Chapter 13
Modelling Classification by Granular Computing

Yan Zhao
University of Regina, Canada

ABSTRACT

A granular structure includes granules, levels, hierarchies and multiple hierarchies. Classification can be modelled by granular computing regarding these components. More specifically, classification tasks can be understood as a search in a certain search space represented by a granule network. This chapter discusses the basic components of a granular structure, followed by the modelling of classification in terms of these components. The top-down, bottom-up strategies for searching classification solutions within different granule networks are discussed.

1. INTRODUCTION

Granular Computing (GrC), as a newly developed computing paradigm, is studied by researchers in computational intelligence to explore varying levels of granularity in problem solving and information management (Bargiela & Pedrycz, 2002; Bargiela & Pedrycz, 2007; Lin, 2003; Lin et al., 2002, Inuiguchi et al., 2003; Polkowski & Artiemjew, 2007; Yao, 2007; Yao, 2006; Yao, 2007; Yao & Zhou, 2007; Zadeh, 1997; Zadeh, 1998). A granule is a collection of entities that can represent basic knowledge and concepts of human intelligence, and thus is regarded as the primitive notion of granular computing. A family of granules forms a granulated view at a certain level of resolution or scale. Granules at different levels are linked together by an order relation to compose a hierarchical structure, which provides a structured description of a system or an application under consideration. For a problem, one may need to construct multiple hierarchies for various descriptions and interpretations. Granular computing provides
a systematic and effective means for building conceptual models, organizing and discovering knowledge, as well as solving real-world problems by practical heuristics and strategies.

Based on granular computing, a concept can be described by a granule directly or be defined by a logical formula. Once the basic granules and formulas are properly identified, one is able to investigate the relations among concepts and define computational operations upon concepts (Yao, 2006). The relationships, such as sub- and super-concepts, along with disjointed and overlapping concepts, can be conveniently expressed in forms of rules, with quantitative measures to indicate the strength. Rule mining, therefore, can be viewed as a process of forming concepts and finding relationships between concepts in terms of granules and their formulas (Yao & Yao, 2002).

Classification, also known as supervised learning, is a common method used in data mining, machine learning and pattern recognition (Breiman et al., 1984; Cendrowska, 1987; Cestnik et al., 1987; Clark & Niblett, 1989; Grzymala-Busse, 1992; Brzymala-Busse, 2005; Michalski & Larson, 1980; Mitchell, 1987; Pal, 2004; Quinlan, 1983; Quinlan, 1992; Zhong et al., 2001). For classification tasks, concept formation involves the identification of granules and the description of granules regarding their classes, and concept relationship identification involves the connections of granules regarding their classes. The classification problem can be properly modelled by granular computing theory.

This chapter investigates classification tasks with respect to the basic components of granular computing, and discusses the classification of a granule, a granulation, in addition to the construction of a set of classification rules within a granule network. Two searching strategies, the top-down approach and the bottom-up approach, are studied to cope with two granule networks, a formula-based network and a granule-based network, respectively.

2. GRANULAR STRUCTURES

Granular computing exploits structures in terms of granules, granulations and hierarchies based on multilevel and multiview representations. In this section, we use an information table to demonstrate a concrete granular structure.

2.1 Information Tables and a Logic Language

An information table provides a convenient way to describe a finite set of objects, called a universe, by a finite set of attributes (Pawlak, 1982). It represents all available information and knowledge.

**Definition 1**

An information table is the following tuple:

\[ S = (U, At, \{V_a \mid a \in At\}, \{I_a \mid a \in At\}) \]

where \( U \) is a finite nonempty set of objects, \( At \) is a finite nonempty set of attributes, \( V_a \) is a nonempty set of values of \( a \in At \), \( I_a : U \rightarrow V_a \) is an information function.
Related Content

A Novel Hybrid Algorithm Based on Firefly Algorithm and Differential Evolution for Job Scheduling in Computational Grid

Healthcare Collaborative Framework Based on Web 2.0, Grid Computing and SOA
www.igi-global.com/chapter/healthcare-collaborative-framework-based-web/64495?camid=4v1a

A Global Survey on Data Deduplication
www.igi-global.com/article/a-global-survey-on-data-deduplication/210174?camid=4v1a

Serendipity Reloaded: Fair Loading in Event-Based Messaging
Daniel Cutting and Aaron Quigley (2010). *Principles and Applications of Distributed Event-Based Systems* (pp. 90-118).
www.igi-global.com/chapter/serendipity-reloaded-fair-loading-event/44397?camid=4v1a