Knowledge-Based Systems 91 (2016) 93–101

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Variable precision multigranulation decision-theoretic fuzzy rough sets

Tao Feng [a,*], Ju-Sheng Mi [b]

[a] School of Sciences, Hebei University of Science and Technology, Shijiazhuang 050018, China
[b] College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang 050024, China

ARTICLE INFO

Article history:
Received 2 February 2015
Revised 4 October 2015
Accepted 5 October 2015
Available online 23 October 2015

Keywords:
Multigranulation fuzzy approximation space
Variable precision rough sets
Lower and upper approximations
Three-way decisions

ABSTRACT

This paper studies variable precision multigranulation fuzzy decision-theoretic rough sets in an information system. We firstly review definitions and properties of multigranulation fuzzy rough sets. A novel membership degree based on single granulation rough sets is proposed. Then two operators based on this membership degree are defined. By employing these operators, two types of variable precision multigranulation fuzzy rough sets in an information system are proposed. Finally, inspired by three-way decisions, we propose Type-1 variable precision multigranulation decision-theoretic fuzzy rough sets.

1. Introduction

Human brain is a very great machine to capture useful information quickly and accurately, so it is an interesting subject to simulate the thinking way of human brain. There are many attempts. One of the methods is Pawlak's rough set theory [21], which deals with insufficient and incomplete data. It is one of the characteristics of rough set theory that uncertain concepts and phenomena are approximated by the existing knowledge [22]. From both theoretical and practical viewpoints, Pawlak's rough approximation is very stringent, and may limit application scopes. With more than thirty years' development, many authors have generalized Pawlak's rough set theory by using nonequivalence binary relations [1,13,32,34,43,45,49], and developed these rough set models based on reasoning and knowledge acquisition in incomplete information tables. Moreover, many authors generalized rough approximations to fuzzy environments, for example, rough fuzzy sets and fuzzy rough sets [3,7,23,27,36,52,53,55]. These models have been employed to handle fuzzy and quantitative data.

Previously, many scholars used granular computing to analyze information sources. The method of granular computing is proposed by Zadeh [54], which is based on a single granulation structure. Recently, rough set theory becomes a popular mathematical framework for granular computing. In this theory, concepts are expressed by upper and lower approximations induced by a single granulation structure [8,9,24,44]. Thus, it is called single granulation rough set model.

Since we can catch an element from different aspects [26] or different levels [18,37], and we always meet different useful information sources for the same element, so we need to give an overall consideration for these information sources. Thus, the theory of granular computing should be generalized to suit multiple information sources. In order to meet actual needs, Qian et al. [27] first proposed multigranulation rough sets (MGRS, in short). It has a more widely application scope, for example, decision making, feature selection, and so on [12,28–30,42]. Since for different requirements, a concept can be described by different multiple binary relations, many extensions of MGRSs have been proposed. For example, Qian and Liang et al. generalized classical multigranulation rough sets to neighborhood-based ones [11] and covering ones [7]. The neighborhood multigranulation rough sets are useful for hybrid data sets. Dou et al. [2] investigated variable precision multigranulation rough sets. Yao et al. discussed rough set models in multigranulation spaces [50]. Qian et al. [31] discussed multigranulation decision-theoretic rough sets. The topological structures of multigranulation rough sets were discussed by She et al. [35]. Li et al. [6] made a detailed comparison between multigranulation rough sets and concept lattices via rule acquisition. Furthermore, multigranulation rough sets based on fuzzy binary relations [41] and multigranulation fuzzy rough sets based on classical tolerant relations [38] were defined. Liu et al. [14,15] proposed fuzzy covering multigranulation rough sets. Liang et al. [10] proposed an efficient algorithm for feature selection in large-scale and multiple granulation data sets. Since there always exist some noises, uncertainty and fuzziness in an information system, it is difficult for us to deal with all the problems only using the above MGRS theories. It is necessary for us to study the fuzzy MGRSs.

There always exist a few errors in approximations, but some uncertainty in classification process is admitted in making decision. To handle uncertain and imprecise information, in 1993, Ziarko proposed variable precision rough set model [58], which is directly

* Corresponding author. Tel.: +8613785119408.
E-mail addresses: fengtao_new@163.com (T. Feng), mijsh@263.net (J.-S. Mi).

http://dx.doi.org/10.1016/j.knosys.2015.10.007
0950-7051/

derived from Pawlak's rough set model without any additional assumptions. It may make a better utilization of data being analyzed, and a lower likelihood of incorrect decision. Moreover, this model is also useful for eliminating noise attributes. In order to weaken the errors generated by incompleteness, uncertainty and noises, we try to construct variable precision rough sets based on multigranulation fuzzy rough sets.

