自动化测试大作业

南京大学 软件学院 iSE实验室

目录

01. 作业概述

02. 文献综述

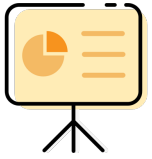
03. 工具实现

04. 联系方式

01

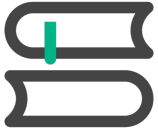
作业要求

软工研究四维



Presenting

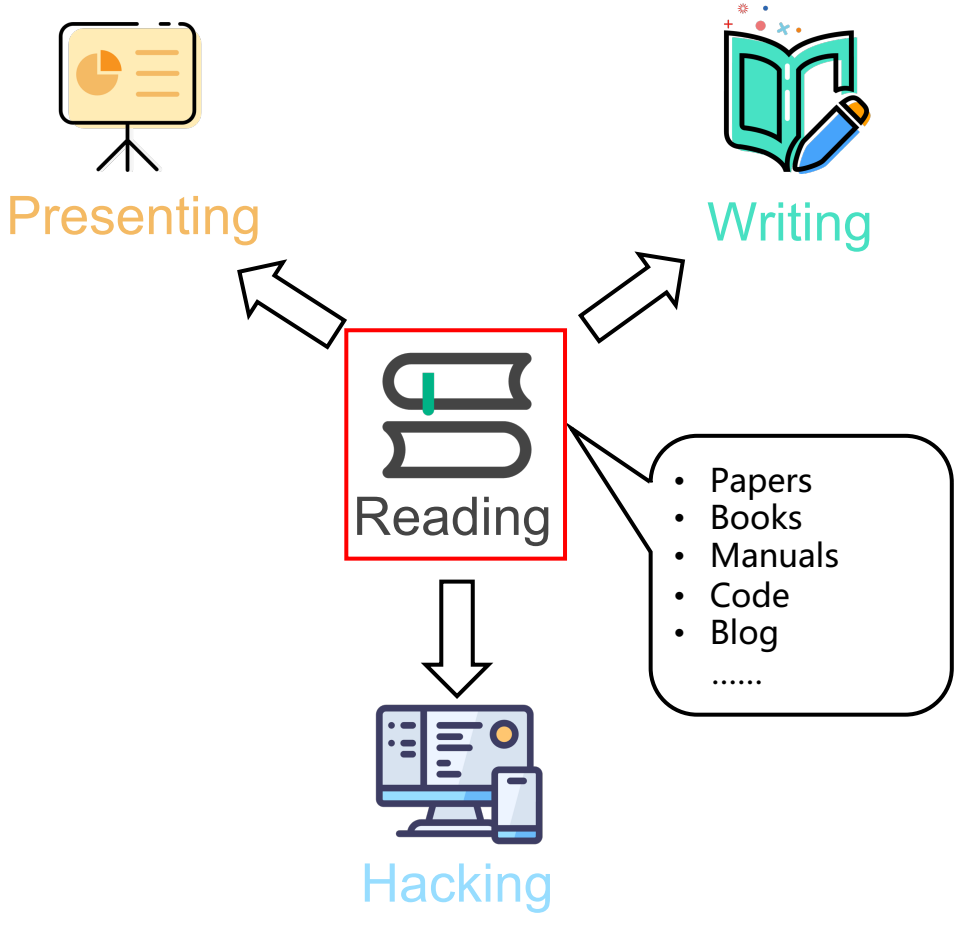
Writing



Reading

Hacking

软工研究四维


Presenting
Writing
Reading
Papers
Books
Manuals
Code
Blog
.....
Hacking

作业概述

- 小组作业：每组不超过4人
- 作业组成：文献综述 + 工具/流程实现
- 作业方向：经典、AI、移动应用
- 分数组成：20%综述 + 30%工具 + 5%课堂汇报
- 选题方式：问卷收集
- 提交方式：提交GitLink[1]仓库链接，先私有后公开

[1] https://www.gitlink.org.cn/

作业概述

- 时间节点（DDLs，当日的23:59）
  - 选题（10月18日）
  - 综述汇报（10月25日 & 10月28日）
  - 综述提交（11月18日）
  - 代码提交（12月2日）

