\underline{P_2}^{(\alpha,\beta)_2}(X) = \{x_5, x_6, x_7\} \cup \{x_1, x_9\} = \{x_1, x_5, x_6, x_7, x_9\}.$$

Similar to the decision-theoretic rough set model, we can extend the concept of probabilistic approximations and regions of a single decision to a partition U/D. For simplicity, we assume that the same loss functions are used for all decisions. The detailed definition is as follows.

Definition 3. Let $S = (U, At = C \cup D)$ be a decision table, $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l = 1, 2, \ldots, n,$ and $U/D = \{D_1, D_2, \ldots, D_m\}$ a decision partition on U. Then, the (α, β)-Lower approximation and the (α, β)-upper approximation of D related to P_l are defined as

$$\underline{P_l}^{(\alpha,\beta)_l}(D) = \{\underline{P_l}^{(\alpha,\beta)_l}(D_1), \underline{P_l}^{(\alpha,\beta)_l}(D_2), \ldots, \underline{P_l}^{(\alpha,\beta)_l}(D_m)\},$$

$$\overline{P_l}^{(\alpha,\beta)_l}(D) = \{\overline{P_l}^{(\alpha,\beta)_l}(D_1), \overline{P_l}^{(\alpha,\beta)_l}(D_2), \ldots, \overline{P_l}^{(\alpha,\beta)_l}(D_m)\}.$$

Correspondingly, the positive region, the boundary region and the negative region of the target decision D in the DG-DTRS model can be respectively represented as follows:

$$POS_{P_l}^{(\alpha,\beta)_l}(D) = \bigcup_{i=1}^{m} POS_{P_l}^{(\alpha,\beta)_l}(D_i) = \bigcup_{i=1}^{m} \underline{P_l}^{(\alpha,\beta)_l}(D_i),$$

$$BND_{P_l}^{(\alpha,\beta)_l}(D) = \bigcup_{i=1}^{m} BND_{P_l}^{(\alpha,\beta)_l}(D_i),$$

$$NEG_{P_l}^{(\alpha,\beta)_l}(A) = U - POS_{P_l}^{(\alpha,\beta)_l}(D) \cup BND_{P_l}^{(\alpha,\beta)_l}(D).$$

Please cite this article in press as: Y. Sang et al., Decision-theoretic rough sets under dynamic granulation, Knowl. Based Syst. (2015), http://dx.doi.org/10.1016/j.knosys.2015.08.001

In what follows, we can extend the monotonicity of a single target concept to a decision partition $U/D = \{D_1, D_2, \ldots, D_m\}$ of the universe, which is shown in the following theorem.

Theorem 3 (*Decision monotonicity*). *Let $S = (U, At = C \cup D)$ be a decision table, $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C$, and $U/D = \{D_1, D_2, \ldots, D_m\}$ a decision partition on U. Given $P_l = \{A_1, A_2, \ldots, A_l\}$, then for any P_l, $l = 1, 2, \ldots, n$, we have*

$$POS_{P_1}^{(\alpha,\beta)_1}(D) \subseteq POS_{P_2}^{(\alpha,\beta)_2}(D) \subseteq \cdots \subseteq POS_{P_l}^{(\alpha,\beta)_l}(D).$$

In the following, we want to illustrate that the positive region of a target decision can also be recursively computed on the gradually reduced universe by the below theorem.

Theorem 4. *Let $S = (U, At = C \cup D)$ be a decision table, $P = \{A_1, A_2, \ldots, A_n\}$ a family of attribute sets with $A_1 \succcurlyeq A_2 \succcurlyeq \cdots \succcurlyeq A_n, A_l \in 2^C, l = 1, 2, \ldots, n$, and $U/D = \{D_1, D_2, \ldots, D_m\}$ a decision partition on U. Then, given $P_l = \{A_1, A_2, \ldots, A_l\}$, we have*

$$POS_{P_{l+1}}^{U_{(\alpha,\beta)_{l+1}}}(D) = POS_{P_l}^{U_{(\alpha,\beta)_l}}(D) \cup POS_{A_{l+1}}^{U_{l+1}(\alpha_{l+1},\beta_{l+1})}(D),$$

where $U_1 = U$ and $U_{l+1} = U - POS_{P_l}^{U_{(\alpha,\beta)}}(D)$.

The recursive computation principle is explained by the following example.

Example 5. Let $S = (U, C \cup D)$ be a decision table, where $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be a universe, $C = \{a_1, a_2\}$, $U/D = \{\{x_1, x_3, x_5, x_6, x_7, x_9\}, \{x_2, x_4, x_8, x_{10}\}\}$, $U/a_1 = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7, x_8\}, \{x_9, x_{10}\}\}$, and $U/C = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}, \{x_9\}, \{x_{10}\}\}$.

