ARTICLE IN PRESS IJA:7614

International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Fuzzy probabilistic rough set model on two universes and its applications

Hai-Long Yang [a,b,c,*], Xiuwu Liao [a], Shouyang Wang [c], Jue Wang [c]

[a] *The School of Management, Xi'an Jiaotong University, Xi'an 710049, PR China*
[b] *College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, PR China*
[c] *Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China*

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form 19 March 2013
Accepted 1 May 2013
Available online xxxx

Keywords:
Rough sets
Fuzzy relations
Fuzzy probabilistic approximation spaces
Fuzzy probabilistic rough sets

ABSTRACT

The classical probabilistic rough set model is established based on a crisp binary relation. As a generalization of crisp binary relation, fuzzy relation makes descriptions of the objective world more realistic, practical, and accurate in some cases. Thus probabilistic rough set model based on a crisp binary relation limits its application domain. In this paper, based on a fuzzy relation, we propose a fuzzy probabilistic rough set model on two universes. Meanwhile, the concepts of the inverse lower and upper approximation operators are presented. We also study some properties of these approximation operators. Finally, a numerical example of the clinical diagnosis systems is applied to illustrate the validity of the proposed model. And we compare the proposed model with other models to show the superiority of the proposed model.

1. Introduction

Rough set theory, proposed by Pawlak [19,20] in 1982 is a new mathematical approach for representing and processing information in database. A lack of consideration of the degree of overlap between an equivalence class and the set motivates many researchers investigated probabilistic generalizations of the theory [22,30–32,36].

Probabilistic rough set approximations can be formulated based on the notions of rough membership functions [21] and rough inclusion [18]. Both notions can be interpreted in terms of conditional probabilities or a posteriori probabilities. Three probabilistic models have been proposed and studied intensively. They are the decision-theoretic rough set model [1,31,32,34,35,39,45], the variable precision rough set model [5–7,44], and the Bayesian rough set model [3,24,25,46]. In [36], Yao brought together the three models' studies on probabilistic rough set approximations in a unified and comprehensive framework. Threshold values, known as parameters, are applied to a rough membership function or a rough inclusion to obtain probabilistic or parameterized approximations. The probabilistic rough set model treats the required parameters as a primitive notion. The interpretation and the process of determining the parameters are based on rather intuitive arguments and left to empirical studies [36]. There is a lack of theoretical and systematic studies and justifications on the choices of the threshold parameters. A solution to this problem was reported earlier in a decision-theoretic framework for probabilistic rough set approximations (see [32,35]), based on the well-established Bayesian decision procedure for classification. Within the decision-theoretic framework, the required threshold values can be easily interpreted and calculated based on more concrete notions, such as costs and risks. Liu et al. [12] studied probabilistic model criteria with decision-theoretic rough sets. Li and Yang [9] studied an axiomatic characterization of probabilistic rough sets. Yao [37] provided an analysis of three-way decision rules in the classical rough set model and the decision-theoretic rough set model. The connections established

* Corresponding author at: College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China.
E-mail address: yanghailong@snnu.edu.cn (H.-L. Yang).

0888-613X/$ – see front matter
http://dx.doi.org/10.1016/j.ijar.2013.05.001

Please cite this article in press as: H.-L. Yang et al., Fuzzy probabilistic rough set model on two universes and its applications, International Journal of Approximate Reasoning (2013), http://dx.doi.org/10.1016/j.ijar.2013.05.001

between the levels of tolerance for errors and costs of incorrect decisions make the rough set theory practical in applications. Yao [38] also compared probabilistic three-way decisions, probabilistic two-way decisions, and qualitative three-way decisions of the standard rough set model. It is shown that, under certain conditions when considering the costs of different types of miss-classifications, probabilistic three-way decisions are superior to the other two. In a probabilistic rough set model, the positive, negative and boundary regions are associated with classification errors or uncertainty. The uncertainty is controlled by a pair of thresholds defining the three regions. The problem of searching for optimal thresholds can be formulated as the minimization of uncertainty induced by the three regions. In [13–15], Liu et al. applied decision-theoretic rough sets to three-way investment decisions, multiple-category classifications, and three-way government decisions. By using Shannon entropy as a measure of uncertainty, Deng and Yao [2] presented an information-theoretic approach to the interpretation and determination of thresholds.

