



智能软件创新赋能新质生产力发展
2024 CCF ChinaSoft
中国软件大会

智能化芯片设计程序测试

李晓鹏¹ 闫明¹ 樊兴宇¹ 唐振韬² 开昱雄² 郝建业^{1,2} 袁明轩² 陈俊洁^{1*}

1



天津大学
Tianjin University

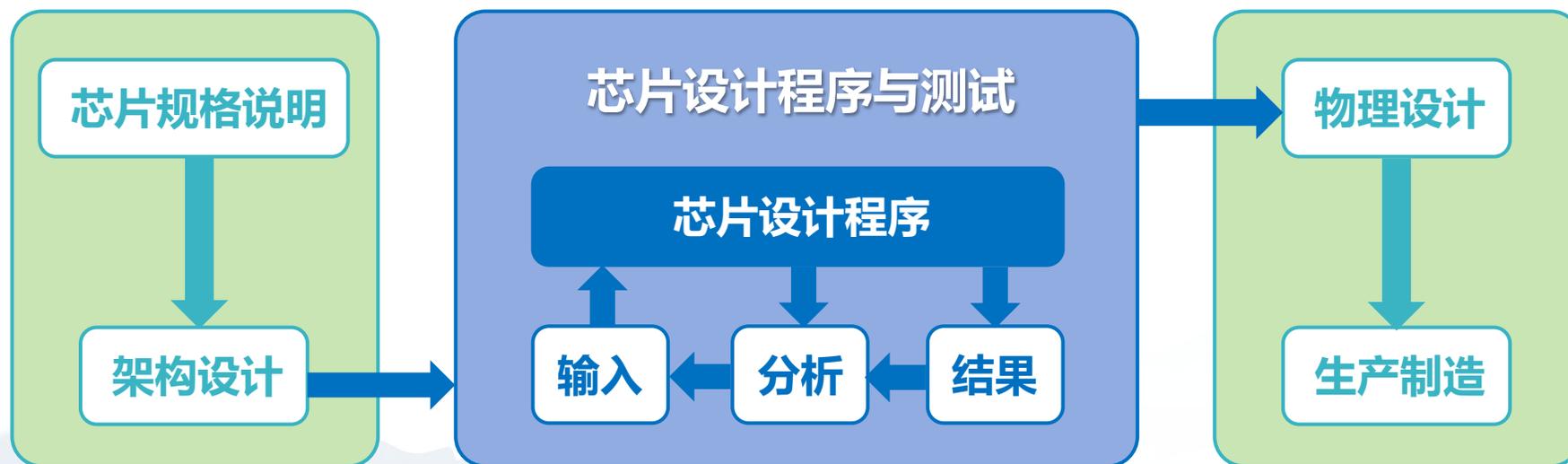
2



李晓鹏

2024年11月16日

■ 芯片设计流程



■ 保障芯片设计程序质量至关重要

■ 测试技术发展

- 早期阶段：人工设计测试用例
- 发展阶段：基于仿真的自动化测试方法
- 智能时代：机器学习、深度学习、大语言模型

■ 测试维度

测试输入生成

覆盖结果

输入特征

语义结构

测试预言

差分测试

蜕变测试

测试优化

测试排序

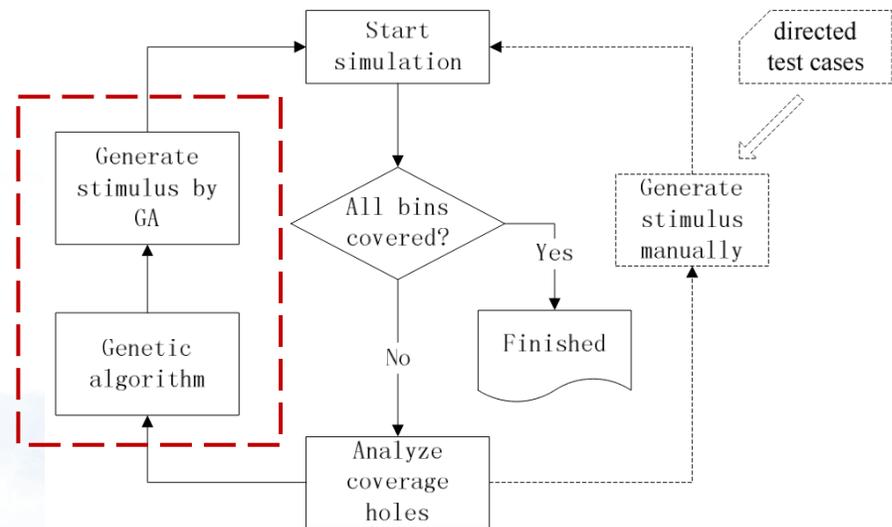
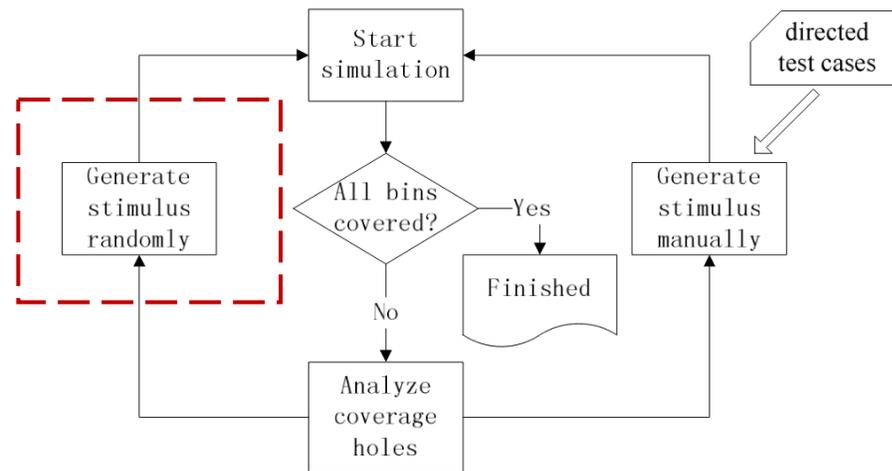
测试选择

