Work Pattern Data Mining to Assist Students in Software Development Practices

Dale E. Parson
Department of Computer Science

ABSTRACT

Many course projects in computer science require students to create software programs. The process of compiling programs into executable forms, running them and checking output validity is often automated. In the current project we are extending this automation to collect data showing work patterns that include time of project start and completion, time and duration of work sessions, time of day in which students work, locality or distribution of a student’s efforts in a code base, and other work attributes. This project is collecting these data from a series of student programming courses, augmented with data from brief questionnaires, and will analyze them using modern data mining tools to discover patterns of course assignment requirements and student work habits that correlate with successful completion of projects and retention in the major. Student participation is voluntary. The project goal is to develop automated means for detecting productive patterns of student work habits for reinforcement and unproductive patterns for correction.

The PI recently taught Kutztown University’s first offering of a graduate course in data mining. The current project will apply the insights and software tools acquired in that course to correlate student work patterns with patterns of successful project and course completion. Information that is output from the tools includes statistical tables, rules and decision trees that illustrate inference paths from student work patterns to patterns of success. This output is useful in two ways. First, it illustrates correlations in statistical and symbolic terms for instructor understanding of student behavior. Second, it produces inference data structures that the PI plans to incorporate into an automated early warning system for future students. The project is employing undergraduate research assistants, and we anticipate co-publication of results at an international conference on computer science education.