

# A Rough Set-Based Prediction Approach to Medical Application

Jia-Yuarn Guo<sup>1</sup>, \*Chun Pei<sup>2</sup>, Yu-Tung Dai<sup>3</sup>, Han Hsin Chou<sup>4</sup>,  
Chung-hsiu Ku<sup>5</sup>

<sup>1</sup>Dept. of Industrial Engineering and Management,

<sup>2</sup>The Graduate School of Gerontic Technology and Service Management,  
Nan Kai Institute of Technology

<sup>3</sup>Dept. of Occupational Safety and Health, Chung Haw College of Medical

<sup>4</sup>Department of Logistics and Shipping Management, Kainan University

<sup>5</sup>Ordnance Readiness and Development Center

Correspondence: Chun Pei, pantonpei@hotmail.com

## Introduction

It is noted that due to the typically huge size of today's information systems, real-world data tend to be incomplete due to missing some values. Hence, discover knowledge from incomplete information systems has received more and more attention in recent years.

The rough sets theory, proposed by Pawlak, provides a natural method to cope with incomplete or inconsistent information which has been the main impediment to the classification and rule induction of objects. Several techniques have been developed to extract decision rules from an incomplete information system. A key factor among them is using different methods to manage the missing data (unknown values). The simplest is removing the objects with unknown values. Other simple methods include replacing missing values with possible values calculated by statistics analysis. These methods all try to transform an incomplete system into a complete system by smoothing or extending the data. Other groups of techniques deal with the incomplete systems without changing the size of the data sets or making assumption of the missing values. These methods intend to induce every certain rule directly from the original data sets. Like the second group, our approach uses a rule generation algorithm to induce all certain rules and possible rules from the original incomplete data. No matter what the missing values will be, they won't affect the induction rules.

Compared with most of the classification and rule induction methods which induce knowledge from sets approximation concepts, the rule induction technique proposed in this paper applies a new modified rule-reduct generation algorithm (MRGI) and rule induction program (RIPI) to mine knowledge directly from the minimal set of decision rules (rule-reduct).

## Conclusions

In this article, a new rule-generation algorithm (MRGI) based on rough set theory has been proposed to generate the minimal set of rule-reduct which also represent the certain rules from incomplete information system. Some

examples have been illustrated to demonstrate the validation of the methods we presented.

Compared with other complex rough sets based rule induction approaches, MRGI present not only the capacity to deal with uncertainty or inconsistent information but the high ability to generate concise and simple rules as well, and makes it easier to analyze and induce knowledge from large incomplete information systems. Without changing the size of the original system or adding possible values to the null values, the technique proposed in this paper generate knowledge rules directly from the original incomplete information systems.

We strongly believe that the rule induction technique presented in this paper will provide an efficient access for rule generation and knowledge induction, and will play a competent role to complement other existed rule induction Techniques for incomplete information systems.

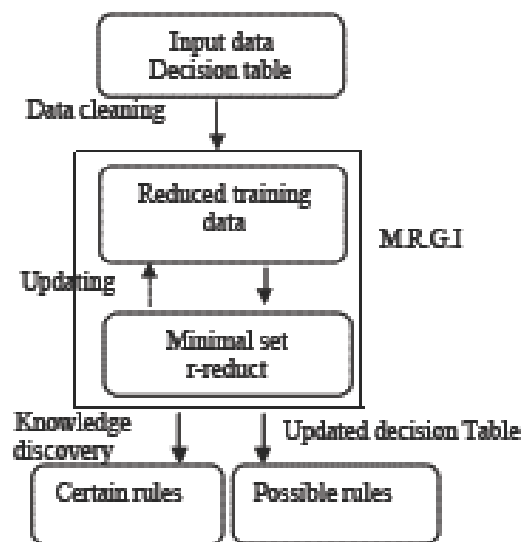


Fig.1 Mechanism of RGIPI including MRGI

