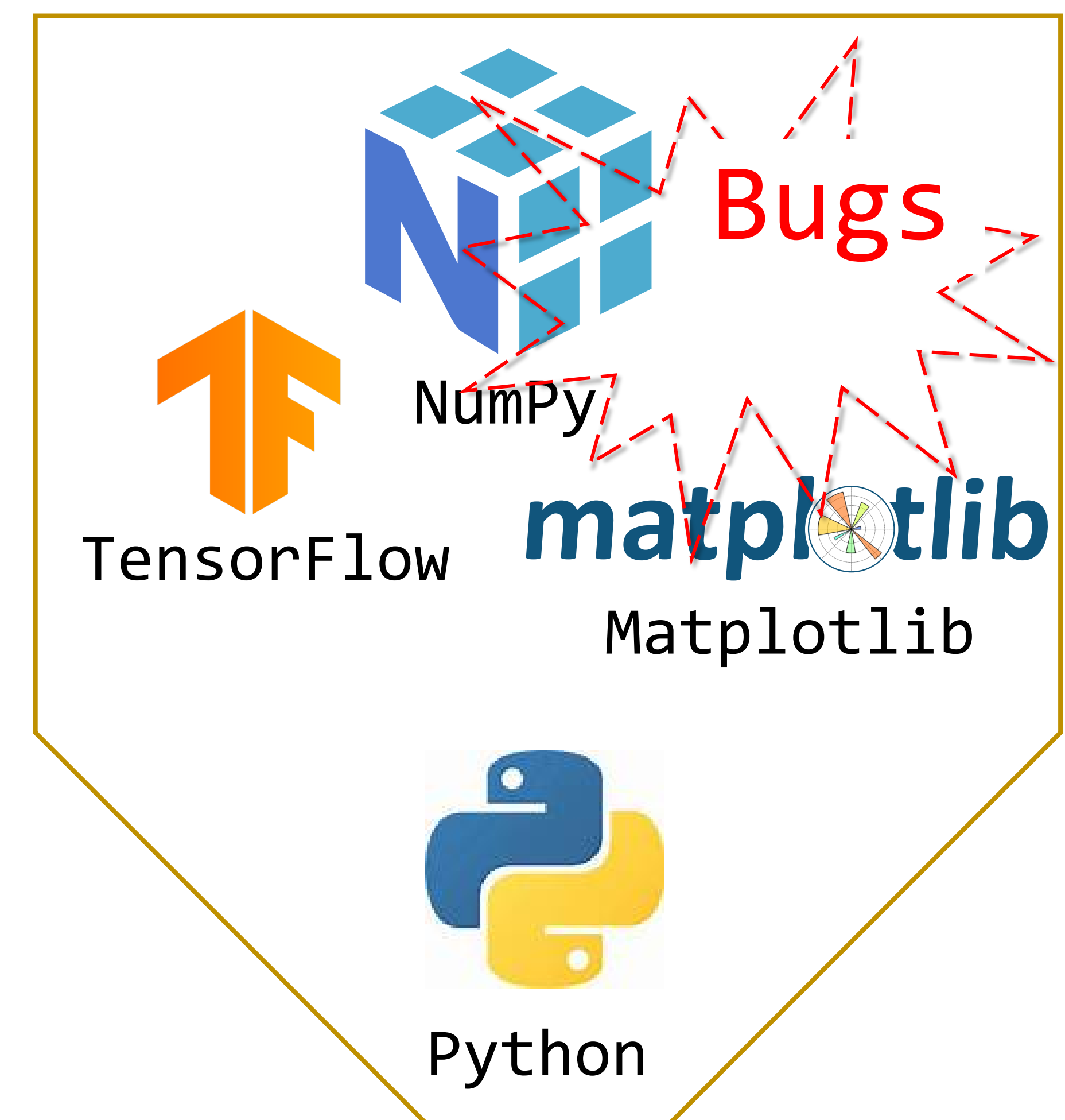


A Python Benchmark for Bugs with Python Built-in Types

Xin Zhang, Rongjie Yan, Jiwei Yan, Baoquan Cui, Jun Yan, Jian Zhang
{zhangxin19, yrj, yanjw, cuibq, yanjun, zj}@ios.ac.cn

Introduction

- The need for bug detection tools is growing in tandem with the adoption of the Python programming language.
- It is critical to understand how effective existing Python tools are in detecting crash bugs in practice.
- Some mature benchmarks are widely adopted in practice, but no benchmark in Python.



