Neurocomputing 139 (2014) 336–344

Contents lists available at ScienceDirect

Neurocomputing

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

# Novel algorithms of attribute reduction with variable precision rough set model

Yanyan Yang [a,*], Degang Chen [b], Ze Dong [a]

[a] School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, PR China
[b] Department of Mathematics and Physics, North China Electric Power University, Beijing 102206, PR China

ARTICLE INFO

*Article history:*
Received 25 October 2013
Received in revised form
4 January 2014
Accepted 10 February 2014
Communicated by D. Wang
Available online 12 April 2014

*Keywords:*
Variable precision rough sets
Discernibility matrix
Minimal element
Equivalence class pair relative to condition attributes

ABSTRACT

The variable precision rough set model resists noise in data by introducing a parameter to relax the strict inclusion in approximations of the classical rough set model, and attribute reduction with the variable precision rough set model aims at deleting dispensable condition attributes from a decision system by considering this relaxation. In the variable precision rough set model, the approach of the discernibility matrix is the theoretical foundation of attribute reduction. However, this approach has always heavy computing load, so its effective improvement is clearly of importance in order to find reducts faster. In this paper, we observe that only minimal elements in the discernibility matrix are sufficient to find reducts, and each minimal element is at least determined by one equivalence class pair relative to condition attributes. With this motivation, the relative discernibility relation of a condition attribute is defined to characterize minimal elements in the discernibility matrix, and the algorithm of finding all minimal elements is developed by this characterization. Based on the algorithm of finding all minimal elements, we develop two algorithms to find all reducts and one reduct in variable precision rough sets. Several experiments are performed to demonstrate that our methods proposed in this paper are effective to reduce the computing load.



## 1. Introduction

The variable precision rough set model (VPRS) [30,31], which is one of the improvements of Pawlak's rough set model [7,10,24,30], shows certain robustness to misclassification and noise in data. It relaxes the strict inclusion in approximations of Pawlak's rough set model [17,18] to the partial inclusion by considering a parameter as an inclusion degree. Essentially, this parameter represents a threshold that determines the portion of the objects in an equivalence class which are classified to the same decision class. Many researchers have focused on the selection of the parameter [1–5,8,21] and developing some extensions of VPRS [11,13,20,26,28]. On the other hand, more researches have been put forward on the investigation of attribute reduction with VPRS [3,9,12,14,25,29]. Similar to the purpose of attribute reduction with Pawlak's rough set model, attribute reduction with VPRS also aims at selecting a set of less condition attributes to provide the same classification information as the full condition attributes set by introducing this parameter.

Attribute reduction with VPRS was originally proposed in [30] to preserve the sum of objects in the $\beta$-lower approximations of all decision classes by defining the dependency function. This kind of reduct is called approximate reduct or $\beta$-reduct, and was noted not to preserve the $\beta$-lower approximation of each decision class in VPRS [3,9]. Therefore, the derived decision rules from $\beta$-reduct may be in conflict with the ones from the original decision system [14]. Considering these drawbacks of $\beta$-reduct, the concepts of $\beta$-lower and upper distributed reducts based on VPRS were presented in [14] to preserve $\beta$-lower and upper approximations, respectively. Meanwhile, the derived decision rules from $\beta$-lower and upper distributed reducts are compatible with the ones derived from the original system. Furthermore, in [14] $\beta$-lower and upper distributed discernibility matrices were also provided to find $\beta$-lower and upper distributed reducts, respectively.

As noted in [29], the dependency function, the positive region and the $\beta$-lower approximation may not be monotonic in VPRS after deleting a condition attribute. Consequently, the algorithm of finding $\beta$-reduct may not be convergent by applying the measure of dependency function. In [22], the algorithm of attribute reduction was proposed by using $\beta$-variable precision rough entropy, where both the precision parameter and the correction coefficient were given in advance. VPRS-QuickReduct algorithm was proposed in [16] to improve the current QuickReduct algorithm by selecting the biggest dependency attribute during each iteration and using VPRS's dependency as a criterion which admitted some small

