Dominance Relations in Rough Sets Approximations for Assessing Students Knowledge

SYLVIA ENCHEVA
Stord/Haugesund University College
Faculty of Technology, Business and Maritime Sciences
Bjørnsonsg. 45, 5528 Haugesund
NORWAY
sbe@hsh.no

Abstract: A lot of research has been done with respect to automated evaluation of students’ knowledge. In this article we apply dominance relations in rough sets approximations for assessing students knowledge.

Key–Words: Decision support services, knowledge evaluation, intelligent systems, automated tests

1 Introduction

Extracting knowledge from huge databases often requires application of classifications rules that can be automatically discovered from the database. This can be achieved using methods from rough sets theory. In this case the database will be first divided into mutually exclusive elementary sets. A rough logic classifier, operating with lower and upper approximations, will be further employed for determining the class of the objects.

Intelligent tutoring systems (ITS) are generally understood as computer systems facilitating learning by supplying automated instructions and feedback to students. A lot of work has been done since the appearance of the first ITSs in the late 1970s. The majority of such systems relied heavily on domain experts knowledge and skills for customizing hints and feedback. According to [22] cognitive and constrained tutors are based on rules established by experts. Current ITSs are expected to both encourage students creative work and point direct them to a correct solution if necessary. [7] and [9] employ ontologies in designing ITSs due to ontologies’ abilities to support reasoning. This approach needs a lot of work for encoding actual knowledge into ontology. Learning correlations across domains and modalities were investigated by [18]. Partially correct responses are recognized by the work of [10] and [21]. Student knowledge models are built by examples in [23]. An agent based tool is used to predict students’ knowledge reading from log data. Machine learning was used by [3] to predict how much time students needed for solving problems.

New users often lack in-depth understanding of the subject they study and cannot really judge to which extend a particular help function is facilitating the learning process. What they actually express is their overall likes of the tool. Among other factors that influence their responses are friends’ opinions on the matter, student’s degree of interest in that subject, honesty, i.e. are their responses anonymous or a sender can be tracked down and so on.

Evaluation of students’ knowledge has been a subject of special interest to various research communities. In this respect automated tests appear to be among the most popular ways of providing immediate feedback to both students and test designers. A test designer has to give serious considerations to logical reasonings involved in the process decision making. Use of Boolean logic limits system’s responses to true or false and cannot therefore recognize other occurrences like f. ex partially correct or incomplete answers. Boolean logic appears to be quite sufficient for most everyday reasonings, but it is certainly unable to provide meaningful conclusions in presence of inconsistent and/or incomplete input, [12].

Researchers from different field have been working together for many years in order to further develop methods for knowledge transmission and knowledge evaluation. The process of delivering hints in an intelligent tutoring system has been discussed in [34]. A taxonomy for automated hinting is developed in [31]. The role of hints in a Web based learning systems is considered in [6].

The aim of this paper is to present an approach for following students’ progress in obtaining new knowledge based on rough sets approximations. The responses to automated tests of each individual learner

ISOB: 978-960-474-273-8 162
are placed in appropriate rough sets approximations. The resulting path provides strong indication about the current level of learning outcomes and points possible inconsistencies related to tests’ contents.

The theory of rough sets [27] is designed to handle imprecise and vague data among other things. Data reduction in qualitative analysis can be obtained applying dominance-based rough set approach, [13]. Its application does not depend on additional information about the data, can extract dependencies among attributes and present them in a natural language. Further more, it has been used by researchers in assessment of bankruptcy risk assessment [30], business failure prediction [4], and accident preventions [32].

The rest of the paper is organized as follows. Section 2 contains definitions of terms used later on. Section 3 explains how to combine personal responses. Section 4 contains the conclusion of this work.

2 Background

Let $P$ be a non-empty ordered set. If $\sup\{x, y\}$ and $\inf\{x, y\}$ exist for all $x, y \in P$, then $P$ is called a lattice, [8]. In a lattice illustrating partial ordering of knowledge values, the logical conjunction is identified with the meet operation and the logical disjunction with the join operation.

