Knowledge-Based Systems 91 (2016) 204–218

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Generalized attribute reduct in rough set theory

Xiuyi Jia[a], Lin Shang[b,*], Bing Zhou[c], Yiyu Yao[d]

[a] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[b] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
[c] Department of Computer Science, Sam Houston State University, Huntsville 77341, USA
[d] Department of Computer Science, University of Regina, Regina S4S 0A2, Canada

ARTICLE INFO

Article history:
Received 18 December 2014
Revised 13 April 2015
Accepted 16 May 2015
Available online 21 May 2015

Keywords:
Attribute reduction
Rough set
Generalized definition

ABSTRACT

Attribute reduction plays an important role in the areas of rough sets and granular computing. Many kinds of attribute reducts have been defined in previous studies. However, most of them concentrate on data only, which result in the difficulties of choosing appropriate attribute reducts for specific applications. It would be ideal if we could combine properties of data and user preference in the definition of attribute reduct. In this paper, based on reviewing existing definitions of attribute reducts, we propose a generalized attribute reduct which not only considers the data but also user preference. The generalized attribute reduct is the minimal subset which satisfies a specific condition defined by users. The condition is represented by a group of measures and a group of thresholds, which are relevant to user requirements or real applications. For the same data, different users can define different reducts and obtain their interested results according to their applications. Most current attribute reducts can be derived from the generalized reduct. Several reduction approaches are also summarized to help users to design their appropriate reducts.

1. Introduction

Attribute reduction is a key concept in rough set theory [28]. It plays an important role in many areas including machine learning, data mining, and knowledge representation. Specifically, rough sets can be used to construct granular structures in the area of granular computing [35,58]. Attribute reduction has been drawing broad attention in recent years, which can be classified into two groups: One group concentrates on seeking quick reduction algorithms to compute the reducts efficiently [4,6,8,9,11,12,23,17,18,22,25,27,31,32,44,52,54,60,67]; The other group focuses on the definition of reduct to find appropriate reducts for different applications [7,13,14,16,20,21,24,26,28,33,36,40,45–47,56,64–66]. In this paper, we aim to conduct an investigation on the definition of attribute reduct.

Why do we have so many different definitions of attribute reduct? In real applications, for the same data, different users may have different learning tasks, leading to the fact that the learned results are possibly different. Thus, many kinds of attribute reducts have been defined to meet different needs.

What are the differences among these reducts? Generally speaking, an attribute reduct can be interpreted as a minimal set of attributes that can preserve or improve one or several criteria. Different attribute reducts were defined based on different criteria.

Although so many different reducts have been studied, they still cannot be applied directly in some simple applications with user requirement. For example, the rules derived based on the Pawlak's reduct are all certain rules, which is the result of requiring the positive region remain unchanged. Assuming a typical situation, a rule is acceptable to users if its confidence is greater than 80%, then the Pawlak's reduct is no longer suitable, and we cannot find any appropriate reduct from existing definitions for this application. For the sake of a better understanding of this problem, we convert it to the following question.

How to choose or define appropriate reducts for different users in different applications?

In general, previous studies on definition of attribute reduct focus on selecting what kinds of criteria or properties of data to keep unchanged or to extend, such as distribution of objects, quality of classification, and so on. However, all these attribute reducts are relevant to the data only, but irrelevant to the application problem.

* Corresponding author.
E-mail addresses: jiaxy@njust.edu.cn (X. Jia), shanglin@nju.edu.cn (L. Shang), zhou@shsu.edu (B. Zhou), yyao@cs.uregina.ca (Y. Yao).

http://dx.doi.org/10.1016/j.knosys.2015.05.017
0950-7051/

In this paper, we argue that the criteria for definition of reduct should be connected with both the data and the user requirements. Actually, the requirements represent the user's preference and demand for the problem. Based on reviewing existing studies, we propose a generalized attribute reduct which considers both properties of data and user requirements in real applications. The contribution of the generalized reduct is that it can instruct users to define appropriate attribute reducts to meet their requirements.

By considering the user preference on learning data, the generalized attribute reduct is interpreted as a minimal subset of attributes which satisfies a specific condition. The condition is represented by a set of measures and a set of thresholds. Most existing attribute reducts can be derived from our definition through constructing corresponding measures and thresholds. Moreover, the measures and the thresholds can be provided by users or domain experts, which indicate the user requirements in real applications.

The rest of the paper is organized as follows. Section 2 summarizes existing attribute reducts and compares them through experiments. In Section 3, we introduce the definition of our generalized attribute reduct, and discuss some properties of the definition. In Section 4, we briefly review some measures based on rule induction and give a detailed illustration to show how to define an appropriate reduct that users really want. Section 5 derives the existing reducts from our definition. Section 6 introduces several reduction approaches. Section 7 concludes the paper.