测试覆盖指标

- 代码覆盖：程序代码被执行程度
- 功能覆盖：芯片设计功能检测充分性
- 断言覆盖：程序代码执行正确性

基于遗传算法的测试输入生成

- 结合遗传算法和约束随机测试方法
- 编码测试输入、覆盖率作为适应度函数
- 交叉、变异测试输入，提高覆盖率

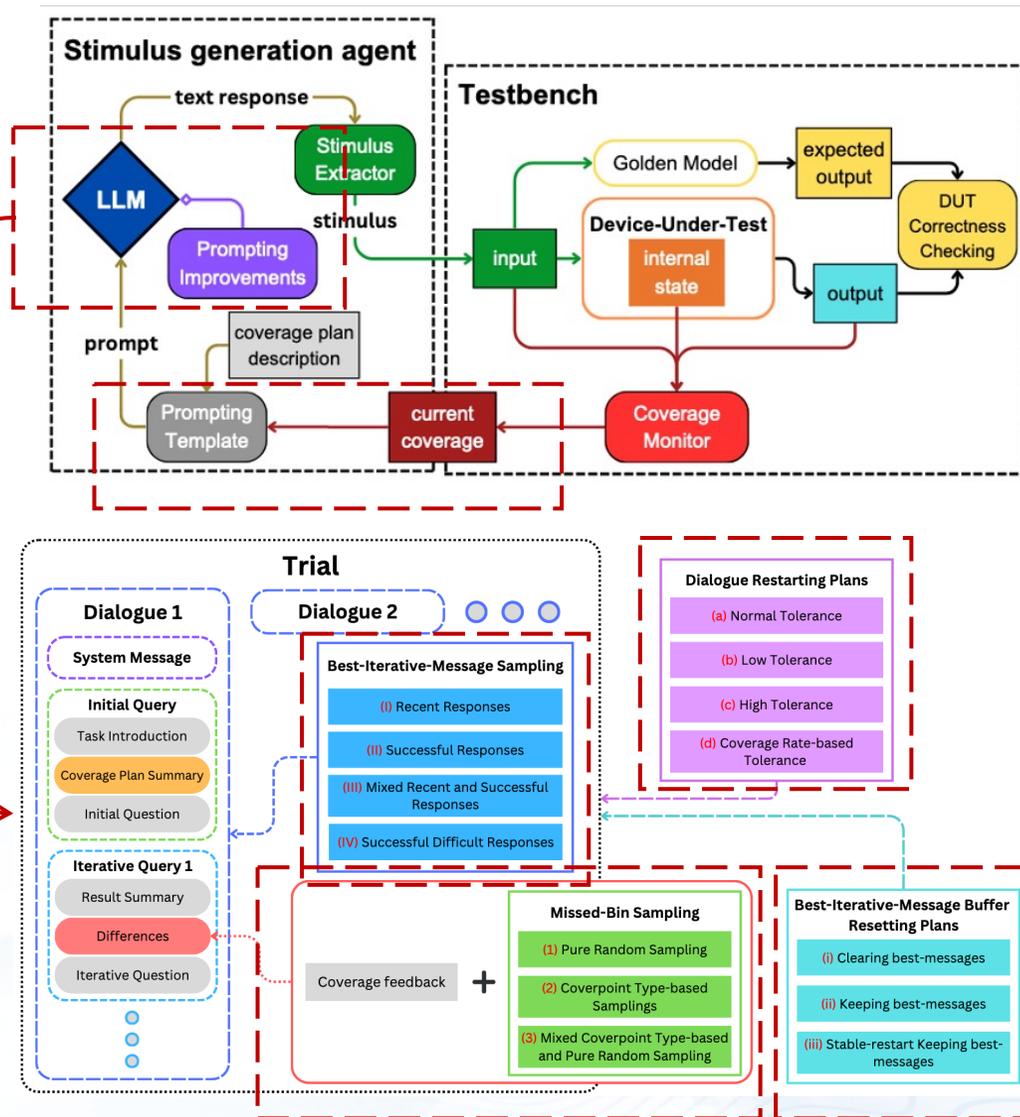


测试输入生成：以覆盖率为引导



基于大语言模型的覆盖率引导测试输入生成

- 提示词模板结合覆盖率信息
- 改进提示词策略
- 考虑大语言模型输入长度限制
- 未覆盖点采样
- 最佳迭代消息采样
- 防止重复历史错误
- 对话重启
- 最佳迭代消息缓冲区重置