From classical stand point of view a concept is well defined by a pair of intention and extension. Existence of well defined boundaries is assumed and an extension is uniquely identified by a crisp set of objects. In real life situations one has to operate with concepts having grey/gradual boundaries, like f. ex. partially known concepts, [35], undefinable concepts, and approximate concepts, [19].

Rough Sets were originally introduced in [26]. The presented approach provides exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space. An approximation space is a pair $A = (U, R)$, where $U$ is a set called universe, and $R \subseteq U \times U$ is an indiscernibility relation.

Equivalence classes of $R$ are called elementary sets (atoms) in $A$. The equivalence class of $R$ determined by an element $x \in U$ is denoted by $R(x)$. Equivalence classes of $R$ are called granules generated by $R$.

The following definitions are often used while describing a rough set $X$, $X \subseteq U$:

- the $R$-upper approximation of $X$, $\bar{X} := \bigcup_{x \in U_1} \{R(x) : R(x) \cap X \neq \emptyset\}$
- the $R$-lower approximation of $X$, $\underline{X} := \bigcup_{x \in U_1} \{R(x) : R(x) \subseteq X\}$
- the $R$-boundary region of $X$, $R\Lambda_R(X) := \bar{X} - \underline{X}$

Attributes reduction stands for removal of attributes that do not effect the primary system.

Suppose $\phi$ and $\psi$ are logical formulas representing conditions and decisions respectively. The certainty factor $\mu$ of a decision rule $\phi \rightarrow \psi$ is $\mu(\phi, \psi) = \frac{\text{card}(\phi \cap \psi)}{\text{card}(\phi)}$, where $\mu(\phi, \psi) \in [0, 1]$. The rule is deterministic if $\mu(\phi, \psi) = 1$ and nondeterministic otherwise.

Rough sets attribute analysis is usually applied in the process of establishing the relative importance of an attribute and consequently remove it if it contains redundant information.

A data table is the four-tuple $S = (U, Q, V, f)$, where $U$ is a finite set of objects (universe), $Q = \{q_1, q_2, ..., q_m\}$ is a finite set of attributes, $V_q$ is the domain of the attribute $q$, $V = \bigcup_{q \in Q} V_q$, and $f : U \times Q \rightarrow V$ is a total function such that $f(x, q) \in V_q$ for each $q \in Q, x \in U$, called the information function.

If the objects in the data table are classification examples, then the set of attributes is divided into condition attributes and a decision attribute. In multicriteria classification, condition attributes are criteria. The notion of a criterion involves a preference order in its domain, while the domains of attributes, usually considered in machine discovery, are not preference-ordered.

A decision attribute $d$ makes a partition of $U$ into a finite number of classes $Cl = \{Cl_t, t \in T\}$ and $T = \{1, ..., n\}$. Each $x \in U$ belongs to one and only one class, $Cl_t \in Cl$. The classes from $Cl$ are preference-ordered according to increasing order of class indices, that is, for all $r, s \in T$ such that $r > s$, the objects from $Cl_r$ are preferred to the objects from $Cl_s$. In other words, the classes $Cl$ represent a comprehensive evaluation of the objects in $U$ : the worst objects are in $Cl_1$, the best objects are in $Cl_n$, and the other objects belong to the remaining classes $Cl_r$, according to an evaluation improving with the index $r \in T$.

In multicriteria classification, due to the preference order in the set of classes $Cl$, the sets to be approximated are not the particular classes but upward unions and downward unions of the classes.

Given $P \subseteq C$ and $x \in U$, the ”granules of knowledge” used in DRSA for approximation of the units $Cl^{\leq}_t$ and $Cl^{\geq}_t$ are:

- A set of objects dominating $x$, called $P - \text{dominating set}$, $D^P_f(x) = \{y \in U : yD_f x\}$
A set of objects dominated by \( x \), called \( P - \text{dominated set}, D_P^-(x) = \{y \in U : xD_Py\} \)

With respect to the boundary, there holds a property of identity saying that if an object is doubtful with respect to \( Cl^<_T \), it is doubtful also with respect to \( Cl^<_T \).

\( P \)-lower approximations of unions of classes represent certain knowledge provided by criteria from \( P \subseteq C \), while \( P \)-upper approximations represent possible knowledge and the \( P \)-boundaries contain doubtful knowledge.