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2014.02.023
0925-2312/

errors. However, the above two heuristic algorithms in [16,22] share a common trouble that they usually cannot find a proper reduct but an over-reduct or sub-reduct due to their stop criteria. Here, an over-reduct may include redundancy condition attributes, while a sub-reduct may ignore some key condition attributes. An over-reduct or sub-reduct is acceptable to save running time, but it is not enough to further capture key attributes for practical problems. To find proper reducts, the commonly used discernibility matrix method was developed in [14], where a discernibility function was constructed to find all reducts. Besides, the binary discernibility matrix [23,6] was successively developed to find attribute reduction for VPRS. In essence, the binary discernibility matrix [6,23] is equivalent to the distributed discernibility matrix proposed in [14], and can be also used to find all reducts. Even though finding all reducts with this technique is an NP-hard problem [19], this method provides a mathematical foundation of finding attribute reduction for VPRS. On the other hand, a search algorithm based on the discernibility matrix can be developed to find a proper reduct but requires heavy computational load since it has to employ all non-empty elements in the discernibility matrix.

We observe in this paper that it is dispensable to calculate all elements of a $\beta$-distributed discernibility matrix, and only some selected elements called minimal elements are sufficient to find reducts. We also find that every minimal element can be determined by one pair of equivalence class relative to condition attributes (ECPC) at least. These facts motivate us in this paper to develop novel algorithms to compute $\beta$-distributed reducts based on minimal elements in the $\beta$-distributed discernibility matrices. As a result, the computational load can be effectively reduced.

In this paper, we develop a technique to find $\beta$-lower and upper distributed reducts for VPRS by improving the existing approach of the $\beta$-upper and lower distributed discernibility matrices, where only minimal elements in the $\beta$-upper and lower distributed discernibility matrices are considered. First, $\beta$-upper and lower distributed relative discernibility relations of each condition attribute are defined. The minimal elements in the $\beta$-upper and lower distributed discernibility matrices are characterized by the $\beta$-upper and lower distributed relative discernibility relations, respectively. Then, we develop an algorithm to find all minimal elements in the $\beta$-upper and lower distributed discernibility matrices. Finally, we design novel algorithms to compute all $\beta$-distributed reducts and one $\beta$-distributed reduct based on the minimal elements, and several experiments are also performed to demonstrate our idea in this paper is effective.

The remainder of this paper is structured as follows. In Section 2, we review some basic notions of attribute reduction for VPRS. In Section 3, we characterize minimal elements of the $\beta$-distributed discernibility matrix by $\beta$-distributed relative discernibility relations and develop the algorithms of finding all minimal elements, all $\beta$-distributed reducts and one $\beta$-distributed reduct. In Section 4, taking the $\beta$-lower distributed reduct as an example, we perform numerical experiments to demonstrate the effectiveness of our methods proposed in this paper. We conclude this paper in Section 5.

## 2. Attribute reduction for VPRS

In this section, we mainly review attribute reduction with VPRS from [14].

An information system is a pair $S=(U,A)$, where $U=\{x_1, x_2, \ldots, x_n\}$ is a non-empty, finite universe of discourse and $A=\{a_1, a_2, \ldots, a_m\}$ is a non-empty, finite set of attributes. With every $B\subseteq A$, we associate a binary relation $IND(B)$, which is called the $B$-indiscernibility relation and defined as $IND(B)=\{(x,y)\in U\times U: a(x)=a(y), \forall a\in B\}$. Then, $IND(B)$ is an equivalence relation and can partition $U$ into a family of disjoint subsets $U/IND(B)=\{[x]_B: x\in U\}$, where $[x]_B$ denotes the equivalence class of $IND(B)$ including $x$.

For an inclusion degree $\beta\in(0.5,1]$ and $X\subseteq U$, we can characterize $X$ by a pair of $\beta$-upper and lower approximations: $\overline{IND(B)}_\beta(X)=\cup\{[x]_B: P(X|[x]_B)>1-\beta\}$ and $\underline{IND(B)}_\beta(X)=\cup\{[x]_B: P(X|[x]_B)\geq\beta\}$. Here, $P(X|Y)=|X\cap Y|/|Y|$ if $|Y|>0$, and $P(X|Y)=1$ otherwise. $|X|$ is the cardinality of set $X$.

A decision table (DT) is an information system $A^*=(U, A\cup D)$, where $A\cap D=\phi$. $A$ is the set of condition attributes, while $D=\{d\}$ is the decision attribute set. For the sake of simplicity in the sequel, suppose $V_d=\{1,2,\ldots,r\}$ is the range of $d$ and $U/IND(D)=\{D_1, D_2, \ldots, D_r\}$, where $D_j=\{x\in U, d(x)=j\}$ is the decision class for $j=1,2,\ldots,r$.

