

Integrating Large Language Models in Automated Program Verification

Nina Narodytska 

VMware Research by Broadcom

USA

n.narodytska@gmail.com

Abstract—The demonstrated code-understanding capabilities of large language models (LLMs) raise the question of whether they can be used for automated program verification—a task that typically demands high-level, abstract reasoning about program properties, which remains challenging for existing verification tools. In this talk, we overview several research directions, developed over the past two years, that propose methodologies for combining the strengths of LLMs and automated reasoners in program verification. We then take a closer look at Lemur, a tool that formally describes such a methodology as a set of derivation rules and proves its soundness. Lemur instantiates the calculus as a sound automated verification procedure and demonstrates practical improvements on a suite of synthetic and competition benchmarks.