2. Summary of existing definitions of attribute reducts

In this section, we summarize existing definitions of attribute reducts, and compare these definitions through experiments.

2.1. Existing definitions of attribute reducts

To summarize existing different definitions of attribute reducts, we classify these definitions into two groups from a decision perspective. One group contains those definitions that are decision-independent. Zhang et al. [64] proposed the theory of attribute reduction in the concept lattice and examined the judgement theorems of consist sets. Wu [47] discussed the attribute reduct in incomplete information systems and incomplete decision systems based on Dempster–Shafer theory of evidence, and introduced the plausibility reduct and belief reduct. Quafafou [33] defined an α-reduct in *Alpha Rough Set Theory* based on α-dependency, which preserves the dependency relation unchanged. Chen [3] introduced a concept of part reduct to describe the minimal description of a definable set by attributes of the given information system.

The other group contains definitions that are decision-dependent. These definitions are usually applied in classification problems. This group can be further divided into two categories. In the first category, the purpose of attribute reduction is to obtain a minimal subset of attributes that has the same classification power as the entire set of condition attributes. Pawlak [30] defined a quantitative reduct which ensures the classification ability unchanged, in which γ is used to represent the quality of classification. Śle¸zak proposed a concept of reduct to find the majority decision rules, he also defined the attribute reduct that keeps the class membership distribution unchanged for all objects by using the membership distribution function [41,42]. Zhang et al. [63] proposed the notions of the distribution reduct and maximum distribution reduct. Mi et al. [24] introduced the concepts of β lower distribution reduct and β upper distribution reduct based on variable precision rough sets. Their reducts preserve the lower distribution and the upper distribution of the decision class unchanged. These attribute reducts are summarized as a minimal subset of attributes that has the same classification power in terms of generalized decision, majority decision, or maximum distribution for all objects in the universe, they concentrated on the decision class or classes to which an equivalence class belongs [66].

In the second category, the reduct is interpreted as a minimal subset of attributes that keeps the positive, boundary and negative regions of decision classes, or other criteria unchanged or extended [66]. Pawlak defined a reduct of knowledge be the essential part which suffices to define all basic concepts. The Pawlak's reduct keeps the positive region unchanged [28,30]. Miao et al. [25] studied the mutual information as the reduction criterion, which is actually a kind of conditional entropy. Wang et al. [45] discussed attribute reduct from an algebra viewpoint and an information viewpoint. In the algebra view, a reduct is a minimal subset of attributes that keeps the positive region unchanged. In the information view, a reduct is a minimal subset of attributes that keeps the conditional entropy unchanged. Hu et al. [13] defined the consistency based attribute reduct, considered the distribution of each decision class under the precondition of keeping positive region unchanged. Jiang and Lu [16] gave two new definitions of reducts based on two concepts: mean decision power and decision information entropy. Xu and Sun [49] constructed a new conditional entropy based reduct to reflect the change of decision ability objectively in a decision table. For the definitions of reducts in Pawlak rough set model, the boundary region and the negative region were usually not considered. However, for the reducts defined in probabilistic rough sets, all three regions were considered with different decision region rules [56]. In decision-theoretic rough set model, Li et al. [20] proposed a positive region expanded attribute reduct, as the monotonicity of positive region does not always hold. Jia et al. [14,15] introduced a minimum cost attribute reduct which can minimize the decision cost.

Additionally, some researchers focused on the relationships between different definitions. Kryszkiewicz [19] compared several different reducts, analyzed the relationships between them, and generalized that a reduct is a minimal subset of attributes that satisfying some specific criteria. Miao et al. [26] discussed several definitions of reducts based on consistent and inconsistent data, and introduced corresponding algorithms. Some researchers also have done some works on the generality of reducts. Yao and Zhao [56] introduced a generalized reduct in probabilistic rough set models, which is a minimal subset of attributes that satisfying some criteria. Wang et al. [46] generalized the equivalence relation to a binary relation, and defined an attribute reduct based on the binary relation. Śle¸zak [40] also suggested that some measures can be the criteria for defining a reduct, such as information entropy.

Existing definitions of attribute reducts are summarized in Table 1.

2.2. Experimental comparisons of attribute reducts

In this section, we check the performances of different kinds of reducts on several criteria through some comparison experiments.

2.2.1. Comparison reducts

There are 22 different reducts in Table 1, they can be further grouped according to some criteria. A coarse and a fine approach to group all reducts are shown in Tables 2 and 3. Several typical reducts will be selected as the comparison algorithms.

The addition–deletion method[1] is applied to implement these definitions. The method starts with an empty set and uses inner significance to rank the attributes. As the fuzzy information relevant definition needs more expert opinions and the formal context relevant definition is not easy to apply the addition–deletion method, we

[1] The reduction approach will be explained in Section 6.