Ain Shams University
Faculty of Engineering
Computer and Systems Engineering Department

Knowledge Discovery in Intelligent Tutoring Systems

A Thesis
Submitted in Partial Fulfillment of the Requirements of the
Degree of Master of Science in Electrical Engineering
(Computer and Systems Engineering Department)

Submitted By

Sally Sameh Abd El Ghaffar Attia

B. Sc., Electrical Engineering
(Computer and Systems Engineering)
Ain Shams University, 1998

Supervised by

Prof. Dr. Hani M. K. Mahdi

Dr. Hoda K. Mohamed

Cairo, Egypt

November, 2004

Faculty of Engineering

Examiners Committee

Name : Sally Sameh Abd El Ghaffar Attia

Thesis : Knowledge Discovery in Intelligent Tutoring Systems

Degree : Master of Science in Electrical Engineering

Name, Title and Affiliation	**Signature**
1. Prof. Dr. Mohamed Zaki Abdel-Megeed Head of Systems Engineering and Computer Dept. Faculty of Engineering, Al-Azhar University, Cairo	……………
2. Prof. Dr. Abdel-Moneim A. Wahdan Head of Computer and Systems Engineering Dept. Faculty of Engineering, Ain Shams University, Cairo	……………
3. Dr. Hoda Korashi Mohamed Computer and Systems Engineering Dept. Faculty of Engineering, Ain Shams University, Cairo (Supervisor)	……………

Date:21/11/2004

Acknowledgements

I owe a special acknowledgment to Prof. Hani M. K. Mahdi for spending a lot of his precious time reading my work, writing his comments, suggesting improvements and discussing technical points with me specially during preparing my first paper with him and Dr. Hoda to be published in ICEEC'04 in Cairo.

I am deeply grateful to Dr. Hoda K. Mohammad for supporting me during the process of writing the thesis, reading it several times and giving me her valuable comments and advice and helping me to finish the thesis in time.

I will never be able to thank my parents enough for supporting me during all my life, especially during the last period of preparing the thesis. They gave me hope and help when I really needed them and had to bear a lot so that I could finish this work. I would like to dedicate this work to them and also to my sister who helped me in drawing the figures.

I am very grateful to my husband for helping me processing the thesis whenever I needed his help and for bearing that I was always busy during the last period. I would also like to dedicate this work to my lovely two kids, Youssef and Nour.

Abstract

Sally Sameh Abd El Ghaffar Attia

Knowledge Discovery in Intelligent Tutoring Systems

Master of Science dissertation

Ain Shams University, 2004

Knowledge discovery through data mining aims at searching for meaningful information like patterns and rules in large volumes of data. Our objective is to discover useful knowledge in Intelligent Tutoring Systems (ITS). These are tutoring systems which offer the ability to respond to individualized student needs. A major challenge for today's schools is to create individualized instruction in inexpensive, expandable ways. To achieve this goal, educational systems must adapt to each student by mining student data to determine what a student knows and does not know.

An experiment was conducted over a tutorial for binary relations. Students' answers to questions at the end of the lesson were collected. Data mining using Rough Sets technique was implemented to extract important rules from the data (students' answers) and hence the student can be directed to which parts of the lesson he should take again, thus helping to adopt the tutoring systems to each student individual needs and hence the tutoring systems are called intelligent. Using this knowledge, teaching systems can guide students in the learning process.

In this research, three different approaches of Rough Sets are applied to detect the decision rules. These approaches provide a powerful foundation to discover important structures in data. These approaches are unique in the sense that they only use the information given by the data and do not rely on other model assumptions. We found that, rough sets can be used to understand

large sets of student data, pinpointing problem areas for student learning.

The results obtained were in the form of decision rules that showed what concepts the student understood and which he did not understand depending on which questions he answered correct and which questions he answered wrong. Also some questions of the quizzes were found to be useless. It was concluded that data mining was able to extract some important patterns and rules from the students' answers which were hidden before and which are helpful to both the students and the experts.

In conclusion, this research has provided an in-depth analysis of the use of rough sets approaches as a mining tool in computer-based tutorials. The application of rough sets methods in diagnosing student misconceptions makes a contribution to the fields of computer-based education, fault-tolerant teaching, and data mining.

Keywords:

Knowledge Discovery; Data Mining; Intelligent Tutoring Systems; Rough Sets; Fault Tolerant Teaching.

Statement

This dissertation is submitted to Ain Shams University for the degree of Master of Science in Computer and Systems Engineering Department.

The work included in this thesis was out by the author at Computer and Systems Engineering Department, Ain Shams University.

No part of this thesis has been submitted for a degree or qualification at other university or institution.

Date : 21/ 11/ 2004

Signature :

Name : Sally Sameh Abd El Ghaffar Attia

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

As the world becomes more interconnected, information can be shared between people as never before. While this opens up exciting possibilities for communication and research, it also results in an overwhelming amount of information that each person must process every day in order to function. With each new person joining the global village, the need to develop automated methods to filter, sort, search, and share information in meaningful ways grows exponentially. Therefore, there is a big need for data mining in society today.

Data mining is a set of methods used as a step in the Knowledge Discovery (KD) process to distinguish previously unknown relationships, rules and patterns within large volumes of data. One of data mining tasks is description i.e. to describe databases in terms of patterns which human can understand and make use of [1].

Our schools are facing an even greater need for the benefits that data mining promises to deliver. Both teachers and students are faced with teaching and learning an ever-growing number of ideas all in the same school day as before. Ideally, as more issues are addressed, schools would add more teachers and classes to cover new topics. Instead, schools face teacher shortages, inadequate teacher training, and growing classroom sizes.

At the same time, the typical trend in universities is changing from young on-campus students to diverse professionals of all ages taking classes in the evenings, online, or even at a distance. Colleges and universities are now challenged to deliver quality service to students of widely ranging abilities and backgrounds across all distances. To rise to this challenge, information technology assisted education must pair individuals with expert instruction that is interactive, adaptive, and accessible. The promise of information technology assisted education is great, but research into methods for the delivery of education with information technology is still a rich, open field [2].

In our research, we are trying to mine databases resulting from Intelligent Tutoring Systems (ITS). These are Computer-based tutoring systems which achieve their intelligence by representing pedagogical decisions about how to teach as well as information about the learner. This allows for greater versatility by altering the system's interactions with students. Intelligent tutoring systems have been shown to be highly effective at increasing student's motivation and performance [3].

We are trying to investigate if the relatively new area of research, termed data mining and knowledge discovery, can be applied to educational problems to achieve our goal of individualized instruction, at hopefully much lower costs than traditional adaptive teaching systems that use knowledge-based models to understand and direct student learning.

Many excellent intelligent tutoring systems exist today. However, the majority of these intelligent tutoring systems require the construction of complex models that are applicable only to a specific tutorial in a specific field, necessitating a large number of experts to create and then test these models on students. Data mining and knowledge discovery, on the other hand, might be applied to the problem of understanding student knowledge, and using this understanding to direct knowledge remediation.

The goal of data mining in ITS is to automatically assess student knowledge of the concepts underlying a tutorial topic, and use this