Knowledge-Based Systems 33 (2012) 65–72

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

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

Graded rough set model based on two universes and its properties

Caihui Liu [a,b,*], Duoqian Miao [a], Nan Zhang [a]

[a] Department of Computer Science and Technology, Tongji University, 201804 Shanghai, China
[b] Department of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou, 341000 Jiangxi, China

ARTICLE INFO

Article history:
Received 23 October 2011
Received in revised form 20 February 2012
Accepted 20 February 2012
Available online 9 March 2012

Keywords:
Rough sets
Two universes of discourse
Grade
Approximation operators
Relation matrix

ABSTRACT

In recent years, much attention has been given to the rough set models based on two universes of discourse and different kinds of rough set models on two universes have been developed from different points of view. In this paper, a novel model, i.e., the graded rough set model on two distinct but related universes (GRSTU) is proposed from the absolute quantitative point of view. We study the basic properties of approximation operators in GRSTU, and introduce a relation matrix based algorithm to compute the lower and upper approximations of a set of objects in GRSTU. Furthermore, the relationships between classical rough set model and GRSTU are discussed and some conclusions related to the GRSTU are given. Finally, several examples are employed to demonstrate the conceptual arguments of GRSTU, and an application of GRSTU is also illuminated in details.

1. Introduction

Rough set theory, originally proposed by Pawlak [1,2] as an extension of set theory, is an effective approach to dealing with imprecise, uncertain and incomplete information. It has been successfully used in many research areas, such as pattern recognition, machine learning, knowledge acquisition and data mining [3–7,32–35,38,39,41].

As we know, Pawlak's rough set model has a basic hypothesis, that is, whether an object belongs to a class or not is completely certain. However, in practice, allowing some extent of uncertainty in the classification process may lead to a deeper understanding and a better utilization of the data being analyzed. In order to deal with the uncertainty in such cases, a lot of models have been proposed. For example, based on the Bayesian decision procedure with minimum cost (risk), Yao [8,9] proposed a decision-theoretic rough set model (DTRSM) which brings new insights into the probabilistic approaches to rough set model. DTRSM not only has good semantic interpretation, but also be beneficial for rule acquisition in the applications involving cost and risk. And Yao and Lin [10,11] presented a graded rough set model (GRS) from the absolute quantitative point of view. Moreover, many other models have also been proposed, such as rough set models based on arbitrary binary relations [12–14,40], rough set models based on incomplete systems [15–17,43], covering rough sets [18–20], rough fuzzy sets and fuzzy rough sets [21,42,44], variable precision rough sets [22,23], etc. Through loosening the strict definition of the approximations in Pawlak's rough set model, these models enrich the application scope of rough set theory.

In the real world, we often face some situations in which making a decision is difficult. For example, in the process of identifying or determining the nature and cause of a disease, since a certain disease may simultaneously have several symptoms but the same symptom may be shared by diverse diseases, a doctor (or a decision-maker) often finds it is difficult to distinguish whether a person has suffered from the disease or not. In these kinds of situations, more than one universes of discourse are often involved. However, Pawlak's rough set model and its extensions mentioned above are all based on only one universe, therefore these models may be not suitable to deal with the above problem. Hence, it is meaningful to propose a rough set model based on two universes.

The generalization of rough set model from only one universe of discourse to the two distinct but related universes of discourse has attracted much attention [24–29,36]. Wong et al. [24] first proposed a rough set model on two universes from the viewpoint of compatibility. In [26,28,29], the applications and some interesting properties about the rough set model on two universes were discussed. Wu et al. [25] developed a general framework for the study of the fuzzy rough set models on two universes in which both constructive and axiomatic approaches were considered and surveyed. In [27], four types of rough fuzzy approximation operators on two universes have been proposed. Zhang et al. [36] studied the generalized interval-valued fuzzy rough sets on two universes of discourse.

* Corresponding author at: Department of Computer Science and Technology, Tongji University, 201804 Shanghai, China. Tel.: +86 21 69584157.

E-mail addresses: liu_caihui@163.com (C. Liu), miaoduoqian@163.com (D. Miao), zhangnan0851@163.com (N. Zhang).

0950-7051/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2012.02.012

Although the above models can effectively overcome the limitations of rough set models on one universe, they still lack the adaptability in solving uncertainty problems. To solve this problem, Shen et al. [30] proposed a variable precision rough set model on two universes from the relative quantitative point of view. Ma and Sun [37] introduced a probabilistic rough sets over two universes and used it to deal with the problem of Bayesian risk decision. In this paper, from the absolute quantitative point of view, we propose a graded rough set model defined on two distinct but related universes. Our model is not only an extension of the rough set model on two universes but also an extension of Pawlak's rough set model. To compare with Yao and Lin's graded rough set model, our model may be more appropriate to handle the problems where more than one universe is involved. Paralleling with Pawlak's rough set model, the basic properties of our model are discussed. Meanwhile, a relation matrix based algorithm for computing the lower and upper approximations in GRSTU is proposed.

The remainder of this paper is organized as follows. In the next section, we briefly introduce some notions using in Pawlak's rough set model, graded rough set model on one universe and rough set model on two universes. In Section 3, we define the lower and upper approximation operators in GRSTU and discuss the basic properties of GRSTU. In Section 4, two examples are employed to substantiate the conceptual arguments of GRSTU. An application of GRSTU is discussed in details in Section 5. Finally, Section 6 concludes the paper.

