Title: An Empirical Study on Static Analyzer Toolsets to Reduce False Positives, False Negatives in Python Type Checkers

Benedek Kaibas, *Allegheny College, PA, 16335, USA*

Dr. Gregory M. Kapfhammer, *Allegheny College, PA, 16335, USA*

Abstract—Python's massive scale relies on static type checkers like Mypy to prevent crashes. Yet, verifying these tools remains a slow, manual bottleneck. Who checks the checkers? We introduce Pytifex, a system where AI agents automate this oversight. Unlike random 'fuzzing' tools, our agents analyze historical bug reports to learn specific structural weaknesses. They then generate new, deceptive code examples that trick the latest tool versions into reporting false errors. This exposes hidden flaws standard tests miss. We argue that this AI-driven adversarial testing is critical to ensuring the safety nets of modern software remain reliable.