Information Sciences 370–371 (2016) 239–255

Contents lists available at ScienceDirect

Information Sciences

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

Boundary region-based rough sets and uncertainty measures in the approximation space☆

Zhouming Ma[a,b], Jusheng Mi[a,*]

[a] College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, Hebei 050024, PR China
[b] School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, Fujian 363000, PR China

ARTICLE INFO

Article history:
Received 31 August 2015
Revised 2 June 2016
Accepted 17 July 2016
Available online 1 August 2016

Keywords:
Boundary region
Axiomatization
Uncertainty measure
Rough sets

ABSTRACT

The approximation operators play an important role in rough set theory, which are mainly defined by means of neighborhood systems. In this paper, firstly, we try to propose a class of novel definitions of the approximation operators via a predefined boundary region based on a binary relation. Then we compare the proposed concepts with the originals, the necessary and sufficient conditions of their equipollence are investigated. Secondly, we give the definitions of boundary region based on a covering. By employing the boundary, a class of novel definitions of the approximation operators based on a covering are proposed. It is shown that the proposed operators are equivalent to a class of covering approximation operators introduced by Zakowski. Thirdly, the relationship between general binary relations and coverings based approximation operators is investigated with the aid of the novel boundary region. Finally, the more reasonable characterizations of the accuracy and the roughness are proposed by employing the boundary operators. Meanwhile, we study uncertainty measures of approximation spaces based on a partition and a covering.

1. Introduction

Rough set theory, as a powerful mathematical tool to deal with uncertainty, granularity, and incompleteness of knowledge in information systems, was proposed by Pawlak [30–32] in 1982. Many scholars have developed the theory from different perspectives, such as the coverings [45,57,59], the fuzzy sets [24,26,28,46,48], the intuitionistic fuzzy sets [7,9,60], the multigranulation rough sets [19,20,39–41,55] and the knowledge discoveries [15–17,27]. As a result, it has been successfully applied in many fields [1,5,6,10,12,13,21,22,34] involving artificial intelligences [35,36,47,49,56].

The lower and upper approximation operators are the most important concepts in the theory of rough set. Many scholars studied them in different environment. There are two issues of the generalized rough sets: the relations and the coverings based generalized rough sets. The classical rough set is based on an equivalence relation, it is natural to extend equivalence relations to non-equivalence relations such as similarity relations (reflexive and transitive) [43], tolerance relations (reflexive and symmetric) [14,29,37,42], and arbitrary binary relations [23,53,54,61]. Given that an equivalence relation means a partition, which is a special covering, many types of coverings based rough sets were presented and investigated [3,4,38,57]. The axiomatization of these rough sets is another hot issue attracting many researchers. It is mainly studied for the properties of

☆ This is an extended version of the paper published at 2015 International Conference on Machine Learning and Cybernetics, July 12–15, 2015, Guangzhou, China.

* Corresponding author. Fax: +8631180787202.
E-mail addresses: mazhouming@sina.com (Z. Ma), mijsh@263.net (J. Mi).

http://dx.doi.org/10.1016/j.ins.2016.07.040
0020-0255/

the approximation operators, which are the important carriers of the rough set theory, and the brief properties characterize the essential of the theory. Consequently, the axiomatizations of a variety of binary relations based generalized rough sets have basically been solved [25,44,50–53,61], as well as the axiomatizations of the coverings based generalized rough sets. However, there is certain limitation in the extending of the above mentioned rough sets based on an arbitrary relation and a covering by means of the classical definitions of approximation operators. For instance, when a relation is not a tolerance relation, the approximation operators may lose many the basic properties of the classical operators. In addition, it is an important issue for the correlation between the general relations and the coverings based generalized rough sets. And it is not apparent to connect the two types of generalized rough sets for the reason that the coverings are more complex than the partitions.

As one of the most important issues in rough set theory, uncertainty measures have been widely studied. The granulation, as a powerful mechanism to evaluate uncertainties of the approximation spaces, was first introduced by Zadeh [58]. An axiomatic definition of the knowledge granulation was proposed [18], and a variety of forms of the knowledge granulation were introduced in [8,11,17]. All these studies were dedicated to evaluating uncertainty of a set in terms of the partition ability of a classification, and the measure methods are more intuitive and easily understood. To evaluate uncertainty of a set, several numerical measures were presented by Pawlak [33], which are accuracy and roughness of a set and approximation accuracy of a rough classification. Although the measures are effective in most cases, they have some restrictions in some special cases. For instances, when the boundary (i.e., difference between upper and lower approximations) of some sets are equal, it is not well to characterize the degree of closeness to lower or upper approximations of a set associated with a kind of classification task.

In this paper, we first review several existing definitions of the lower and upper approximation operators, and show the similarity and difference among them. In many situations, one always want to extend the classical rough sets to generalized rough sets but remain some basic properties. However, the approximation operators based on general binary relations may not have the basic properties. At this point, our purpose is to avoid the limitation of the original definitions that are straightly extended to the non-tolerance relations based generalized rough sets. Based on these considerations, we introduce a measure of the predefined boundary regions based on the binary relations, and give a novel definitions of the lower and upper approximation operators. We consider the properties of the novel operators and their axiomatic characterizations, and draw a comparison between the novel operators and the original ones. We also give the necessary and sufficient conditions of their equipollence. Secondly, similar to the approach of the binary relations based generalized rough sets, the novel definitions of coverings based approximation operators are proposed. Comparing the proposed definitions with the existing ones, we point out that the covering based novel operators and the covering approximation operators proposed by Zakowski [57] are equivalent. We also investigate the properties and axiomatic characterizations of the covering based novel operators. Thirdly, we study the relationship between the covering and binary relation based generalized rough sets. In fact, Zhu have established the relationship between a class of covering and a similarity relation based rough sets [62]. Bartol et al.characterized the covering generated by a tolerance relation [2]. In our works, we show that the novel approximation operators based on an arbitrary relation are equivalent to the ones based on a reduction of the corresponding covering. Meanwhile, the novel operators based on a covering are equivalent to the ones based on a tolerance relation. In addition, with the aid of the novel definitions, we also study relationship between the axiomatic characterization of the approximation operators based on the binary relations and the ones based on the coverings. Finally, we give the improved axiomatic definition of the knowledge granulation, and introduce a more reasonable knowledge granulation based on the novel boundary operators. It is more effective to characterize the approximation spaces based on a partition and a covering. For a object set, to distinguish the different effect of the same boundaries on various sets, we also introduce some measure methods in greater detail.

The construction of this paper is organized as follows. In Section 2, we review and compare serval definitions of the classical approximation operators, which are based on an equivalence relation. The novel definitions of the approximation operators based on a relation are then presented, and their properties and axiomatic characterizations are also examined. We also study novel definitions of a type of covering based approximation operators in Section 3. In Section 4, we combine the two approximation operators based on the general binary relations and the coverings. Finally, we study the uncertain measures in the approximation spaces in Section 5. The conclusion and prospect of this paper are given in the last section.

2. The approximation operators based on binary relations via the boundary region

In this section, we discuss the general binary relations based approximation operators, and investigate the advantages and disadvantages. To remove the defect of the approximation operators, we introduce a definition of boundary region based on a relation and propose the novel definitions of the approximation operators by using the boundary region. And we also study their properties and axiomatic characterizations.

2.1. The classical approximation operators

We recall basic definitions of the classical approximation operators and some other important concepts, such as neighborhood, precision, precise set, etc. Then we review the basic properties of the approximation operators based on a binary relation.