Obviously, $\{a_1\} \succcurlyeq C$ holds. Thus, we can construct two granulation orders $P_1 = \{a_1\}$ and $P_2 = \{\{a_1\}, C\}$.

Suppose $(\alpha, \beta)_2 = \{(0.7, 0.2), (0.8, 0.2)\}$. From Algorithm 1, one has the lower approximation of D. Then it follows

$$POS_{P_1}^{U_{(\alpha,\beta)_1}}(D) = \underline{P_1}^{(\alpha,\beta)_1}(D_1) \cup \underline{P_1}^{(\alpha,\beta)_1}(D_2) = \{x_5, x_6, x_7\}.$$

Let $U_1 = U$, and update the universe

$$U_2 = U_1 - POS_{P_1}^{U_{(\alpha,\beta)_1}}(D) = \{x_1, x_2, x_3, x_4, x_8, x_9, x_{10}\}.$$

Through computing, we have

$$POS_C^{U_2(\alpha_2,\beta_2)}(D) = \{x_1, x_2, x_8, x_9, x_{10}\} \text{ and } POS_{P_2}^{U_{(\alpha,\beta)_2}}(D) = \{x_1, x_2, x_5, x_6, x_7, x_9, x_{10}\}.$$

Hence, $POS_{P_2}^{U_{(\alpha,\beta)_2}}(D) = POS_{P_1}^{U_{(\alpha,\beta)_1}}(D) \cup POS_C^{U_2(\alpha_2,\beta_2)}(D)$.

That is to say, the target decision D can be recursively approximated by using dynamic granulation orders P_1 and P_2 on the gradually reduced universe.

4. Conclusions and future studies

As an important model within rough set theory, the decision-theoretic rough sets have been largely enriched. However, the non-monotonicity of its positive region may lead to an overlapping problem for attribute reduction. To solve this problem, in this paper we have proposed a new decision-theoretic rough set model based on the local rough set and the dynamic granulation principle, called a decision-theoretic rough set under dynamic granulation (DG-DTRS) which satisfies the monotonicity of the positive region of a target concept (or decision). To achieve the risk minimization under each granulation, based on the Bayesian decision procedure, we have also given an approach to update the required parameters α and β in the proposed model for each granulation. This dynamic decision-theoretic rough set model can ensure the monotonicity of positive region and the local risk minimization as information granulation becomes finer besides providing sound semantic interpretation. Hence, the modified version with several better properties can be regarded as an important improvement of the original decision-theoretic rough set model.

Acknowledgments

This work was supported by National Natural Science Fund of China (Nos. U1435212, 61432011, 61322211, 61202018), the Research Fund for the Doctoral Program of Higher Education (No. 20121401110013), Program for New Century Excellent Talents in University (No. NCET-12-1031).

