Information Sciences 241 (2013) 101–118

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Multigranulation rough sets: From partition to covering ☆

Guoping Lin [a,b], Jiye Liang [a,*], Yuhua Qian [a]

[a] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] Department of Mathematics and Information Science, Zhangzhou Normal University, Zhangzhou, 363000 Fujian, China

ARTICLE INFO

Article history:
Received 2 December 2011
Received in revised form 18 March 2013
Accepted 25 March 2013
Available online 2 April 2013

Keywords:
Covering
Multigranulation
Granular computing
Rough set

ABSTRACT

The classical multigranulation rough set (MGRS) theory offers a formal theoretical framework for solving the complex problem under multigranulation environment. However, it is noticeable that MGRS theory cannot be applied in multi-source information systems with a covering environment in the real world. To address this issue, we firstly present in this paper three types of covering based multigranulation rough sets, in which set approximations are defined by different covering approximation operators. Then, by using two different approximation strategies, i.e., seeking common reserving difference and seeking common rejecting difference, two kinds of covering based multigranulation rough set are presented, namely, a covering based optimistic multigranulation rough set and a covering based pessimistic multigranulation rough sets. Finally, we develop some properties and several uncertainty measures of the covering based multigranulation rough sets. These results will enrich the MGRS theory and enlarge its application scope.

1. Introduction

Rough set theory, proposed by Pawlak [30,31], is a well-established mechanism for dealing with vagueness and uncertainty in data analysis. It is an efficient method employed in many areas: feature selection [6,12,13,15,19,45], knowledge reduction [17,20–23,35], rule extraction [1,46], uncertainty reasoning [9,33], granular computing [3,16,24,32,50,52], and others [5,7,8].

Rough set theory is originally constructed on the basis of an indiscernibility relation (or an equivalence relation) or a partition of the universe. However, it is restrictive for many real-world applications. To overcome this limitation, there are two main methods to generalize the classical rough sets. One method is to extend the equivalence relation to other binary relations, such as similarity relation, tolerance relation, and dominance relation [14,42,43,47,53]. The other important method is to replace a partition of the universe with a covering [2,4,11,26,27,34,41,54–60]. In 1983, Zakowski [55] has first employed the covering of a universe for establishing a covering based generalized rough set. Since then, many researchers have proposed a great number of diversity upper and lower approximation operators and studied them extensively [2,4,11,26,27,34,41,54,56–60]. For example, Yao [54] investigated approximation operators by using coverings produced by the predecessor and/or successor neighborhoods of serial or inverse serial binary relations. Zhu et al. [56–60] systematically studied six types of approximation operators and investigated their properties and relationships of them. Particularly, Yao [54] studied a unified framework and a more systematic formulation of covering based rough sets from three aspects: the element, the granule, and the subsystem. In fact, the existing approximation operators have either dual property or non-

☆ This is an extended version of the paper published at 2011 International Conference on Intelligent Computing, August 11–14, 2011, Zhengzhou, China.

* Corresponding author. Tel./fax:+86 0351 7018176.
E-mail addresses: guoplin@163.com (G. Lin), ljy@sxu.edu.cn (J. Liang), jinchengqyh@126.com (Y. Qian).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2013.03.046

dual property. Under the covering application background of rough sets, Chen et al. [4] presented a new covering to construct the upper and lower approximations of an arbitrary set. Covering based generalized rough sets are important improvements among these extensions, which can handle more complex practical problems. And they have obtained much attention in many domains including machine learning and uncertainty reasoning. Actually, in the view of granular computing [52], either a partition or a covering of the universe can be considered as a granular space.

From the above, we can see that set approximations in the above rough sets are described only by a single binary relation (a single granulation [52]) or a single covering (or a single covering granulation) on a given universe, which cannot be applied in some practical multigranulation backgrounds [36,37]. Qian et al. [36] first took multiple binary relations into account and proposed multigranulation rough sets, in which a target concept was described by multiple binary relations on a universe according to a user's different requirements. Up to now, many extensions of MGRS have been proposed. For example, Liu et al. [28,29] proposed covering fuzzy rough set based multigranulation rough sets. Xu et al. [48] investigated another generalized version, called variable precision multigranulation rough sets. Yang et al. [51] proposed a multigranulation rough set based on a fuzzy binary relation. Lin et al. [25] investigated neighborhood-based multigranulation rough sets, which can be used to deal with data sets with hybrid attributes. She et al. [44] explored topological structures of multigranulation rough sets, which further enriches the theory of MGRS. It is deserved to mention that Liang et al. [15] proposed an efficient feature selection algorithm for large-scale data sets from the perspective of multiple granulations, which has shown an important implication of MGRS theory. Accordingly, MGRS theory has displayed its advantages in knowledge discovery from large-scale data sets. In fact, in a Pawlak's approximation space, each object can be classified into a certain concept as shown in Fig. 1. However, in real-world applications, such as a multi-source covering information system [10] and computing with words, different subsets of the universe usually overlap, as shown in Fig. 2, in which these basic information granules form a covering of the objects, rather than a Pawlak's approximation space. It is difficult for the classical MGRS theory to deal with this issue. To address this issue, it is necessary to generalize the classical MGRS to covering based multigranulation rough sets for enriching its application domains.

In this paper, we introduce covering into the multigranulation environment and present covering based optimistic and pessimistic multigranulation rough sets.

Additionally, lots of researchers suggested some possible applications of the uncertainty measures in the fields of pattern recognition and image analysis in the literature [9,18,39–41,49]. The concept of entropy was originally introduced by Shannon in [40], which is a very useful mechanism for characterizing information content in various modes. It has been applied in many diverse fields. Furthermore, Shannon entropy and its variants were adopted for rough set theory in the literature [9,18,39,41,49]. Similarly, in this paper, in order to make wide applications of the covering based multigranulation rough set theory, we propose several uncertainty measures for covering based multigranulation rough sets, including degree of rough membership, approximation measure, and rough entropy.

The main objective of this paper is to establish three types of rough sets based on multiple coverings by using different approximation strategies due to the practical different applied backgrounds. The rest of this paper is organized as follows. Some basic concepts of classical multigranulation rough sets are briefly reviewed in Section 2. In Section 3, three types of covering based optimistic and pessimistic multigranulation rough sets are constructed and some of their important properties are investigated. In Section 4, several uncertainty measures for covering based multigranulation rough sets are presented, such as degree of rough membership, approximation measure, and rough entropy. We then conclude the paper with a summary and direction for the further research in the last section.

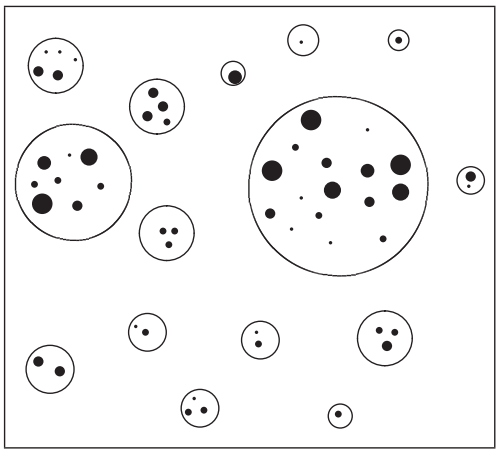

Fig. 1. A partition on a universe of discourse.