Most existing research about rough set theory is based on one universe. But in reality, rough set models based on one universe may limit the applications of rough set theory. Two or multi-universes can describe the reality more reasonably and effectively. Nowadays, more and more researchers have realized it and some efforts have been made based on two different universes [8,10,11,16,23,26,27,40–43]. For example, Pei and Xu [23] studied rough set models on two universes and revealed some new properties. Li and Zhang [8] studied rough fuzzy approximations on two universes of discourse. Liu et al. [11] studied graded rough set model based on two universes and its properties. Shen and Wang [26] studied variable precision rough set model on two universes and its properties. Sun and Ma [27] studied fuzzy rough set model on two different universes and its application by using the cut-relation of a fuzzy relation. Yan and Zheng [40] proposed the rough set model on dual-universes and studied its properties using character function and relation matrix. Yang et al. [42] fully studied transformation of bipolar fuzzy rough set models and its applications. In [16], Ma and Sun generalized the classical probabilistic rough set on one universe proposed in [32] to two-universes case. They discussed the properties of the probabilistic rough set on two universes in detail. The probabilistic rough set on two universes can be applied to the disease diagnosis decision, the personalized marketing, spam filtering, etc. It is worth pointing out that the probabilistic rough set model discussed in [16] is established based on a crisp binary relation. However, in the practical applications, the information we face may be fuzzy. Thus, the probabilistic rough set model based on a crisp binary relation limits its application domain. So it is necessary to generalize the probabilistic rough set model given in [16] to fuzzy case. In the present paper, we will study the fuzzy probabilistic rough set model on two universes and apply it to clinical diagnosis systems [29].

The remainder of the paper is structured as follows. In the next section, we introduce some basic definitions and properties which will be used in this paper. In Section 3, based on a fuzzy relation, we introduce the notion of the fuzzy probabilistic rough set model on two universes. Meanwhile, the concepts of the inverse lower and upper approximation operators are presented. Then some interesting properties and related results are also established. Finally, an example is applied to illustrate the validity of the proposed model. And we compare the proposed model with other models to show the superiority of the proposed model. The last section summarizes the conclusions and presents some topics for future research.

2. Preliminaries

In this section, we recall some basic notions which will be used in this paper.

2.1. Pawlak rough sets and generalized rough sets based on binary relation

Let U be a non-empty finite universe, R be an equivalence relation on U. We use U/R to denote the family of all equivalence classes of R (or classifications of U), and $[x]_R$ to denote an equivalence class of R containing the element $x \in U$. The pair (U, R) is called an approximation space. For any $X \subseteq U$, we can define the lower and upper approximations of X [19,20] as follows:

$$\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\}, \qquad \overline{R}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}.$$

The pair $(\underline{R}X, \overline{R}X)$ is referred to as the rough set of X. The rough set $(\underline{R}X, \overline{R}X)$ gives rise to a description of X under the present knowledge, i.e., the classification of U.

Furthermore, the positive region, negative region, and boundary region of X about the approximation space (U, R) are defined as follows, respectively,

$$\mathrm{pos}(X) = \underline{R}(X), \qquad \mathrm{neg}(X) = \sim\overline{R}(X), \qquad \mathrm{bn}(X) = \overline{R}(X) - \underline{R}(X),$$

where $\sim X$ stands for the complementation of the set X.

Definition 2.1 *(Rough set based on a relation).* (See [33].) Suppose R is a binary relation on a universe U. A pair of approximation operators, $L(R), H(R) : P(U) \longrightarrow P(U)$, are defined by

Please cite this article in press as: H.-L. Yang et al., Fuzzy probabilistic rough set model on two universes and its applications, International Journal of Approximate Reasoning (2013), http://dx.doi.org/10.1016/j.ijar.2013.05.001