

"All we like sheep: Cloning as a software engineering tool"

Michael W. Godfrey, PhD  
Software Architecture Group (SWAG)  
David R. Cheriton School of Computer Science  
University of Waterloo

Every programmer understands the value of the copy-paste idiom. Furthermore, studies suggest that 10-20% of the codebase of industrial software systems consists of code that has been duplicated, or ``cloned'', from elsewhere within the system. However, the wide scale use of "code cloning" is usually characterized as evidence of poor design or lazy developer practice. In the last 10 years, research into software clone detection has provided several techniques for locating software code clones; in turn, these techniques have allowed for detailed case studies of many large software systems, which have led to new insights on how cloning may be used as a principled software engineering tool. In this talk, I will further discuss the issue of code cloning; I will give a technical overview of some of the clone detection techniques in use, and I will discuss some of the lessons we have learned from studying the use of cloning in industrial software systems.

Michael W. Godfrey is Associate Professor at the David R. Cheriton School of Computer Science in the University of Waterloo, where he is also a member of SWAG, the Software Architecture Group. Between 2001 and 2006, he held an Associate Industrial Research Chair in telecommunications software engineering sponsored by Nortel Networks and NSERC. He holds a PhD in Computer Science from the University of Toronto (1997) and was previously a faculty member at Cornell University. His main research area is software evolution: understanding how and why software changes over time. His research interests include empirical studies, software tool design, reverse engineering, and program comprehension.