02

文献综述

总体要求

- 题目选择：每组一个
- 文献来源：CCF-**A,B**,C类会议与期刊[1,2]
- 文献年限：时间近5~10年，越新越好
- 文献数量：每组20~40篇**初级研究**（Primary Study）
- 模板要求：计算机学报模板
- 页数要求：6页正文+2页参考文献
- 编写工具：Word、Overleaf[3]、NJU Latex[4]

[1] 中国计算机学会推荐国际学术会议和期刊目录（第五版），2019年修订， https://ccf.atom.im/
[2] CCF推荐期刊会议下载链接： https://www.ccf.org.cn/ccf/contentcore/resource/download?ID=99185
[3] https://www.overleaf.com/
[4] https://tex.nju.edu.cn/

- Mutation Testing
  - 推荐期刊&会议
    - 软件工程领域的顶级期刊和会议
    - 期刊：TSE, TOSEM, IST, JSS, ESE
    - 会议：ISSTA, FSE/ESEC, ASE, ICSE, ISSRE, ICST

- Mutation Testing
  - 变异测试优化技术综述
    - 变异体的选择、约简、执行、分析
  - 变异测试应用综述
    - 优化回归测试、引导测试生成、AI测试

- Fuzzing
  - 推荐期刊&会议
    - 软件工程&安全领域的顶级期刊和会议
    - 期刊：TSE, TOSEM, TDSC
    - 会议：ISSTA, FSE/ESEC, ASE, ICSE, PLDI, Usenix Security, CCS, S&P, NDSS

- Fuzzing
  - 模糊测试技术中种子调度技术综述
    - 种子选择、种子排序、能量分配
  - 定向模糊测试技术综述
    - 白盒、灰盒、调度方式、程序分析技术
  - 基于生成的模糊测试技术综述
  - 内核模糊测试技术综述

- 有效警告识别（Actionable Warning Identification, AWI）技术综述
  - 文献检索：要求2020年以后的文献不少于10篇，总文献数量不得少于20篇
  - 推荐文献&期刊：软件工程领域顶会顶刊

- 有效警告识别（Actionable Warning Identification, AWI）技术综述

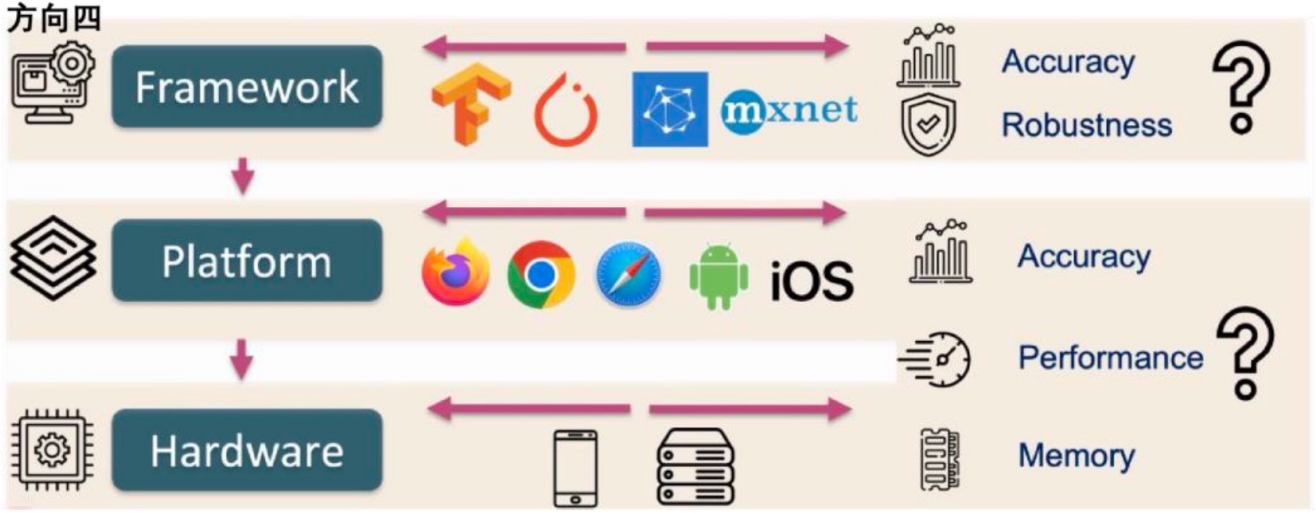
序号	目的	关键字
1	后处理警告	1) elimination, 2) reduction, 3) simplification, 4) ranking, 5) classification, 6) reviewing, 7) inspection
2	静态分析	1) static analysis, 2) automated code analysis,3) source code analysis, 4) automated defects detection
3	警告	1) alarm, 2) warning, 3) alert, 4) violations

