The University of Alabama
Department of Computer Science
Colloquium Series Speaker

Nicholas Kraft
Clemson University

An Infrastructure to Support Interoperability in Reverse Engineering

Monday, March 5
11:00a.m., Houser 108

Abstract:
Reverse engineering tools are critical for controlling the high cost and risk of legacy system evolution. Wider adoption of these tools is a significant challenge to their increased effectiveness. Failure to adopt reverse engineering tools is caused in part by their lack of interoperability with common academic and industrial tools. This lack of interoperability hinders accessibility, reproducibility, and reusability of experimental results, and obstructs researchers' attempts to compare competing techniques. For example, Das has reported that it took his team at Microsoft Research several months to synchronize the outputs of tools implementing points-to analysis algorithms, and Murphy et al. have reported considerable difficulty comparing results from tools for extracting call graphs.

To address the problems of accessibility, comparability, reproducibility, and reusability, we have created an infrastructure that supports interoperability among reverse engineering tools and other software tools. In this talk, I will present the three major components of our infrastructure: (1) a hierarchy of schemas for low- and middle-level program representation graphs, (2) g4re, a tool chain for reverse engineering C++ programs, and (3) a repository of reverse engineering artifacts, including the previous two components, a test suite, and tools, GXL instances, and XSLT transformations for graphs at all levels of the hierarchy. In addition, I will present the results of two empirical studies that we performed using g4re. The studies investigate the space and time costs incurred by the infrastructure using a test suite of 12 popular, open source C++ applications and libraries that consist of approximately one million C++ lines of non-commented, non-preprocessed code.