**Definition 2.1.** [14] Let $(U, A\cup D)$ be a DT, $B\subseteq A$, $H_B^\beta=(\overline{IND(B)}_\beta(D_1),\ldots,\overline{IND(B)}_\beta(D_r))$, $L_B^\beta=(\underline{IND(B)}_\beta(D_1),\ldots,\underline{IND(B)}_\beta(D_r))$. $B$ is called a $\beta$-upper (lower) distributed reduct of $(U, A\cup D)$ iff $B$ is a minimal subset of $A$ such that $H_B^\beta=H_A^\beta(L_B^\beta=L_A^\beta)$.

According to Definition 2.1, we can see that $\beta$-upper (lower) distributed reduct preserves $\beta$-upper (lower) approximation of each decision class. However, as indicated in [26], $\beta$-upper (lower) distributed may not be monotonic with the parameter, and thus in our paper we will not discuss the relationship between this type of reduct and the parameter. To find $\beta$-upper (lower) distribute reducts, $\beta$-upper (lower) distributed discernibility matrices are given as follows.

**Definition 2.2.** [14] Let $(U, A\cup D)$ be a DT, $B\subseteq A$, $U/IND(D)=\{D_1,\ldots,D_r\}$, $U/IND(B)=\{M_1,\ldots,M_n\}$, $M_B^\beta(x)=\{D_j: x\in\overline{IND(B)}_\beta(D_j)\}$ and $G_B^\beta(x)=\{D_j: x\in\underline{IND(B)}_\beta(D_j)\}$, $\forall x\in U$. We denote $D_1^{*\beta}=\{([x]_A,[y]_A): M_A^\beta(x)\neq M_A^\beta(y)\}$, $D_2^{*\beta}=\{([x]_A,[y]_A): G_A^\beta(x)\neq G_A^\beta(y)\}$ and denote by $a_k(M_i)$ the value of the condition attribute $a_k$ with respect to the objects in $M_i$.

Define

$$D_l^\beta(M_i,M_j)=\begin{cases}\{a_k\in A: a_k(M_i)\neq a_k(M_j)\}, & (M_i,M_j)\in D_l^{*\beta},\\ A, & (M_i,M_j)\notin D_l^{*\beta},\end{cases}$$

$(l=1,2)$.

Then, $D_l^\beta(M_i,M_j)$, $(l=1,2)$ are called $\beta$-upper distributed and $\beta$-lower distributed discernibility attribute sets respectively. $D_l^\beta=(D_l^\beta(M_i,M_j),\ i,j\leq n)$, $(l=1,2)$ are called $\beta$-upper distributed and $\beta$-lower distributed discernibility matrices, respectively.

**Remark 2.1.** In the remainder of this paper, we assume that notions with the subscripts $l$ are related to $\beta$-upper distributed and $\beta$-lower distributed discernibility matrices respectively and the mark $(l=1,2)$ located by these notions is ignored.

Clearly, $D_l^\beta$ is symmetric and $D_l^\beta(M_i,M_i)=A$. For $\forall D_l^\beta(M_i,M_j)\in D_l^\beta$, if there does not exist another element in $D_l^\beta$ as its proper subset, $D_l^\beta(M_i,M_j)$ is called a minimal element in $D_l^\beta$.

Denoted by $M_l^\beta=\wedge\{\vee\{a_k: a_k\in D_l^\beta(M_i,M_j)\}: i,j\leq n\}$, $M_l^\beta$ are referred to the $\beta$-upper and lower distributed discernibility functions, respectively. Let $M_l^{*\beta}$ be the reduced disjunctive form of $M_l^\beta$ obtained from $M_l^\beta$ by applying the distributed and absorption laws as many times as possible, then there exist $t$ and $A_{li}\subseteq A$ for $i=1,2,\ldots,t$, such that $M_l^{*\beta}=(\wedge A_{l1})\vee\cdots\vee(\wedge A_{lt})$. Therefore, $RED_l^\beta=\{A_{l1}, A_{l2},\ldots,A_{lt}\}$ are the collections of all $\beta$-upper and lower distributed reducts, respectively.

Moreover, we denote the intersection of all $\beta$-upper or lower distributed reducts by $I_l^\beta=\{a\in A: D_l^\beta(M_i,M_j)=\{a\}$, where $(M_i,M_j)\in D_l^{*\beta}\}$ [14].