NUS co-leads S$10 million international programme to advance AI-agent based software reliability and verification

The National University of Singapore (NUS) is co-leading a new S$10 million programme to revolutionise software reliability and verification using AI, with the aim of addressing one of the most critical challenges in computer science: ensuring the correctness and security of software systems in an era increasingly dominated by AI-generated code.

Known as *AI for Program Reasoning*, the new programme is co-led by Professor Abhik Roychoudhury from the Department of Computer Science in NUS School of Computing, and Professor Cristian Cadar from the Imperial College London. They will work closely with leading experts from the Singapore Management University, Massachusetts Institute of Technology, and ETH Zürich.

Software powers modern society, from scientific discovery to critical infrastructure, and errors can lead to severe consequences, such as research retractions and security breaches. As AI systems generate more code, the need for strong and reliable verification becomes even more urgent.

The four-year *AI for Program Reasoning* programme will develop advanced AI technologies for program reasoning, combining symbolic analysis with generative AI to automate both formal and informal reasoning about software.

Revolutionising AI-driven program reasoning

The programme will focus on three areas:

- Making code checking easier across programming languages and domains through AI-generated analysis tools.
- Advancing code-focused large language models (LLMs) so that their reasoning can be easily verified.
- Automating the generation of specifications and proofs for large-scale software verification, including critical systems such as the Linux operating system. The infrastructure will also be applied for AI-based verification of AI-generated code.

Prof Roychoudhury remarked, “Applying formal, logic-based methods to prove the correctness of computer systems has been a long-standing ambition. A key difficulty in achieving this goal has been the lack of specifications to prove correctness. With the rise of AI agents that can combine the reasoning about text as well as code representations - there remains a real possibility of extracting non-trivial specifications which are then proved automatically via proof search automation. The impact of such research would be huge, leading to the validation of widely used software like the Linux kernel.”

“Modern infrastructure and technology rely on software, and reliability failures can lead to serious consequences that impact the economy and society. As AI begins to write significant portions of code, we must ensure that this code is not only functional but also safe and secure. Our research will create AI systems that can reason about programs, verify their correctness, and even repair them when necessary,” added Prof Cadar.

The team will deliver demonstrators such as customised reasoning engines, benchmarks for code verification, and AI agents capable of repairing both code and proofs. These innovations will support Singapore’s rapidly growing digital economy and strengthen its position as a global leader in software science.