Abstract

Studies continue to report that more time is spent reading, locating, and comprehending code than actually writing code. The increasing size and complexity of software systems makes it significantly more challenging for humans to perform maintenance tasks on software without automated and semi-automated tools to support them, especially in the error-prone tasks. Thus, software engineers increasingly rely on software engineering tools to automate maintenance tasks as much as possible.

The program analyses that drive today's software engineering tools have historically focused on analyzing the program's data and control flow, dependencies, and other structural information about the program to uncover and prove program properties. Yet, a software system is more than just the source code and its structure. To build effective software tools, the underlying automated analyses need to use all the information available to make the tools as intelligent and useful as possible. By adapting natural language processing (NLP) to source code analysis, and integrating information retrieval (IR), NLP, and traditional program analyses, we can expect significant improvement in automated and semi-automated software engineering tools for many different software engineering tasks.

In this talk, I will overview our research in text analysis of software and related software artifacts and discuss our achievements to date, the challenges faced in text analysis, and the opportunities for text analysis of software in the future.

Bio

Lori Pollock is a Professor in Computer and Information Sciences at University of Delaware. She earned her Ph.D. and M.S. in CS at University of Pittsburgh in 1986 and 1983, respectively, and her B.S. in CS and Economics at Allegheny College in 1981. Her research focuses on program analysis for building better software maintenance tools, software testing, energy-efficient software, and computer science education. She is an ACM Distinguished Scientist and was awarded the University of Delaware's Excellence in Teaching Award and the University of Delaware's E. A. Trabant Award for Women's Equity. She has actively worked for improving the participation of women and other underrepresented groups in computer science for many years. She serves on the Executive Board of the Computing Research Association's Committee on the Status of Women in Computing (CRA-W) and CRA Committee on Education. She is also leading Partner4CS, a project towards meeting the CS10K goals in the Delaware region.