Attribute Core of Decision Table

G.Y. Wang

Institute of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, 400065, P. R. China
wanggy@cqupt.edu.cn

Abstract. Attribute core of a decision table is often the start point and key of many information reduction procedures. In this paper, we will study the problem of calculating the attribute core of a decision table. We find some errors in some former results by Hu and Ye [1, 2]. Both definitions of attribute core in the algebra view and information view are studied and their difference is discovered. An algorithm for calculating the attribute core of a decision table in the information view and a systemic method for calculating the attribute core of a decision table in different cases are developed.

1 Introduction

Rough set theory has been applied in such fields as machine learning, data mining, etc., successfully since Professor Z. Pawlak developed it in 1982. Reduction of decision table is one of the key problems of rough set theory. The attribute core of a decision table is always the start point of information reduction. In this paper, we will study the problem of calculating the attribute core of a decision table. There are some theory results on this problem by Hu [1] and Ye [2]. Unfortunately, there are some errors in these results. We will study these problems and develop some new theory and algorithm through examining the difference between the algebra view and information view of rough set.

2 Problem in Calculating Attribute Core of a Decision Table

Hu developed a method to calculate the attribute core of a decision table based on Skowron's discernibility matrix [1].

Def. 1. For a set of attributes $B \subseteq C$ in a decision table $S=(U, C \cup D, V, f)$, the discernibility matrix can be defined by $C_D(B)=\{C_D(i,j)\}_{n \times n}$, $1 \le i,j \le n = |U/\text{Ind}(B)|$, where

$$C_D(i,j) = \begin{cases} \{a \in B \mid a(O_i) \neq a(O_j), O_i \in U, O_j \in U\} & d(O_i) \neq d(O_j) \\ 0 & d(O_i) = d(O_j) \end{cases},$$

for $i,j=1,2,\ldots,n$.

Hu drew the following conclusion in [1]: $|C_D(i,j)|=1$ if and only if the attribute in it belongs to $\text{CORE}_D(C)$. This conclusion is used in many later documents.

J.J. Alpigini et al. (Eds.): RSCTC 2002, LNAI 2475, pp. 213–217, 2002.

Ye proved Hu's conclusion might not be true under some conditions by giving a counterexample. He developed an improved method to calculate the attribute core through improving the definition of discernibility matrix in the following way.

Def. 2. For a set of attributes $B\subseteq C$ in a decision table $S=(U, C\cup D, V, f)$, the discernibility matrix can be defined by $C'_D(B)=\{C'_D(i,j)\}_{n\times n}$, $1\le i,j\le n=|U/\mathrm{Ind}(B)|$, where

$$C'_D(i,j)=\begin{cases} C_D(i,j) & ,\min\{|D(x_i)|,|D(x_j)|\}=1 \\ & ,else \end{cases},$$

for $i,j=1,2,\ldots,n$, $|D(x_i)|=|\{d_y|y\in[x_i]_C\}|$.

In [2], Ye drew another conclusion: $|C'_D(i,j)|=1$ if and only if the attribute in it belongs to $\mathrm{CORE}_D(C)$.

Unfortunately, Ye did not find the real reason leading to the error of Hu's conclusion and method, the inconsistency in the decision table. The methods for calculating the attribute core of a decision table in [1] and [2] are not complete.

3 Information View of Rough Set

For the convenience of later illustration, we discuss some basic concepts about the information view of rough set at first [3].

Def. 3. Given an information system $S=(U, C\cup D, V, f)$, and a partition of U with classes X_i, $1\le i\le n$. The entropy of attributes B is defined as $\mathrm{H}(B)=-\sum_{i=1}^{n}p(X_i)\log(p(X_i))$, where, $p(X_i)=|X_i|/|U|$.

Def. 4. Given an information system $S=(U, C\cup D, V, f)$, the conditional entropy of D ($U/\mathrm{Ind}(D)=\{Y_1,Y_2,\ldots,Y_m\}$) given $B\subseteq C$ ($U/\mathrm{Ind}(B)=\{X_1,X_2,\ldots,X_n\}$) is $\mathrm{H}(D|B)=-\sum_{i=1}^{n}p(X_i)\sum_{j=1}^{m}p(Y_j|X_i)\log(p(Y_j|X_i))$, where $p(Y_j|X_i)=|Y_j\cap X_i|/|X_i|$, $1\le i\le n$, $1\le j\le m$.

The following 3 theorems can be proved easily.

Theorem 1. Given a relatively consistent decision table $S=(U, C\cup D, V, f)$, an attribute $r\in C$ is relatively reducible if and only if $\mathrm{H}(D|C)=\mathrm{H}(D|C-\{r\})$.

Theorem 2. Given a relatively consistent decision table $S=(U, C\cup D, V, f)$, the attribute set C is relatively independent if and only if $\mathrm{H}(D|C)\ne\mathrm{H}(D|C-\{r\})$ for all $r\in C$.

Theorem 3. Given a relatively consistent decision table $S=(U, C\cup D, V, f)$, attribute set $B\subseteq C$ is a relatively reduct of condition attribute set C if and only if $\mathrm{H}(D|B)=\mathrm{H}(D|C)$, and the attribute set B is relatively independent.

From the above 3 theorems we can find that the definition of reduct of a relatively consistent decision table in the information view is equivalent to its definition in the algebra view. Unfortunately, it will be different for a relatively inconsistent decision table. Now, let's have a look at the definitions of reduct and attribute core for a relatively inconsistent decision table in the information view.

Def. 5. Given a decision table $S=(U, C\cup D, V, f)$, attribute set $B\subseteq C$ is a relatively reduct of condition attribute set C if and only if

(1) $\mathrm{H}(D|B)=\mathrm{H}(D|C)$, and

(2) $\mathrm{H}(\{d\}|B)\ne\mathrm{H}(\{d\}|B-\{r\})$ for any attribute $r\in B$.

Def. 6. $\mathrm{CORE}_Q(P)=\cap\mathrm{RED}_Q(P)$ is called the Q-core of attribute set P.