Chapter 19
Dominance-Based Rough Set Approach to Granular Computing

Salvatore Greco
University of Catania, Italy

Benedetto Matarazzo
University of Catania, Italy

Roman Słowiński
Poznań University of Technology, Poland

ABSTRACT

Dominance-based Rough Set Approach (DRSA) was introduced as a generalization of the rough set approach for reasoning about preferences. While data describing preferences are ordinal by the nature of decision problems they concern, the ordering of data is also important in many other problems of data analysis. Even when the ordering seems irrelevant, the presence or the absence of a property (possibly graded or fuzzy) has an ordinal interpretation. Since any granulation of information is based on analysis of properties, DRSA can be seen as a general framework for granular computing. After recalling basic concepts of DRSA, the article presents their extensions in the fuzzy context and in the probabilistic setting. This permits to define the rough approximation of a fuzzy set, which is the core of the subject. The article continues with presentation of DRSA for case-based reasoning, where granular computing based on DRSA has been successfully applied. Moreover, some basic formal properties of the whole approach are presented in terms of several algebras modeling the logic of DRSA. Finally, it is shown how the bipolar generalized approximation space, being an abstraction of the standard way to deal with roughness within DRSA, can be induced from one of the algebras modeling the logic of DRSA, the bipolar Brower-Zadeh lattice.

DOI: 10.4018/978-1-60566-324-1.ch019
INTRODUCTION

This article describes Dominance-based Rough Set Approach (DRSA) to granular computing and data mining. DRSA was first introduced as a generalization of the rough set approach for dealing with multiple criteria decision analysis, where preference order has to be taken into account. The ordering is also important, however, in many other problems of data analysis. Even when the ordering seems irrelevant, the presence or the absence of a property has an ordinal interpretation, because if two properties are related, the presence, rather than the absence, of one property should make more (or less) probable the presence of the other property. This is even more apparent when the presence or the absence of a property is graded or fuzzy, because in this case, the more credible the presence of a property, the more (or less) probable the presence of the other property. Since the presence of properties, possibly fuzzy, is the basis of any granulation, DRSA can be seen as a general framework for granular computing.

The article is organized as follows. First, we introduce DRSA in the context of decision making. After presenting the main ideas sketching a philosophical basis of DRSA and its importance for granular computing, the article introduces basic concepts of DRSA, followed by their extensions in the fuzzy context and in the probabilistic setting. This prepares the ground for defining the rough approximation of a fuzzy set, which is the core of the subject. It is also explained why the classical rough set approach is a specific case of DRSA. The article continues with presentation of DRSA for case-based reasoning, where granular computing based on DRSA has been successfully applied. Finally, some basic formal properties of the whole approach are presented in terms of several algebras modeling the logic of DRSA. Moreover, we show how the bipolar generalized approximation space, being an abstraction of the standard way to deal with roughness within DRSA, can be induced from one of the algebras modeling the logic of DRSA, the bipolar Brower-Zadeh lattice.

DOMINANCE-BASED ROUGH SET APPROACH

This section presents the main concepts of the Dominance-based Rough Set Approach (DRSA) (for a more complete presentation see, for example, Greco et al. (1999, 2001a, 2004b,c, 2005a); Słowiński et al. (2005)).

Information about objects is represented in the form of an information table. The rows of the table are labeled by objects, whereas columns are labeled by attributes and entries of the table are attribute-values. Formally, an information table (system) is the 4-tuple \( S = \langle U, Q, V, \varphi \rangle \), where \( U \) is a finite set of objects, \( Q \) is a finite set of attributes, \( V = \bigcup_{q \in Q} V_q \) and \( V_q \) is the value set of the attribute \( q \), and \( \varphi : U \times Q \rightarrow V_q \) is a total function such that \( \varphi(x, q) \in V_q \) for every \( q \in Q, x \in U \), called an information function (Pawlak, 1991). The set \( Q \) is, in general, divided into set \( C \) of condition attributes and set \( D \) of decision attributes.

Condition attributes with value sets ordered according to decreasing or increasing preference of a decision maker are called criteria. For criterion \( q \in Q \), \( \succeq_q \) is a weak preference relation on \( U \) such that \( x \succeq_y \) means “\( x \) is at least as good as \( y \) with respect to criterion \( q \)”. It is supposed that \( \succeq_q \) is a