The essential ideas of three-way decisions are commonly used in different fields and disciplines by different names and notations. It is more suitable for decision making of human cognition [19]. Yao first proposed a unified framework of the theory of three-way decisions in 2010 [46]. The theory of three-way decisions is constructed based on the notions of acceptance, rejection and noncommitment, adding a noncommitment notion with respect to two-way decisions. It is also used to interpret three regions of Pawlak's rough sets [47,48]. Corresponding to the three regions, one may construct rules for acceptance from the positive region, construct rules for rejection from the negative region and construct rules for noncommitment from the boundary region. In recent years, three-way decision theory develops in a wide range [49] and is more banausic in many aspects, for example, email spam filtering [57] and social networks [25]. Three-way decision rough sets in interval-value and fuzzy circumstances are studied [16,17,56]. Three-way decision model based on the evidence theory is also studied by Xue [40]. Yu et al. studied a tree-based incremental overlapping clustering method using the three-way decision theory [51]. In this paper, we would like to combine the method of three-way decisions and variable precision rough sets based on multigranulation fuzzy approximation spaces, which can help us to make a reasonable and suitable decision for every element.

The rest of this paper is organized as follows: Section 2 reviews definitions and properties of optimistic multigranulation fuzzy approximations and pessimistic multigranulation fuzzy approximations. Section 3 proposes a novel membership degree, and two operators based on this membership degree in multigranulation fuzzy approximation space are also defined. Then, we define two types of variable precision multigranulation fuzzy rough sets in an information system. Section 4 studies decision-theory of the Type-1 variable precision fuzzy rough sets. We then conclude the paper with a summary and give an outlook for further researches in Section 5.

2. Basic concepts

In this section, we first review definitions and propositions of fuzzy rough sets [20] and multigranulation fuzzy rough sets [39].

Let $S=(U,AT)$ be an information system, in which U is a non-empty and finite universe of discourse, and AT is a non-empty finite set of attributes. The set of all fuzzy sets defined on U is denoted by $F(U)$. Every attribute is a fuzzy set, that is, $\forall x\in U,\ a\in AT$, the value of x on attribute a is $a(x)\in[0,1]$. R: $U\times U\to[0,1]$ is a fuzzy tolerance relation [20] satisfying (1) reflexivity: $R(x,x)=1$, $\forall x\in U$, (2) symmetry: $R(x,y)=R(y,x)$, $\forall x,y\in U$. Given $A\subseteq AT$, R_A is a fuzzy tolerance relation, $\forall x\in U$, $R_A(x)$ is a fuzzy set such that $R_A(x)(y)=R_A(x,y)$, $\forall y\in U$.

Using fuzzy tolerance relation R_A, $\forall X\in F(U)$, the lower and upper approximations of X can be computed by approximation operators $\underline{R_A}(X)$ and $\overline{R_A}(X)$. They are defined as follows: $\forall x\in U$,

$$\underline{R_A}(X)(x)=\bigwedge_{y\in U}((1-R_A(x,y))\vee X(y)),\tag{1}$$

$$\overline{R_A}(X)(x)=\bigvee_{y\in U}(R_A(x,y)\wedge X(y)).\tag{2}$$

The partial relation of two fuzzy tolerance relations is defined as: for two fuzzy tolerance relations R_1 and R_2, $R_1\preceq R_2$ if and only if $R_1(x)\subseteq R_2(x)$ for each $x\in U$. In this case, R_2 is coarser than R_1.

Multiple granulation structures can be obtained by different fuzzy binary relations. Combining the granulation structures, multigranulation fuzzy rough sets can be defined.

2.1. Optimistic multigranulation fuzzy rough sets

Associating fuzzy tolerance rough sets with the theory of granular computing, Xu et al. defined optimistic and pessimistic multigranulation fuzzy rough set models on fuzzy tolerance relations, and discussed their properties in 2011 [39]. These models are reasonable generalizations of crisp multigranulation rough set models. In this subsection, we review the optimistic multigranulation fuzzy rough set model.

Suppose $S=(U,AT)$ is a fuzzy information system, $A_1,A_2,\ldots,A_m\subseteq AT$, and R_{A_i} is a fuzzy tolerance relation with respect to A_i, then $\forall X\in F(U)$, $\underline{\sum_{i=1}^m R_{A_i}^O}(X)$ and $\overline{\sum_{i=1}^m R_{A_i}^O}(X)$ are the optimistic multigranulation fuzzy lower and upper approximations, respectively. They are defined as follows: $\forall x\in U$,

$$\underline{\sum_{i=1}^m} R_{A_i}^O(X)(x)=\bigvee_{i=1}^m\left(\bigwedge_{y\in U}((1-R_{A_i}(x,y))\vee X(y))\right),\tag{3}$$

$$\overline{\sum_{i=1}^m} R_{A_i}^O(X)(x)=1-\underline{\sum_{i=1}^m} R_{A_i}^O(\sim X))(x)=\bigwedge_{i=1}^m\left(\bigvee_{y\in U}(R_{A_i}(x,y)\wedge X(y))\right),\tag{4}$$

where $\sim X(x)=1-X(x)$, $\forall x\in U$. If $\underline{\sum_{i=1}^m R_{A_i}^O}(X)=\overline{\sum_{i=1}^m R_{A_i}^O}(X)$, then X is an optimistic multigranulation fuzzy definable set. Otherwise, X is an optimistic multigranulation fuzzy rough set.