- 基于机器学习的测试用例排序综述
  - 测试排序：对回归测试用例进行排序，目的是今早地发现缺陷、减少回归测试开销
  - 目标文献：测试用例排序与机器学习技术的交叉领域
    - AI for TCP：使用机器学习技术完成传统回归测试任务
    - TCP for AI：测试排序在机器学习任务中的应用
  - 综述内容：排序技术、机器学习技术、排序准则、评估标准、被测集以及应用场景

- 研究框架
  - 从智能模型模糊测试的角度着手，按照不同的阶段，分别关注**测试数据的选择、生成**以及**测试结果的评估**

- 综述题目
  - 方向一：深度学习模型测试数据选择技术研究
    - 关键词：测试数据选择、测试数据排序、测试数据度量
  - 方向二：深度学习模型测试数据生成技术研究
    - 关键词：数据变异、蜕变关系、数据扩增
  - 方向三：深度学习模型测试结果评估技术研究
    - 关键词：鲁棒性、公平性、后门检测、可解释性分析

- 综述题目
  - 方向四：深度学习框架缺陷检测和原因分析技术研究
    - 关键词：深度学习框架/库测试、深度学习框架/库实证研究



- 综述题目
  - 基于机械臂的非侵入式测试技术
  - 移动应用自动化测试结果复现技术
  - 基于GUI界面分析的移动应用测试脚本优化技术
    - 录制回放、脚本修复等技术
  - 面向移动应用GUI界面控件的识别、分析、理解技术
  - 移动应用GUI自动化探索测试技术

综述选题汇总

- Classic Testing
  - 变异测试优化技术综述
  - 变异测试应用综述
  - 模糊测试技术中种子调度技术综述
  - 定向模糊测试技术综述
  - 内核模糊测试技术综述
  - 有效警告识别技术综述
  - 基于机器学习的测试用例排序综述

综述选题汇总

- AI Testing
  - 深度学习模型测试数据选择技术研究
  - 深度学习模型测试数据生成技术研究
  - 深度学习模型测试结果评估技术研究

综述选题汇总

- GUI Testing
  - 基于机械臂的非侵入式测试技术
  - 移动应用自动化测试结果复现技术
  - 基于GUI界面分析的移动应用测试脚本优化技术
  - 面向移动应用GUI界面控件的识别、分析、理解技术
  - 移动应用GUI自动化探索测试技术

03

工具实现

- Fuzzing & Mutation testing工具实现
  - 提交内容
    - 代码、脚本（Process & Analysis）
    - 过程报告：一些配置和流程上的坑、配置的汇总、设计框架&流程等 – PDF
      - 选题描述、项目结构、环境/实验设置、Fuzzing配置、构建流程（引导）、框架设计、结果分析
    - 运行结果（图、表）-- 一系列PDF，CSV，xlsx等
    - 压缩成Zip

- Fuzzing & Mutation testing工具实现
  - Fuzzers：AFL（Base）
    - A组：AFL, AFL++, AFLGo, AFLFast, AFLSmart, Mopt, FairFuzz, EcoFuzz, Neuzz, MTFuzz
    - B组：Vuzzer, Angora, LibFuzzer, Entropic

- Fuzzing & Mutation testing工具实现
  - Subjects
    - Real-world Projects[1]：官网、Git仓库，尽量下载最新版本
    - DARPA CGC dataset[2]
    - LAVA-M[3]
  - C/C++项目的构建
    - gcc, clang, make, cmake, autoconf...

