Work in Progress - Bringing SOCRATES into Computer-Assisted Instruction

Stephen Zahorian, Radim Belohlavek, Scott Craver, Roy McGrann, and Lei Yu
Binghamton University, zahorian@binghamton.edu, rbelohla@binghamton.edu, scraver@binghamton.edu, mcgrann@binghamton.edu, lyu@binghamton.edu

Abstract - SOCRATES (Student Oriented Creative Resource for the Assessment and Teaching of Engineering Skills) will be an automated computer-based system for simulating the Socratic Method of learning in the subject area of Probability and Statistics. The goal is for the student to be an active partner in the learning process using directed self-reflections. In SOCRATES, a knowledge space is created from data gathered from expert practitioners and students. An automatic question poser is then used to lead students through the knowledge space by presenting them with additional questions.

We describe several methods for gathering data from domain experts (instructors and students). This dataset is then analyzed to produce a basis set of generic questions and actions that appear commonly within the tree structure of instructor responses. These various sources of data are combined and analyzed to identify prerequisite relationships and dependencies between the topics. We also show how the collected data can be used to guide an automated question poser. Student interactions with the system can be monitored to develop a model of how students learn.

Index Terms – Computer-assisted learning, Knowledge space.

INTRODUCTION

This paper describes the first step of a larger project to develop a computer-assisted instruction system based on a knowledge space. The specific objectives of the overall project are:

1. To formulate the structure of a knowledge space that is a model for student learning and that codifies domain knowledge regarding student learning.
2. To design an intelligent question poser to guide student learning.
3. To investigate strategies for efficiently guiding student learning in engineering.
4. To assess the benefits to student learning using the learning model implemented in this study.

I. Overview: Using a Knowledge Space as an Experimental Environment for Interactive Learning

The structure of our experimental environment is illustrated in Figure 1 below. The general idea is to create an objective knowledge space of an academic subject, using artificial intelligence methods and database search techniques. The specific academic subject chosen for a prototype system is statistics and probability theory.

FIGURE 1
SOCRATES: THE INTERACTIVE LEARNING ENVIRONMENT

This knowledge space is used to both assess student ability through a question poser and guide an interactive process of student learning. We have named the overall system SOCRATES: Student Organized Creative Resource for Assessment and Teaching of Engineering Skills.

The centerpiece of our system is the knowledge space, or knowledge map. It acts as a judgment simulator to assess student ability from question responses as would a teacher. All functions of the overall system depend on the effectiveness of the knowledge space representation, and therefore considerable effort and expertise are focused on a faithful representation of expert judgment. A judgment simulation vector space, an AI construct pioneered by Ossorio and Jeffrey and used to solve similar judgment problems [1, 2], is the primary mathematical tool used to derive the knowledge map. The knowledge map has two primary purposes: assessment and direction of future learning, both utilizing an interactive question poser. The poser will have the ability to generate questions, and, moreover, it will be able to generate questions relevant to specific areas of the knowledge map. An assessment engine will use the poser to estimate a student’s position in the map; a learning engine will use the poser to direct students away from that position, along an effective path towards a greater coverage and understanding of the subject.
II. Preliminary Work: Determining a Hierarchical Relation of Concepts in the Knowledge Map

In this paper we describe the development of a prototype system that will be used for developing a hierarchical structure embedded in the knowledge map. This structure models the interrelation of concepts involved in understanding probability. As a first step in this development, several students were interviewed to determine their knowledge of probability and to make a preliminary assessment of the interrelation of concepts.

Several methods are envisioned for collecting, analyzing, and organizing domain knowledge. The basic assumption is that knowledge maps can be created from patterns of student answers to questions. Four methods are under consideration: (1) evaluation of student responses to carefully chosen questions; (2) an evaluation of dependencies among questions, as judged by surveying experts; (3) an evaluation of expert opinions about the relationships of different topics to questions; and (4) a hierarchy of questions as suggested by experts, arranged as “Socratic response graphs.” The data from these three methods can be organized in matrix form and then analyzed with methods such as relational factor analysis, to determine knowledge map representations as given in Figure 1.

In the rest of this paper, we focus on the fourth method listed above, using an online data collection method. “Socratic response graphs” are artificial representations of Socratic dialogs, which can be built collaboratively by many visitors to a web site. This data can then be analyzed for patterns and relationships, or used directly in the automated learning system.

![Image](http://example.com/image.png)

**FIGURE 2**

The instructor view allows the browsing of a question’s response graph.

---

**SYSTEM DESIGN FOR SOCRATIC DIALOG DATA COLLECTION**

The SOCRATES system uses a web browser to display graphs of questions, and an interface for fast and convenient data entry. The system design uses an SQL database to hold the questions. These questions may be in one of three basic relationships:

- A question that is answered incorrectly may have a child question, whose exploration is intended to shed light on its parent question.
- A question that is answered correctly may have a next-in-sequence question. In its expected usage, a student who fails to answer a question is given an ordered sequence of sub-questions to consider.
- Multiple questions may exist in the same location in the graph as alternatives or revisions of one another.

Other data in this system follows a similar hierarchical pattern, which can all be represented using a very simple data model.

**SCIENTIFIC VALUE OF RECORDED Dialogs**

The data will be collected in various forms, for the four different basic methods described previously. Two ways of organizing data are as graphs and matrices. The analysis of such data will reveal relevant information regarding the relationships between various topics of a given subject of study, such as probability and statistics; help identify important hidden factors such as “technical calculation skills” or “abstract reasoning;” reveal relationships between subject topics and abstract factors; and will help map students into an abstract knowledge space [3, 4].

**CONCLUSIONS**

SOCRATES, a Student Oriented Creative Resource for assessment and training of engineering skills has been described. This system is intended as a test bed to investigate how students learn and ultimately as a tool for technology enhanced learning. A centerpiece of the system is a knowledge space hierarchical representation of a given topic such as probability theory and statistics. A web based online system has been developed and will be made available for instructors worldwide to collaboratively build Socratic dialogs which will then be analyzed for use with SOCRATES.

**REFERENCES**