The optimistic multigranulation fuzzy boundary region of X is

$$BN^O_{\sum_{i=1}^m R_{A_i}}(X)=\overline{\sum_{i=1}^m} R_{A_i}^O(X)\cap\left(\sim\underline{\sum_{i=1}^m} R_{A_i}^O(X)\right).\tag{5}$$

By the definitions of optimistic multigranulation fuzzy lower and upper approximations, we have the following properties of optimistic multigranulation fuzzy rough sets based on a fuzzy tolerance approximation space: $\forall X,Y\in F(U)$,

1. $\underline{\sum_{i=1}^m R_{A_i}^O}(X)\subseteq X\subseteq\overline{\sum_{i=1}^m R_{A_i}^O}(X)$;
2. $\underline{\sum_{i=1}^m R_{A_i}^O}(\emptyset)=\overline{\sum_{i=1}^m R_{A_i}^O}(\emptyset)=\emptyset$, $\underline{\sum_{i=1}^m R_{A_i}^O}(U)=\overline{\sum_{i=1}^m R_{A_i}^O}(U)=U$;
3. $X\subseteq Y\Rightarrow\underline{\sum_{i=1}^m R_{A_i}^O}(X)\subseteq\underline{\sum_{i=1}^m R_{A_i}^O}(Y)$, $\overline{\sum_{i=1}^m R_{A_i}^O}(X)\subseteq\overline{\sum_{i=1}^m R_{A_i}^O}(Y)$;
4. $\underline{\sum_{i=1}^m R_{A_i}^O}(X)=\bigcup_{i=1}^m\underline{R_{A_i}}(X)$, $\overline{\sum_{i=1}^m R_{A_i}^O}(X)=\bigcap_{i=1}^m\overline{R_{A_i}}(X)$;
5. $\underline{\sum_{i=1}^m R_{A_i}^O}(\sim X)=\sim(\overline{\sum_{i=1}^m R_{A_i}^O}(X))$, $\overline{\sum_{i=1}^m R_{A_i}^O}(\sim X)=\sim(\underline{\sum_{i=1}^m R_{A_i}^O}(X))$;
6. $\underline{\sum_{i=1}^m R_{A_i}^O}(X\cap Y)=\bigcup_{i=1}^m(\underline{R_{A_i}}(X)\cap\underline{R_{A_i}}(Y))$,
   $\overline{\sum_{i=1}^m R_{A_i}^O}(X\cup Y)=\bigcap_{i=1}^m(\overline{R_{A_i}}(X)\cup\overline{R_{A_i}}(Y))$;
7. $\underline{\sum_{i=1}^m R_{A_i}^O}(X\cap Y)\subseteq\underline{\sum_{i=1}^m R_{A_i}^O}(X)\cap\underline{\sum_{i=1}^m R_{A_i}^O}(Y)$,
   $\overline{\sum_{i=1}^m R_{A_i}^O}(X\cup Y)\supseteq\overline{\sum_{i=1}^m R_{A_i}^O}(X)\cup\overline{\sum_{i=1}^m R_{A_i}^O}(Y)$;
8. $\underline{\sum_{i=1}^m R_{A_i}^O}(X\cup Y)\supseteq\underline{\sum_{i=1}^m R_{A_i}^O}(X)\cup\underline{\sum_{i=1}^m R_{A_i}^O}(Y)$,
   $\overline{\sum_{i=1}^m R_{A_i}^O}(X\cap Y)\subseteq\overline{\sum_{i=1}^m R_{A_i}^O}(X)\cap\overline{\sum_{i=1}^m R_{A_i}^O}(Y)$.

If $R_{A_1}\preceq R_{A_2}\preceq\cdots\preceq R_{A_m}$, then we have (1) $\underline{\sum_{i=1}^m R_{A_i}^O}(X))(x)=\underline{R_{A_1}}(X)(x)$; (2) $\overline{\sum_{i=1}^m R_{A_i}^O}(X))(x)=\overline{R_{A_1}}(X)(x)$, $\forall X\in F(U)$. Thus the optimistic multigranulation fuzzy lower and upper approximations are dependent on the thinnest fuzzy relation.