Knowledge-Based Systems 91 (2016) 84–92

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

A R T I C L E I N F O

*Article history:*
Received 23 January 2015
Revised 31 July 2015
Accepted 1 August 2015
Available online 6 August 2015

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

A B S T R A C T

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 $\alpha$ and $\beta$ 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

* 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/

since 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 $\alpha$ and $\beta$ 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,

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

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:

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

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

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

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