References

[1] J. Błaszczyński, S. Greco, R. Słowiński, M. Szelag, Monotonic variable consistency rough set approaches, Int. J. Approx. Reason. 50 (7) (2009) 979–999.
[2] J.H. Dai, Q. Xu, Approximations and uncertainty measures in incomplete information systems, Inf. Sci. 198 (2012) 62–80.
[3] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artif. Intell. 106 (1998) 109–137.
[4] S. Greco, B. Matarazzo, R. Słowiński, Parameterized rough set model using rough membership and Bayesian confirmation measures, Int. J. Approx. Reason. 49 (2) (2008) 285–300.
[5] J.P. Herbert, J.T. Yao, Game-theoretic rough sets, Fundam. Inform. 108 (3–4) (2011) 267–286.
[6] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Trans. Fuzzy Syst. 15 (1) (2007) 73–89.
[7] G. Jeon, D. Kim, J. Jeong, Rough sets attributes reduction based expert system in interlaced video sequences, IEEE Trans. Consum. Electron. 52 (4) (2006) 1348–1355.
[8] X.Y. Jia, W.H. Liao, Z.M. Tang, L. Shang, Minimum cost attribute reduction in decision-theoretic rough set models, Inf. Sci. 219 (10) (2013) 151–167.
[9] X.Y. Jia, Z.M. Tang, W.H. Liao, L. Shang, On an optimization representation of decision-theoretic rough set model, Int. J. Approx. Reason. 55 (1) (2014) 156–166.
[10] X.Y. Jia, L. Shang, B. Zhou, Y.Y. Yao, Generalized attribute reduct in rough set theory, Knowl.-Based Syst. doi:http://dx.doi.org/10.1016/j.knosys.2015.05.017.
[11] S.Y. Li, T.R. Li, J. H, Update of approximations in composite information systems, Knowl.-Based Syst. 83 (2015) 138–148.
[12] J.H. Li, C.L. Mei, Y.J. Lv, Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction, Int. J. Approx. Reason. 54 (1) (2013) 149–165.
[13] J.H. Li, C.L. Mei, W.H. Xu, Y.H. Qian, Concept learning via granular computing: a cognitive viewpoint, Inf. Sci. 298 (2015) 447–467.
[14] J.H. Li, Y. Ren, C.L. Mei, Y.H. Qian, X.B. Yang, A comparative study of multigranulation rough sets and concept lattices via rule acquisition, Knowl.-Based Syst. doi:http://dx.doi.org/10.1016/j.knosys.2015.07.024.
[15] H.X. Li, X.Z. Zhou, J.B. Zhao, D. Liu, Attribute reduction in decision-theoretic rough set model: a further investigation, in: LNCS, vol. 6954, 2011, pp. 466–475.
[16] J.Y. Liang, C.Y. Dang, K.S. Chin, C.M. Yam Richard, A new method for measuring uncertainty and fuzziness in rough set theory, Int. J. Gen Syst 31 (4) (2002) 331–342.
[17] J.Y. Liang, F. Wang, C.Y. Dang, Y.H. Qian, An efficient rough feature selection algorithm with a multi-granulation view, Int. J. Approx. Reason. 53 (2012) 912–926.
[18] J.Y. Liang, J.H. Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Inf. Sci. 179 (2009) 458–470.
[19] T.Y. Lin, Granular computing on binary relations I: data mining and neighborhood systems, in: A. Skowron, L. Polkowski (Eds.), Rough Sets in Knowledge Discovery, Physica-Verlag, 1998, pp. 107–121.
[20] T.Y. Lin, Granular computing on binary relations II: rough set representations and belief functions, in: A. Skowron, L. Polkowski (Eds.), Rough Sets in Knowledge Discovery, Physica-Verlag, 1998, pp. 121–140.
[21] D. Liu, Y.Y. Yao, T.R. Li, Three-way investment decisions with decision-theoretic rough sets, Int. J. Comput. Intell. Syst. 4 (1) (2011) 66–74.
[22] X. Ma, G.Y. Wang, H. Yu, T.R. Li, Decision region distribution preservation reduction in decision-theoretic rough set model, Inf. Sci. 278 (10) (2014) 614–640.
[23] Z. Pawlak, Rough sets, Int. J. Comput. Inform. Sci. 11 (5) (1982) 341–356.
[24] Z. Pawlak, S.K.M. Wong, W. Ziarko, Rough sets: probabilistic versus deterministic approach, Int. J. Man Mach. Stud. 29 (1) (1988) 81–95.

