Chapter 9
Rule Extraction and Rule Evaluation Based on Granular Computing

Jiye Liang  
Shanxi University, P.R. China  

Yuhua Qian  
Shanxi University, P.R. China  

Deyu Li  
Shanxi University, P.R. China

ABSTRACT

In rough set theory, rule extraction and rule evaluation are two important issues. In this chapter, the concepts of positive approximation and converse approximation are first introduced, which can be seen as dynamic approximations of target concepts based on a granulation order. Then, two algorithms for rule extraction called MABPA and REBCA are designed and applied to hierarchically generate decision rules from a decision table. Furthermore, to evaluate the whole performance of a decision rule set, three kinds of measures are proposed for evaluating the certainty, consistency and support of a decision-rule set extracted from a decision table, respectively. The experimental analyses on several decision tables show that these three new measures are adequate for evaluating the decision performance of a decision-rule set extracted from a decision table in rough set theory. The measures may be helpful for determining which rule extraction technique should be chosen in a practical decision problem.

1. INTRODUCTION

Granular computing (GrC) is a new active area of current research in artificial intelligence, and is a new concept and computing formula for information processing. It has been widely applied to many branches of artificial intelligence such as problem solving, knowledge discovery, image processing, semantic web services [18, 29-31, 37].

DOI: 10.4018/978-1-60566-324-1.ch009
As follows, for our further development, we briefly review research on GrC. In 1979, the problem of fuzzy information granules was introduced by Zadeh. Then, in [47, 48] he introduced a concept of granular computing, as a term with many meanings, covering all the research of theory, methods, techniques and tools related to granulation. He identified three basic concepts that underlie the process of human cognition, namely, granulation, organization, and causation. “Granulation involves decomposition of whole into parts, organization involves integration of parts into whole, and causation involves association of causes and effects”. Some authors [37, 45] examined granular computing in connection with the theory of rough sets. Yao [45, 46] suggested the use of hierarchical granulations for the study of stratified rough set approximations and studied granular computing through using neighborhood systems. The theory of quotient space had been extended into the theory of fuzzy quotient space based on fuzzy equivalence relation [50]. Liang and Shi [14, 15] established the relationship among knowledge granulation, information entropy, granularity measure and rough entropy in information systems. Liang and Qian [17] developed an axiomatic approach to knowledge granulation in information systems. Granular computing mainly has three important models: (1) computing with words, (2) rough set theory, and (3) quotient space theory. As applications of granular computing, in this chapter, we focus on rule extraction and rule evaluation based on granular computing.

In the view of granular computing, a target concept described by a set is always characterized via the so-called upper and lower approximations under static granulation in rough set theory, and a static boundary region of the concept is induced by the upper and lower approximations [22-25]. However, through this mechanism, we cannot improve the approximation measure of a rough set and more clearly analyze the approximation structure of a rough set [34, 38]. Based on these conclusions, one objective of this chapter is to establish two kinds of structures of the approximation of a target concept by introducing a notion of a granulation order, called positive approximation and converse approximation, and to apply them to rule extracting from decision tables.

Generally speaking, a set of decision rules can be generated from a decision table by adopting any kind of rule extracting methods. In recent years, how to evaluate the decision performance of a decision rule has become a very important issue in rough set theory. In [3], based on information entropy, Duntsch suggested some uncertainty measures of a decision rule and proposed three criteria for model selection. For a decision rule set consisting of every decision rule induced from a decision table, three parameters are traditionally associated: the strength, the certainty factor and the coverage factor of the rule [32, 33]. In many practical decision problems, we always adopt several rule-extracting methods for the same decision table. In this case, it is very important to check whether or not each of the rule-extracting approaches adopted is suitable for the given decision table. In other words, it is desirable to evaluate the decision performance of the decision-rule set extracted by each of the rule-extracting approaches. This strategy can help a decision maker to determine which of rule-extracting methods is preferred for a given decision table. However, all of the above measures for this purpose are only defined for a single decision rule and are not suitable for evaluating the decision performance of a decision-rule set. There are two more kinds of measures in the literature [26-28], which are approximation accuracy for decision classification and consistency degree for a decision table. Although these two measures, in some sense, could be regarded as measures for evaluating the decision performance of all decision rules generated from a complete decision table, they have some limitations. For instance, the certainty and consistency of a rule set could not be well characterized by the approximation accuracy and consistency degree when their values reaches zero. As we know, when the approximation accuracy or consistency degree is equal to zero, it is only implied that there is no decision rule with the certainty of one in the complete deci-
Related Content

Functional Genomics Applications in GRID
[www.igi-global.com/chapter/functional-genomics-applications-grid/64521?camid=4v1a](www.igi-global.com/chapter/functional-genomics-applications-grid/64521?camid=4v1a)

Mining Environmental Data in the ADMIRE Project Using New Advanced Methods and Tools
Ondrej Habala, Martin Šeleng, Viet Tran, Branislav Šimo and Ladislav Hluchý (2012). *Technology Integration Advancements in Distributed Systems and Computing* (pp. 296-308).
[www.igi-global.com/chapter/mining-environmental-data-admire-project/64455?camid=4v1a](www.igi-global.com/chapter/mining-environmental-data-admire-project/64455?camid=4v1a)

A Method of 3-D Microstructure Reconstruction in the Simulation Model of Cement Hydration
[www.igi-global.com/article/method-microstructure-reconstruction-simulation-model/47209?camid=4v1a](www.igi-global.com/article/method-microstructure-reconstruction-simulation-model/47209?camid=4v1a)

Individual Client Strategies for Active Control of Information-Driven Service Construction in IoT-enabled Smart Spaces
[www.igi-global.com/article/individual-client-strategies-for-active-control-of-information-driven-service-construction-in-iot-enabled-smart-spaces/226960?camid=4v1a](www.igi-global.com/article/individual-client-strategies-for-active-control-of-information-driven-service-construction-in-iot-enabled-smart-spaces/226960?camid=4v1a)