2. Preliminaries

In this section, we outline some basic concepts in rough sets and some current rough set models, such as Pawlak's rough set model [1], graded rough set model on one universe [10] and rough set model on two universes [24]. Throughout this paper, we suppose that the universe U or V is a finite non-empty set.

2.1. Pawlak's rough set model

Let U be a universe of discourse, for any binary relation R on U, we call R an equivalence relation on U, if.

(1) R is reflexive if for all $x \in U$, xRx;
(2) R is symmetric if for all $x,y \in U$, xRy implies yRx;
(3) R is transitive if for all $x,y,z \in U$, xRy and yRz implies xRz.

An equivalence relation is a reflexive, symmetric and transitive relation. The equivalence relation R partitions U into disjoint subsets (or equivalence classes). Let U/R denote the family of all equivalence classes of R. For every object $x \in U$, let $[x]_R$ denote the equivalence class of relation R that contains element x, called the equivalence class of x under relation R.

Let U be a universe of discourse, R an equivalence relation on U, for any $X \subseteq U$, one can describe X by a pair of lower and upper approximations defined as follows.

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

$\underline{R}(X)$ is called the lower approximation of X, which is the union of all the equivalence classes which contain in X, and $\overline{R}(X)$ is called the upper approximation of X, which is the union of all equivalence classes which have non-empty intersection with X. Then $(\underline{R}(X), \overline{R}(X))$ is called the rough sets of X. Accordingly, the positive, negative and boundary regions of X on the approximation space (U,R) can be defined as follows: $pos(X) = \underline{R}(X)$, $neg(X) = \sim \overline{R}(X), bnd(X) = \overline{R}(X) - \underline{R}(X)$, where $\sim$ stands for the complement of a set.

2.2. Generalized rough set operators

The Pawlak rough set model may be extended by using an arbitrary binary relation.

Let U be a universe of discourse and R a binary relation on U, the following two operators: $r(x) = \{y \in U | xRy\}, l(x) = \{y \in U | yRx\}$ are called the successor and predecessor neighborhood operator, respectively.

Definition 1 ([12]). Let U be a universe of discourse and R a binary relation on U. For any $X \subseteq U$, its lower and upper approximations based on the successor neighborhood operator are respectively defined as follows:

$$\underline{R}(X) = \{x \in U | r(x) \subseteq X\}$$
$$\overline{R}(X) = \{x \in U | r(x) \cap X \neq \emptyset\}$$

Analogously, for any $X \subseteq U$, one can define the lower and upper approximations based on the predecessor neighborhood operator.

In the remainder of this paper, we shall only take the case of the successor neighborhood operator into consideration.

2.3. Yao and Lin's graded rough set model on one universe [10]

Let U be a universe and R a binary relation on U, $n \in N$, where N is the set of natural numbers. For any subset $A \subseteq U$, the lower and upper approximations of A with respect to n (denoted by $\underline{apr}_n(A)$ and $\overline{apr}_n(A)$, respectively) are defined as follows.

$$\begin{aligned} \underline{apr}_n(A) &= \{x \in U || r(x)| - |r(x) \cap A| \leqslant n\} \\ &= \{x \in U || r(x) - A| \leqslant n\} \\ \overline{apr}_n(A) &= \{x \in U || r(x) \cap A| > n\} \end{aligned}$$

where $|r(x)|$ denotes the cardinality of set $r(x)$.

An element of U belongs to $\underline{apr}_n(A)$ if at most n of its R-related elements do not belong to A, and belongs to $\overline{apr}_n(A)$ if more than n of its R-related elements belong to A.

2.4. Rough set model on two universes

Next, we shall review some basic concepts and properties of the rough set model on two universes. Detailed description of the model can be found in [12,26,28].

The above model can be generalized to the case of two universes.

Definition 2 ([12,28]). Let U and V be two universes of discourse and R a binary relation from U to V, i.e. $R \subseteq U \times V$. The ordered triple (U,V,R) is called a two-universe approximation space. For any $Y \subseteq V$, the lower and upper approximations of Y can be defined as follows.

$$\underline{R}(Y) = \{x \in U | r(x) \subseteq Y\}$$
$$\overline{R}(Y) = \{x \in U | r(x) \cap Y \neq \emptyset\}$$

$\underline{R}(Y)$ is called the lower approximation of Y and $\overline{R}(Y)$ the upper approximation of Y. $(\underline{R}\ (Y), \overline{R}(Y))$ is called the rough sets of Y. Accordingly, the positive, negative and boundary regions of Y over the approximation space (U,V,R) are defined as follows: $Pos(Y) = \underline{R}(Y)$, $Neg(Y) = \sim \overline{R}(Y)$, $Bnd(Y) = \overline{R}(Y) - \underline{R}(Y)$.

Proposition 1. *Given a two-universe approximation space (U,V,R), for any Y, Y_1, $Y_2 \subseteq V$, the approximation operators given in Definition 2 have the following properties:*

(1) $\underline{R}(Y) = \sim \overline{R}(\sim Y), \overline{R}(Y) = \sim \underline{R}(\sim Y)$
(2) $\underline{R}(V) = \overline{R}(V) = U, \underline{R}(\emptyset) = \overline{R}(\emptyset) = \emptyset$