[25] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.
[26] Z. Pawlak, A. Skowron, Rough membership functions, in: R.R. Yaeger, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster Shafer Theory of Evidence, John Wiley Sons, 1994, pp. 251–271.
[27] Y.H. Qian, J.Y. Liang, Rough set method based on multi-granulations, in: Proceedings of 5th IEEE Conference on Cognitive Informatics, vol. 1, 2006, pp. 297–304.
[28] Y.H. Qian, J.Y. Liang, W. pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artif. Intell. 174 (2010) 597–618.
[29] Y.H. Qian, J.Y. Liang, C.Y. Dang, Incomplete multigranulation rough set, IEEE Trans. Syst. Man Cybernet.: Part A 40 (2) (2010) 420–431.
[30] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: a mulit-granulation rough set, Inf. Sci. 180 (2010) 949–970.
[31] Y.H. Qian, S.Y. Li, J.Y. Liang, Z.Z. Shi, F. Wang, Pessimistic rough set based decisions: a multigranulation fusion strategy, Inf. Sci. 264 (20) (2014) 196–210.
[32] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 225–237.
[33] Y.H. Qian, et al., Local rough set: a solution to rough data analysis in big data, Pattern Recognit., submitted for publication.
[34] M. Rebolledo, Rough intervals-enhancing intervals for qualitative modeling of technical systems, Artif. Intell. 170 (8–9) (2006) 667–685.
[35] D. Ślęzak, W. Ziarko, The investigation of the Bayesian rough set model, Int. J. Approx. Reason. 40 (1–2) (2005) 81–91.
[36] W.H. Xu, Q.R. Wang, X.T. Zhang, Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space, Int. J. Fuzzy Syst. 13 (4) (2011) 246–259.
[37] X.P. Yang, H.G. Song, T.J. Li, Decision making in incomplete information systems based on decision-theoretic rough sets, in: Proceedings of International Conference on Rough Sets and Knowledge Technology, Banff, Canada, LNCS, vol. 6954, 2011, pp. 495–503.
[38] X.P. Yang, J.T. Yao, Modelling multi-agent three-way decisions with decision-theoretic rough sets, Inform. Sci. 115 (2–3) (2012) 157–171.
[39] X.B. Yang, X.N. Song, H.L. Dou, J.Y. Yang, Multi-granulation rough set: from crisp to fuzzy case, Ann. Fuzzy Math. Inform. 1 (1) (2011) 55–70.
[40] X.B. Yang, Y. Qi, H.L. Yu, X.N. Song, J.Y. Yang, Updating multigranulation rough approximations with increasing of granular structures, Knowl.-Based Syst. 64 (2014) 59–69.
[41] J.T. Yao, A ten-year review of granular computing, in: Proceedings of IEEE GrC2007, Sillicon Valley, CA, USA, 2007, pp. 734–739.
[42] J.T. Yao, Y.Y. Yao, W. Ziarko, Probabilistic rough sets: approximations, decision-makings, and applications, Int. J. Approx. Reason. 49 (2) (2008) 253–254.
[43] Y.Y. Yao, S.K.M. Wong, P. Lingras, A decision-theoretic rough set model, in: Z. W. Ras, M. Zemankova, M.L. Emrich (Eds.), Methodologies for Intelligent Systems, vol. 5, North-Holland, New York, 1990, pp. 17–24.
[44] Y.Y. Yao, S.K.M. Wong, A decision theoretic framework for approximating concepts, Int. J. Man Mach. Stud. 37 (6) (1992) 793–809.
[45] Y.Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set model, Inf. Sci. 178 (17) (2008) 3356–3373.
[46] Y.Y. Yao, Probabilistic rough set approximations, Int. J. Approx. Reason. 49 (2) (2008) 255–271.
[47] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. 180 (3) (2010) 341–353.
[48] Y.Y. Yao, The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. 181 (6) (2011) 1080–1096.
[49] H. Yu, S.S. Chu, D.C. Yang, Autonomous knowledge-oriented clustering using decision-theoretic rough set theory, Lect. Notes Comput. Sci. 6401 (2010) 687–694.
[50] H. Yu, C. Zhang, G.Y. Wang, A tree-based incremental overlapping clustering method using the three-way decision theory, Knowl.-Based Syst. doi: http://dx.doi.org/10.1016/j.knosys.2015.05.028.
[51] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets Syst. 90 (1997) 111–127.
[52] X.Y. Zhang, D.Q. Miao, Region-based quantitative and hierarchical attribute reduction in the two-category decision theoretic rough set model, Knowl.-Based Syst. 71 (2014) 146–161.
[53] H.R. Zhang, F. Min, Three-way recommender systems based on random forests, Knowl.-Based Syst. doi: http://dx.doi.org/10.1016/j.knosys.2015.06.019.
[54] M. Zhang, Z.M. Tang, W.Y. Xu, X.B. Yang, A variable multigranulation rough sets approach, in: Proceedings of the 7th International Conference on Intelligent Computing: Bio-inspired Computing and Applications, 2011, pp. 315–322.
[55] Y. Zhao, S.K.M. Wong, Y.Y. Yao, A note on attribute reduction in the decision-theoretic rough set model, in: Proceedings of RSKTC2008, LNCS, vol. 5306, 2008, pp. 61–70.
[56] B. Zhou, A new formulation of multi-category decision-theoretic rough sets, in: Proceeding of 6th International Conference on Rough Set and Knowledge Technology, LNAI, vol. 6954, 2011, pp. 514–522.
[57] X.Z. Zhou, H.X. Li, A multi-view decision model based on decision-theoretic rough set, in: Proceedings of RSKT2009, LNCS, vol. 5589, 2009, pp. 650–657.
[58] B. Zhou, Y.Y. Yao, J.G. Luo, A three-way decision approach to email spam filtering, in: Proceedings of the 23th Canadian Conference on Artificial Intelligence, LNCS, vol. 6085, 2010, pp. 28–39.
[59] W. Ziarko, Variable precision rough set model, J. Comput. Syst. Sci. 46 (1993) 39–59.
[60] W. Ziarko, Probabilistic approach to rough sets, Int. J. Approx. Reason. 49 (2) (2008) 272–284.

Please cite this article in press as: Y. Sang et al., Decision-theoretic rough sets under dynamic granulation, Knowl. Based Syst. (2015), http://dx.doi.org/10.1016/j.knosys.2015.08.001