For every \( P \subseteq C \), the quality of approximation of multicriteria classification \( Cl \) by a set of criteria \( P \) is defined as the ratio between the number of \( P \)-correctly classified objects and the number of all the objects in the data table.

Each minimal subset \( P \subseteq C \) such that \( \omega_P (Cl) = \omega_C (Cl) \) is called a reduct of \( Cl \) and is denoted by \( RED_{Cl} \). Let us remark that a data table can have more than one reduct. The intersection of all reducts is called the core and is denoted by \( CORE_{Cl} \). Criteria from \( CORE_{Cl} \) cannot be removed from the data table without deteriorating the knowledge to be discovered. This means that in set \( C \) there are three categories of criteria:

- **Indispensable** criteria included in the core
- **Exchangeable** criteria included in some reducts but not in the core
- **Redundant** criteria being neither indispensable nor exchangeable, thus not included in any reduct.

One of three induction strategies can be adopted to obtain a set of decision rules:

- Generation of a minimal description, that is, a minimal set of rules
- Generation of an exhaustive description, that is, all rules for a given data table
- Generation of a characteristic description, that is, a set of rules each covering relatively many objects, however, all together not necessarily all objects from \( U \)

### 2.1 Assessment

A method enabling the instructor to do a post-test correction to neutralize the impact of guessing is developed in [15]. The theory and experience discussed in the above listed literature was used while developing our assessment tools.

A personalized intelligent computer assisted training system is presented in [25]. An intelligent tutoring system that uses decision theory to select the next tutorial action is described in [20]. A model for detecting student misuse of help in intelligent tutoring systems is presented in [2]. An investigation of whether a cognitive tutor can be made more effective by extending it to help students acquire help-seeking skills can be found in [17].

A proliferation of hint abuse (e.g., using hints to find answers rather than trying to understand) was found in [1] and [17]. However, evidence that when used appropriately, on-demand help can have a positive impact on learning was found in [33].

A level-based instruction model is proposed in [24]. A model for student knowledge diagnosis through adaptive testing is presented in [14]. An approach for integrating intelligent agents, user models, and automatic content categorization in a virtual environment is presented in [29].

The Questionmark system [38] applies multiple response questions where a set of options is presented following a question stem. The final outcome is in a binary form, i.e. correct or incorrect because the system is based on Boolean logic [11].

### 3 Responses

This section is devoted to classifying various responses from students to tests applying rough sets approximations.

Assessment of students’ understanding of a concept is proposed employing multiple choice tests. The system is designed in a way that a new trial brings different questions and/or answer combinations but related to the same concept. A test related to a particular concept can be taken several times. In order to obtain a higher degree of certainty in the decision process on whether a concept is sufficiently understood we involve three different questions related to that concept. This gives an opportunity to the student to apply his/her understanding in different situations and decreases the chances of ‘just a lucky guess’.

A test contains a predetermined number of questions. Each question is followed by alternative answers and a student can choose exactly one of them or skip that question. This implies of course one correct answer only. Tests with several correct answers are a subject of another model.

A change in the understanding status can be observed when test results of a student appear in different rough set approximations.

Students receive automated advices on how to proceed and what further actions should be taken based on their responses. These advises can vary from
Table 1: A decision table

<table>
<thead>
<tr>
<th>O 1</th>
<th>O 2</th>
<th>O 3</th>
<th>O 4</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>ψ</td>
<td>ψ</td>
<td>ω</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>ω</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>ω</td>
<td>P</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>ω</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>ω</td>
<td>P</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>ω</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>ω</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>ω</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>ω</td>
<td>ω</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>ω</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>ω</td>
<td>P</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

A correct class has been predicted 65% of the time with certainty 1. The attribute analysis implies that the last is the least important one.

4 Conclusion

This work is related to evaluation of knowledge based on information about preference order in the domains of criteria and among decision classes. Applying an extension of the rough set theory from multi-attribute classification problems to multicriteria sorting problems allows finer tuning of final decisions.

References:


[38] http://www.leeds.ac.uk/perception/v4_mrq.html