

# Computer Science Colloquium

Thursday, February 28, 2013

Nguyen Engineering 4201

10:00am-11:00am

## Solving the Search for Source Code

Programmers frequently use keyword searches to find source code in large repositories. However, to do this effectively, programmers must specify keyword queries that capture implementation details of their desired code. I propose that code search should be about behavior, not about keywords.

In this talk, I will present an approach to code search that allows programmers to provide inputs and outputs that define the behavior of their desired code. This approach indexes source code repositories by symbolically analyzing the programs and program fragments and transforming them into constraints representing their behavior. Results are identified using an SMT solver, which, given an input/output specification and the constraint representation of a program fragment, determines if the fragment matches the desired behavior. While promoting code reuse, my approach enables reuse where it was not possible before: the constraints can be relaxed, identifying code that approximately matches the specification. Further, the solver can then guide the instantiation of the code to produce the desired behavior. I will illustrate the generality of the approach by showing its instantiation in subsets of three languages, the Java String library, Yahoo! Pipes mashups, and SQL select statements. I will conclude by sharing my vision for new research directions related to this semantic approach to code search.

## Kathryn Stolee

### Biography

Kathryn Stolee is a Ph.D. candidate and NSF Graduate Research Fellow in the Department of Computer Science and Engineering at the University of Nebraska-Lincoln. She has been awarded an ESEM Distinguished Paper Award and two departmental outstanding research awards. Her research is in software engineering with a focus on program analysis. Extracting useful information from software artifact repositories is a broad theme of her research, most recently through semantic code search.