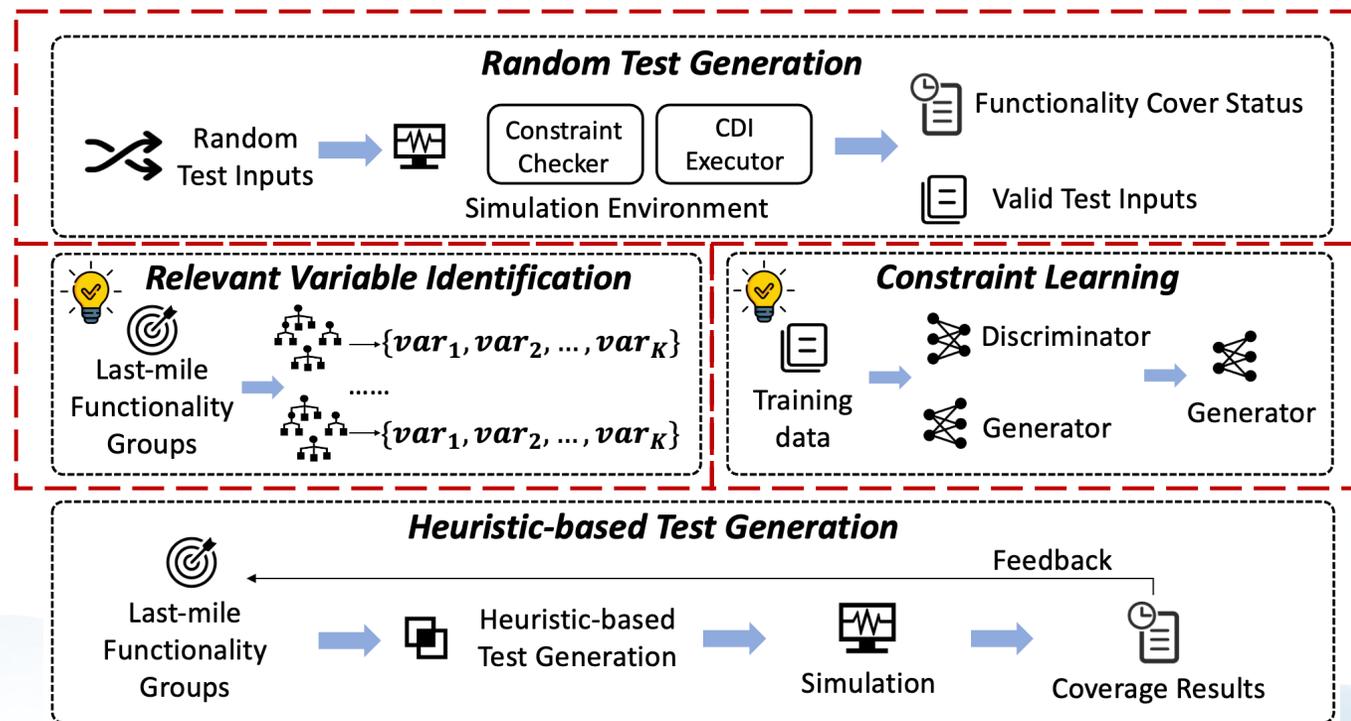


■ 基于随机森林的测试输入特征提取方法

- 黑盒结构，无法分析内部逻辑
- 基于随机测试输入构造训练集
- 学习输入变量和覆盖点的关系
- 压缩搜索空间

■ 基于生成对抗网络的约束特征学习

- 学习不同变量间的复杂约束
- 生成满足约束的测试输入



