Chapter 22

An Evaluation Method of Relative Reducts Based on Roughness of Partitions

Yasuo Kudo
Muroran Institute of Technology, Japan

Tetsuya Murai
Hokkaido University, Japan

ABSTRACT

This paper focuses on rough set theory which provides mathematical foundations of set-theoretical approximation for concepts, as well as reasoning about data. Also presented in this paper is the concept of relative reducts which is one of the most important notions for rule generation based on rough set theory. In this paper, from the viewpoint of approximation, the authors introduce an evaluation criterion for relative reducts using roughness of partitions that are constructed from relative reducts. The proposed criterion evaluates each relative reduct by the average of coverage of decision rules based on the relative reduct, which also corresponds to evaluate the roughness of partition constructed from the relative reduct,

INTRODUCTION

In rough set theory (Pawlak, 1982; Pawlak, 1991), set-theoretical approximation of concepts and reasoning about data are the two main topics. In the former, lower and upper approximations of concepts and their evaluations are the main topics. Accuracy, quality of approximation, and quality of partition are well-known criteria in evaluation of approximations; these criteria are based on the correctness of the approximation. However, the roughness of the approximation is not explicitly treated in these criteria. In reasoning about data, the relative reduct is one of the most important concepts for rule generation based on rough set theory, and many methods for exhaus-

DOI: 10.4018/978-1-4666-1743-8.ch022

tive or heuristic calculation of relative reducts have been proposed (Bao, 2004; Guan, 1998; Heder, 2008; Hu, 2008; Hu, 2003; Ślęzak, 2002;, Pawlak, 1991; Skowron & Rauszer, 1992; Xu, 2008; Xu, 2007; Zhang, 2003). As an evaluation criterion for relative reducts, the cardinality of a relative reduct, i.e., the number of attributes in the relative reduct, is typical and is widely used (for example, in (Heder, 2008; Hu, 2008; Hu, 2003; Xu, 2008; Zhang, 2003)). In addition, other kinds of criteria related to evaluation of partitions are also considered with respect to the following evaluation functions: a normalized decision function generated from a relative reduct *B* (Ślęzak, 2000), the information entropy H(*B*) of *B* (Ślęzak, 2002), and the number of decision rules induced from *B* (Wróblewski, 2001).

In this paper, we consider evaluating relative reducts based on the roughness of partitions constructed from them. The outline of relative reduct evaluation we propose is:

"Good" relative reducts = relative reducts that provide partitions with approximations as rough and correct as possible.

In this sense, we think that evaluation of relative reducts is strictly concerned with evaluation of roughness of approximation.

The paper is structured as follows. First, we review the foundations of rough set theory as background for this paper. Then, we derive some properties related to roughness of partition and the average coverage of decision rules, and propose an evaluation criterion of relative reducts based on roughness of partition. We also demonstrate the proposed method for evaluating relative reducts. Finally, we discuss the results of this paper and present our conclusions.

ROUGH SET

We review the foundations of rough set theory as background for this paper. The contents of this section are based on (Polkowski, 2002).

In rough set data analysis, objects as targets of analysis are illustrated by a combination of multiple attributes and their values and is represented by the following decision table:

$$\left(U, C, d\right),$$

where U is the set of objects, C is the set of condition attributes such that each attribute $a \in C$ is a function $a : U \rightarrow V_a$ from U to the value set V_a of a, and d is a function $d : U \rightarrow V_d$ called the decision attribute.

The indiscernibility relation R_B on U with respect to a subset $B \subseteq C$ is defined by

$$(x, y) \in R_B \Leftrightarrow a(x) = a(y), \forall a \in B. \tag{1}$$

It is easy to confirm that the indiscernibility relation R_B is an equivalence relation on U. The equivalence class $[x]_B$ of $x \in U$ by R_B is the set of objects which are not discernible with x even though they use all attributes in B.

Any indiscernibility relation provides a partition of U. We denote the quotient set of U, i.e., a partition of U, with respect to an equivalence relation R by U / R. In particular, the partition **D**$=\{D_1, \ldots, D_m\}$ provided by the indiscernibility relation R_d with respect to the decision attribute d is called the set of decision classes.

For any decision class $D_i (1 \leq i \leq m)$, the lower approximation $\underline{B}(D_i)$ and the upper approximation $\overline{B}(D_i)$ of D_i with respect to the indiscernibility relation R_B are defined as follows, respectively:

$$\underline{B}\left(D_i\right) = \left\{x \in U \,\middle|\, [x]_B \subseteq D_i\right\}, \tag{2}$$

$$\overline{B}\left(D_i\right) = \left\{x \in U \,\middle|\, [x]_B \cap D_i \neq \varnothing\right\}. \tag{3}$$

A pair $(\underline{B}(D_i), \overline{B}(D_i))$ is called a rough set of D_i with respect to R_B.