

Code Search within Large Repositories: from Universal Libraries to CodeGenie and Beyond

Imagine an infinite code repository with every possible implementation of any conceivable function. With such a code base, programming could be reduced to something like finding the right pieces of software and putting them together to form the desired system. Although we will never have limitless software archives, today we can find vast and growing amounts of code in open source communities like GitHub. This scenario presents great opportunities to improve software reuse and other development tasks. In this talk I will discuss some ideas related to code search within large codebases. In the first part of the talk, I delve into the idea of universal libraries. In the second part of the talk, I discuss CodeGenie, a tool that uses test cases to search for code; present results reached in experiments with interface-driven code search and automatic query expansion; and describe outcomes from experiments into software redundancy. Finally, I conclude the talk by discussing future ideas to be explored in this context.



Dr. Otavio Lemos

Associate Professor

Federal University of São Paulo

Otavio has received his MSc (2005) and DSc (2009) degrees in Computer Science from the University of São Paulo, Brazil. Part of his PhD research was carried out at the University of California, Irvine (UCI), USA, under the supervision of Cristina Lopes. In 2009, he started working at the Federal University of São Paulo (ICT-UNIFESP), Brazil, as an Assistant Professor. In 2016 he went back to the USA for a one-year sabbatical at UCI, and in the end of 2017 he was promoted to Associate Professor at ICT-UNIFESP. During his academic career, he has published a number of papers in important conference proceedings (such as ICSE) and journals (such as JSS). His research interests include software reuse - especially code search -, software testing, empirical software engineering, and agile development.

Wednesday, October 17th, 2018

1:30 PM - 3:30 PM

Gates-Hillman Center, Room 6501