

– Keynote –

Leveraging Machine Learning to Guide Software Evolution

Leon Moonen

Simula Research Laboratory, Norway
<https://leonmoonen.com/> leon.moonen@computer.org

ABSTRACT

Knowledge about dependencies between system artifacts such as modules, methods and variables is essential for a variety of software maintenance and software evolution tasks. Unfortunately, existing approaches to uncover such dependencies by means of static or dynamic program analysis are typically language-specific. Their application is thus largely restricted to homogeneous systems, which is a major drawback given the increasingly *heterogeneity* in modern software systems.

In this talk, we will look at the alternative of using unsupervised machine learning techniques such as *association rule mining*, which can be used to infer knowledge about the relationships between items in a data set. Association rule mining has been successfully used to analyze the change history of a software system and uncover so called *evolutionary coupling* between its artifacts. One of the advantages of this approach is that it is *language-agnostic*, and uncovering dependencies across artifacts written in different programming languages essentially comes for free.

We will explore how association rule mining can be used to derive evidence-based recommendations to guide software maintenance and evolution tasks. Examples include software change impact analysis, recommending related change during development, and conducting targeted regression testing. We survey the state-of-the-art, analyze why and where the applicability of existing techniques falls short, and discuss several avenues for improvement, including novel mining algorithms, methods for aggregating the evidence captured by individual rules, and guidelines for selecting appropriate values for parameters of the mining algorithms.

BIOGRAPHY

Leon Moonen is a senior research scientist in the Software Engineering department at Simula Research Laboratory, Norway. His research aims at creating better techniques and tools to support the understanding, assessment and evolution of large industrial software systems. This work involves the combination of several fields, such as software analytics, program comprehension, software reverse engineering, software

repository mining, machine learning and empirical software engineering. He likes to work in close collaboration with industry to ensure that his research addresses questions of practical value, and to evaluate candidate solutions in real-life circumstances. Current ongoing projects include recommendation systems to support smarter evolution and testing of safety-critical cyber-physical systems, software analytics for continuous software quality and maintainability assessments, methods for creating anti-fragile software systems, and high integrity software engineering.

Dr. Moonen has published over 100 scientific papers and serves on steering-, organizing-, and program committees of international conferences on software maintenance and evolution, reverse engineering, program understanding, and source code analysis. He has (co-)organized various workshops on topics related to these areas. He is co-founder of the Software Improvement Group, a company that specializes in the use of source code analysis to help organizations get control over their software systems. Dr. Moonen received his MSc (cum laude, Computer Science, 1996) and PhD (Computer Science, 2002) from the University of Amsterdam. He is a member of ACM, IEEE Computer Society, EAPLS and the ERCIM Working Group on Software Evolution.



Photo by Bård Gudim