[1] https://docs.qq.com/sheet/DZGtod3FBZ2lXZHhS?tab=BB08J2
[2] https://github.com/CyberGrandChallenge/
[3] Dolan-Gavitt B, Hulin P, Kirda E, et al. Lava: Large-scale automated vulnerability addition[C]//2016 IEEE Symposium on Security and Privacy (SP). IEEE, 2016: 110-121.

- Fuzz-Mut：基于变异测试的模糊器评估
  - 要求
    - 工具：AFL +Mull
    - Subjects：Real world projects链接中的所有项目
    - Campaign设置：每个至少一小时
    - 变异杀死条件：Crash & **Differential Comparing**

- Fuzz-Mut：基于变异测试的模糊器评估
  - 流程：AFL生成测试输入 → Mull复现输入 → 统计分析
    - 生成输入：构建项目（Fuzz）、确定参数、运行afl-fuzz
    - 复现输入：<afl-out>/queue下保存的Test Input、构建项目（Mull）、编写变异测试脚本、运行mull-runner
    - 统计分析：汇总每个在每个Input时间点的变异杀死率，绘制曲线time-mut_score曲线
  - 提交内容：流程&分析脚本、过程报告、分析结果

- Fuzz-Mut：基于变异测试的模糊器评估
  - 参考
    - AFL仓库： https://github.com/google/AFL
    - AFL博客/文档： https://afl-1.readthedocs.io/en/latest/
    - Mull论文：Mull It Over: Mutation Testing Based on LLVM
    - Mull文档：https://mull.readthedocs.io/en/0.19.0/

- Fuzz-Cov：基于覆盖率的模糊器评估
  - 要求
    - Fuzzers：AFL+其他 A Fuzzer*1 + B Fuzzer * 1
    - Subjects：Real world projects链接中的所有项目
    - Campaign设置：每个至少一小时

- Fuzz-Cov：基于覆盖率的模糊器评估
  - 流程：Fuzzer生成输入 → gcov复现 → 统计分析
    - 生成输入：构建项目（Fuzz）、确定参数、运行各种Fuzzer
    - 复现输入：各种Fuzzer产生的Test Input、构建项目（gcov）、重新运行输入、收集分支覆盖信息
    - 统计分析：汇总每个在每个Input时间点的变异杀死率，绘制曲线time-mut_score曲线

- Fuzz-Cov：基于覆盖率的模糊器评估
  - 参考
    - Gcov官网：https://gcc.gnu.org/onlinedocs/gcc/Gcov.html
    - Google Fuzzing Tutorials : https://github.com/google/fuzzing
    - 各种Fuzzer的论文：Neuzz, MTFuzz, Angora

- AWI：基于置信学习的警告数据集去噪技术
  - 警报的有效性：多版本验证
    - 有效警告：在后续版本中警告中消失（正报警告）
    - 无效警告：在后续所有版本警告一直存在（误报警告）

- AWI：基于置信学习的警告数据集去噪技术
  - 任务构成
    - 初始警告数据集收集：Apache开源Java项目
    - 置信学习技术：Confident Learning: Estimating Uncertainty in Dataset Labels
    - 项目交付结果
      - 初始警告数据集收集中的初始标记
      - 置信学习技术的处理结果

- AutoFix：基于迁移学习的人工智能框架缺陷修复技术
  - 背景
    - 人工智能技术应用广泛
    - 人工智能框架不断涌现
    - 人工智能框架中存在的缺陷影响深远
  - 项目目标
    - 通过迁移学习技术学习通用缺陷语料中的修复模型，对框架数据集进行微调，使用提示学习技术使得修复模型能够识别不同语言之间的修复模式

- AutoFix：基于迁移学习的人工智能框架缺陷修复技术
  - 项目实现
    - 通用缺陷数据集和框架缺陷数据集的提取
    - 通用修复模型的设计与生成
    - 框架修复模型的微调
    - 修复效果的评估

