

Regression Tests and the Inventor's Dilemma

Leonardo de Moura

Microsoft Research, Redmond

Reasoning systems are extensively used at Microsoft. They are used in test case generation, model-based testing, static program analysis, program verification, analysis of firewall policies, program synthesis, geometric problem solving, to cite a few. Reasoning systems are complicated pieces of software, which are very often trying to attack undecidable problems. In this talk, we describe different systems used at Microsoft, how they are evaluated, and challenges upon releasing new versions of successful systems. All systems share the same basic evaluation technique: regression tests. They also face the same challenge: the inventor's dilemma. In all these systems, progress is not monotonic. The latest version of the Z3 theorem prover may enable users to automatically prove many new theorems, provide new features and performance improvements, but it inevitably also fails to prove theorems proved by the previous version. More importantly, we very often have to give up some of our progress to be able to reach the next plateau. We conclude describing how we have been addressing this challenge in the Z3 project.