基于图卷积神经网络表征的芯片设计程序结构

学习芯片设计程序语义抽象

控制数据流图：芯片设计程序语义的连续表示

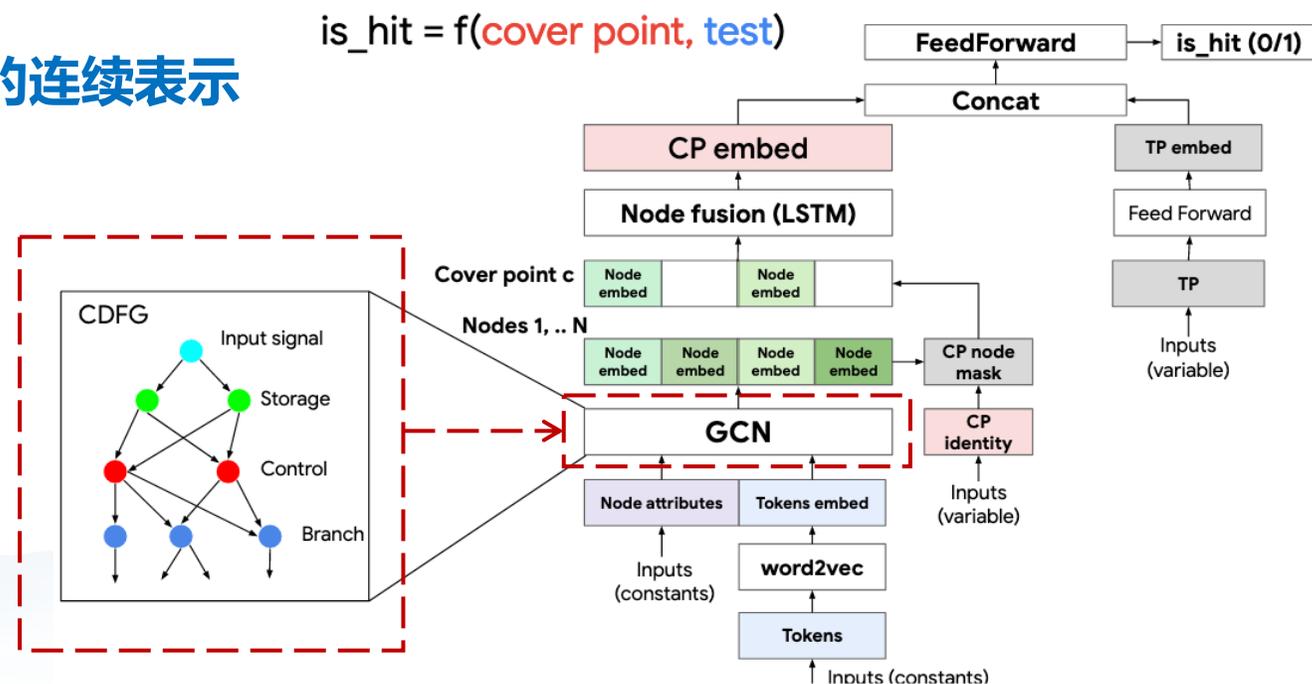
使用图神经网络进行表示学习

预测芯片设计程序的仿真状态

应用：

预测覆盖结果

指导测试输入生成



基于大语言模型的芯片设计程序语义结构理解

芯片设计程序代码阅读器