- ART：面向数值程序的自适应随机测试复现技术
  - 背景
    - 自适应随机测试（ART，Adaptive Random Testing）
    - ART vs. RT：优化随机测试生成产生的测试用例分布
  - 项目目标
    - 针对数值程序，复现经典的自适应随机测试算法，比较不同算法的有效性

- ART：面向数值程序的自适应随机测试复现技术
  - 项目实现
    - 选择待复现的经典的自适应随机测试算法
      - 15种基于数值程序的自适应随机算法
    - 自适应随机算法框架搭建
    - 自适应随机测试算法有效性比较、评估

- AI-1：面向xxx场景的深度学习模型测试技术
  - 背景
    - 问题场景的领域特性影响智能模型测试方法的设计
    - 场景：图像、点云、人脸识别、语音识别
  - 任务构成
    - 提供待测模型和场景数据集，分析特定场景、简历测试需求、设计并实现可用测试技术

- AI-2：基于等价融合算子的深度学习框架差分测试技术
  - 背景
    - 深度学习框架提供了训练模型所需的算子
    - 差分测试是针对深度学习框架算子常用测试方法
  - 任务构成
    - 选择被测的深度学习框架，梳理融合算子列表，实现常见融合算子，生成测试数据，开展深度学习框架差分测试

- GUI-1：基于对话系统引导的移动应用自动化测试结果复现技术
  - 背景
    - 人工缺陷复现是一项复杂且困难的工作
    - 智能化引导可以大大提升人工缺陷的效率
  - 项目目标
    - 构建一种基于对话系统引导的自动化测试结果复现技术，以对自动化测试结果的分析和相关测试知识库为指导，构筑对话系统引导测试人员对自动化测试的结果进行复现

- GUI-1：基于对话系统引导的移动应用自动化测试结果复现技术
  - 任务构成
    - 对话系统的设计与实现
    - 自动化测试执行
    - 自动化测试结果分析
    - 复现引导
    - 交互界面的设计与实现

- GUI-2：基于机械臂视觉识别的非侵入式GUI 界面自动化探索（不超过4 组）
  - 背景
    - 传统GUI自动化探索测试难以模拟真正操作
    - 机械臂可以从用户的视角完成GUI测试
  - 项目目标
    - 利用机械臂模拟真正用户对app 的使用，实现基于机械臂视觉识别的非侵入式GUI界面自动化探索

- GUI-2：基于机械臂视觉识别的非侵入式GUI 界面自动化探索（不超过4 组）
  - 任务构成
    - 目标识别
    - 机械臂操作
    - GUI界面探索策略
    - 机械臂调度
    - 交互界面的设计与实现

工具实现汇总

- Classic Testing
  - 基于变异测试的模糊器评估
  - 基于覆盖率的模糊器评估
  - 基于置信学习的警告数据集去噪技术
  - 基于迁移学习的人工智能框架缺陷修复技术
  - 面向数值程序的自适应随机测试复现技术

工具实现汇总

- AI Testing
  - 面向xxx场景的深度学习模型测试技术
  - 基于等价融合算子的深度学习框架差分测试技术

工具实现汇总

- GUI Testing
  - 基于对话系统引导的移动应用自动化测试结果复现技术
  - 基于机械臂视觉识别的非侵入式GUI 界面自动化探索（不超过4 组）

04

联系方式

联系方式

- Classic Testing
  - 钱瑞祥，qrx_at@163.com
  - 葛修婷，1683245057@qq.com
  - 张犬俊，quanjun.zhang@smail.nju.edu.cn
- AI Testing
  - 刘佳玮， jw.liu@smail.nju.edu.cn
  - 刘关迪，liuguandi@smail.nju.edu.cn
- GUI Testing
  - 虞圣呈，yusc@smail.nju.edu.cn
  - 杜铭哲，nandodu@smail.nju.edu.cn

zychen@nju.edu.cn
fangchunrong@nju.edu.cn

Thank you!