

NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

C S S E M I N A R

Title: 1) Abstract Semantic Differencing via Speculative Correlation 2) Code Similarity via Natural Language Descriptions

Speaker: PhD Student Mr Nimrod Partush
Computer Science Department
Technion - Israel Institute of Technology

PhD Student Ms Meital Ben Sinai
Computer Science Department
Technion - Israel Institute of Technology

Date/Time: 23 March 2015, Monday, 02:00 PM to 03:30 PM

Venue: MR3, COM2-02-26

Chaired by: Dr Roychoudhury, Abhik, Professor, School of Computing
(abhik@comp.nus.edu.sg)

1) Abstract:

Understanding the semantic difference between two versions of a program is invaluable in the process of software development. In this talk we will discuss a new approach towards characterizing of the difference between infinite-state program versions, or establishing their equivalence.

Using abstraction to prove equivalence or find difference requires abstracting relationships between variables in the two programs. Towards that end, we use a correlating abstract domain to compute a sound approximation of these relationships which captures semantic difference. This approximation can be computed over any interleaving of the two programs. We show that the choice of interleaving can significantly affect precision and that in general, the matching problem is cardinal to equivalence checking.

We present a novel speculative search algorithm that aims to find an interleaving of the two programs with minimal abstract semantic difference and evaluate our approach over real-world examples including patches from Git, GNU Coreutils, Linux kernel and the Mozilla Firefox web browser.

Biodata:

Nimrod Partush is a student in the direct PhD course of the Technion CS faculty after graduating summa cum laude from the bachelor program. His research topic is Differential Program Analysis which studies techniques for characterizing program difference and proving equivalence, and applying these towards various uses. His interests include programming languages, static and dynamic program analysis and security.

2) Abstract:

Code similarity is a central challenge in many programming related applications, such as code search, automatic translation, and education. In this talk I'm going to present a new approach tackling this problem. We convert the problem of semantic relatedness between code fragments into a problem of semantic relatedness of textual descriptions.

Our main idea is based on the relationship between code and its textual descriptions as established on the web. Consequently, we can determine semantic relatedness or similarity of code fragments across different libraries and even programming languages, a task considered extremely difficult using traditional approaches.

To experiment with the above approach, we have implemented our approach based on the well known Q&A site, Stackoverflow.

Biodata:

Meital Ben Sinai is currently a graduate student under the supervision of Prof. Eran Yahav, at the Computer Science Department of the Technion in Israel. Her research interests lie in the field of Programming languages, Program Analysis, Natural Languages Processing, Big Data (code) and Security. The topic of her thesis is "Code Similarity via Natural Language Descriptions" which address the question of similarity between code snippets from different programming languages via natural language techniques. She also interested in security and mobile, especially Android.