Education Datamining in Ill-Defined Domains: The case of argument diagrams

M.Sc. Collin Lynch
University of Pittsburgh

When students form a diagram representation, annotate a text, or design a building they are solving ill-defined problems. When tackling ill-defined problems such as design or diagnostic tasks a solver must recharacterize the problem in order to make it solvable, clarifying or specifying missing information, implicit constraints, or potential choices. As a consequence student solutions are typically, often deliberately, unique. Reasonable domain experts often disagree about the quality of individual solutions. This novelty complicates the tutoring process making it difficult for tutors to assess student solutions, identify strengths or weaknesses, and give productive feedback. In many cases the solutions exploit background information or avenues of attack not familiar to the tutor.

In this talk I describe ongoing research into the use of educational datamining in legal education. In a series of studies at the University of Pittsburgh's School of Law we have deployed the LARGO program with law students both in and outside the classroom. LARGO is a diagram-based tutorial system for law. The system tutors students in arguing with tests and hypotheticals via a graph grammar. As part of these studies the diagrams were graded both by the LARGO system and by a set of law school faculty.

We have begun a series of studies focused on the induction of graph rules and the use of induced graph features in tutoring. Preliminary work has shown that some basic graph features are correlated with students' performance. In subsequent work, designed to aid the development of new hand-tooled rules, we used genetic programming to induce rules based upon the existing graph grammars. In future work, to be discussed in detail, we will apply other machine learning methods to both align our existing student assessments: graph grades assigned by experts; graph grades assigned by LARGO; and students' final exam grades. We will also align the graphs themselves and seek to identify salient comparisons between the student graphs.