Bug Collection



1 Project Selection

300 projects

2 Commits Collection

30 projects
2966 commits

3 Bug Reproduction

17 projects
183 bugs

4 Patch Validation

Bug-irrelevant code pruning

15 projects
180 bugs

List of executed
Python files

Code Pruner

Path Tracer

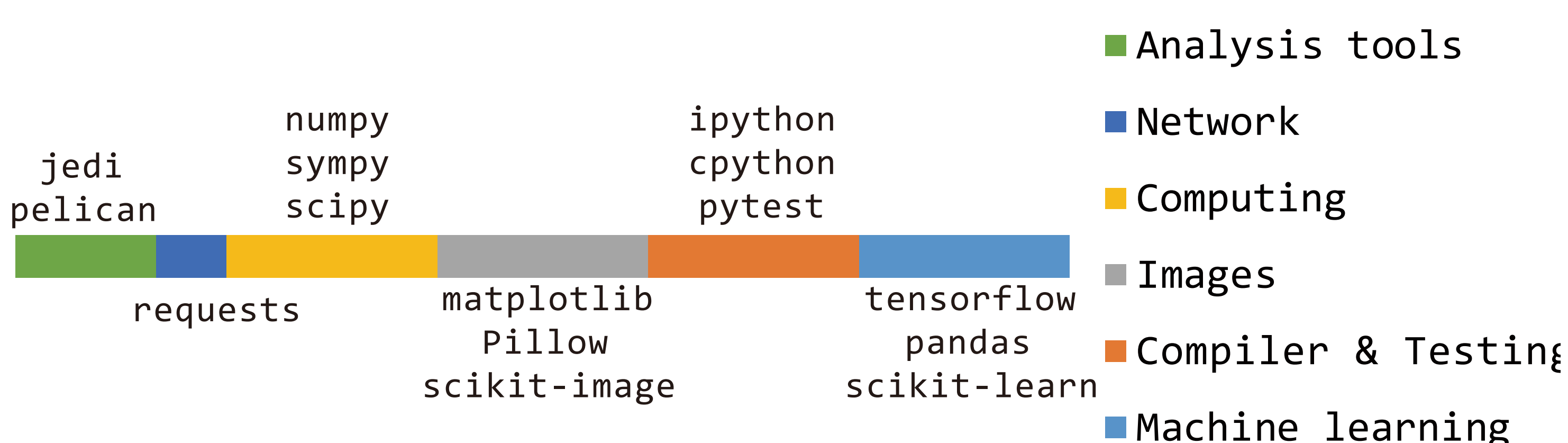
Re-run the test
code

Delete one
unexecuted file

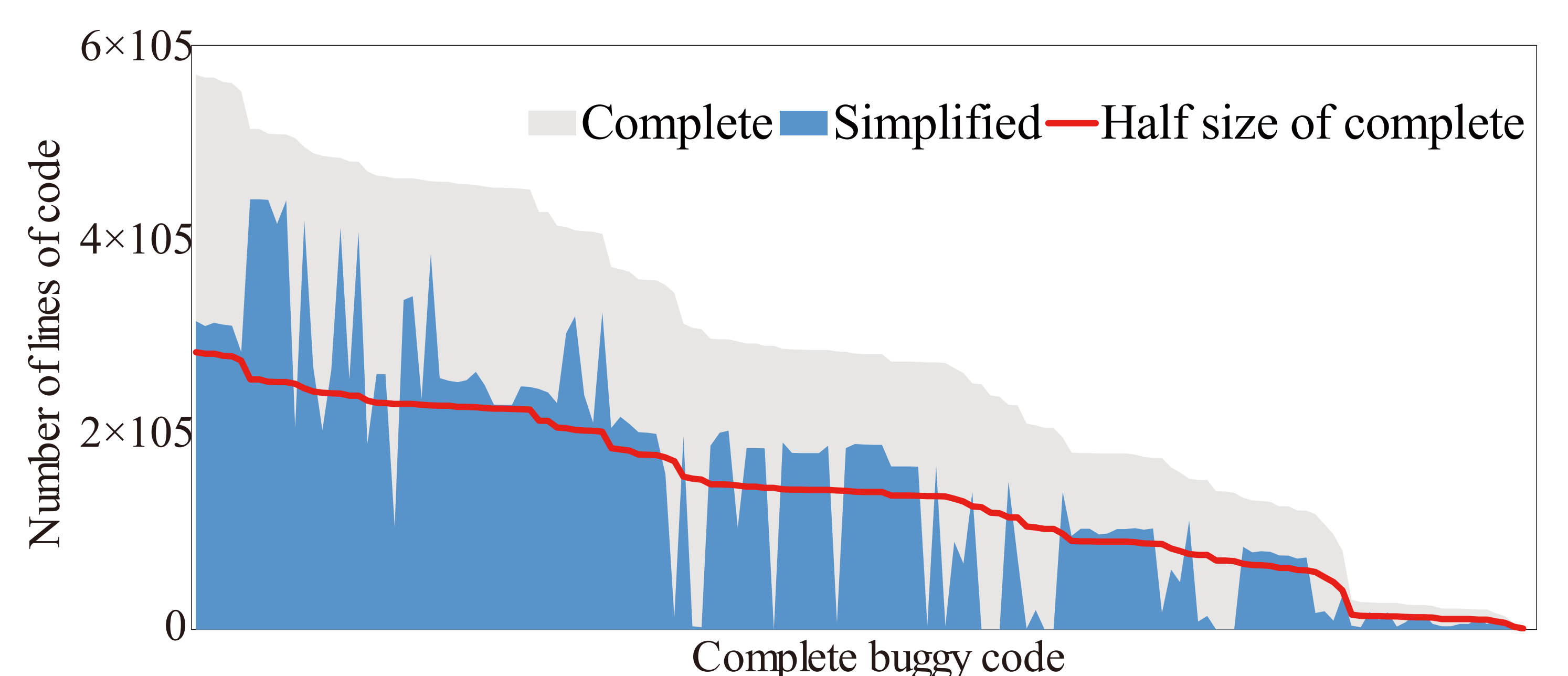
Re-run the test
code

Simplified buggy code

Evaluation



Project Distribution



Scale of reduced code