覆盖率解释器模块

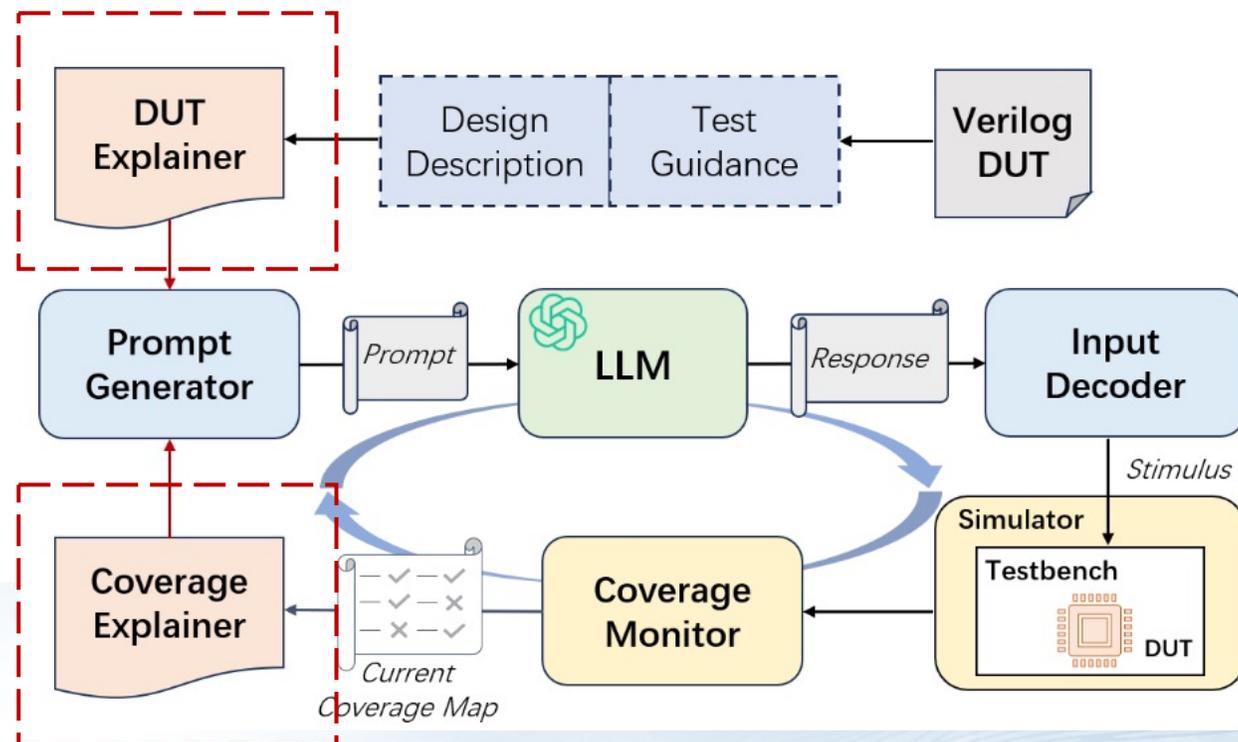
覆盖率报告 -> LLM容易理解的格式

待测程序解释器模块

待测程序 -> 自然语言

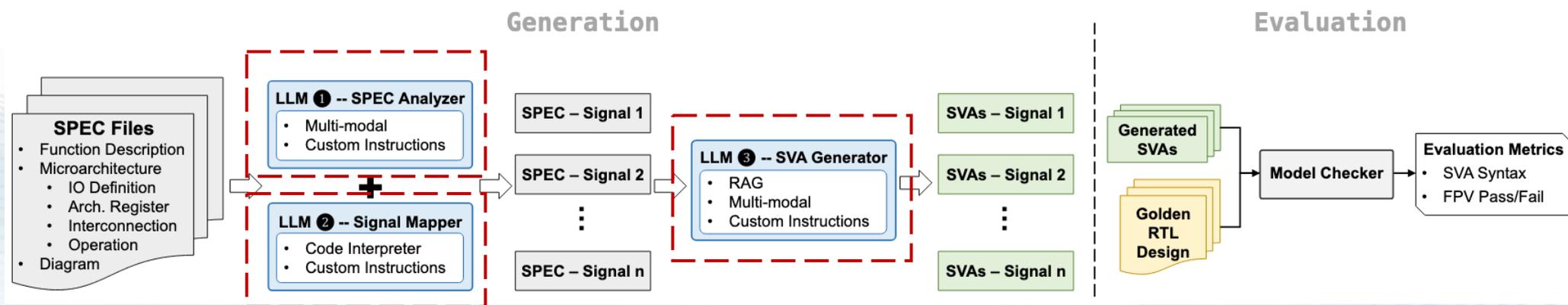
增强LLM对测试意图和程序功能理解

生成到达未探索分支的测试输入



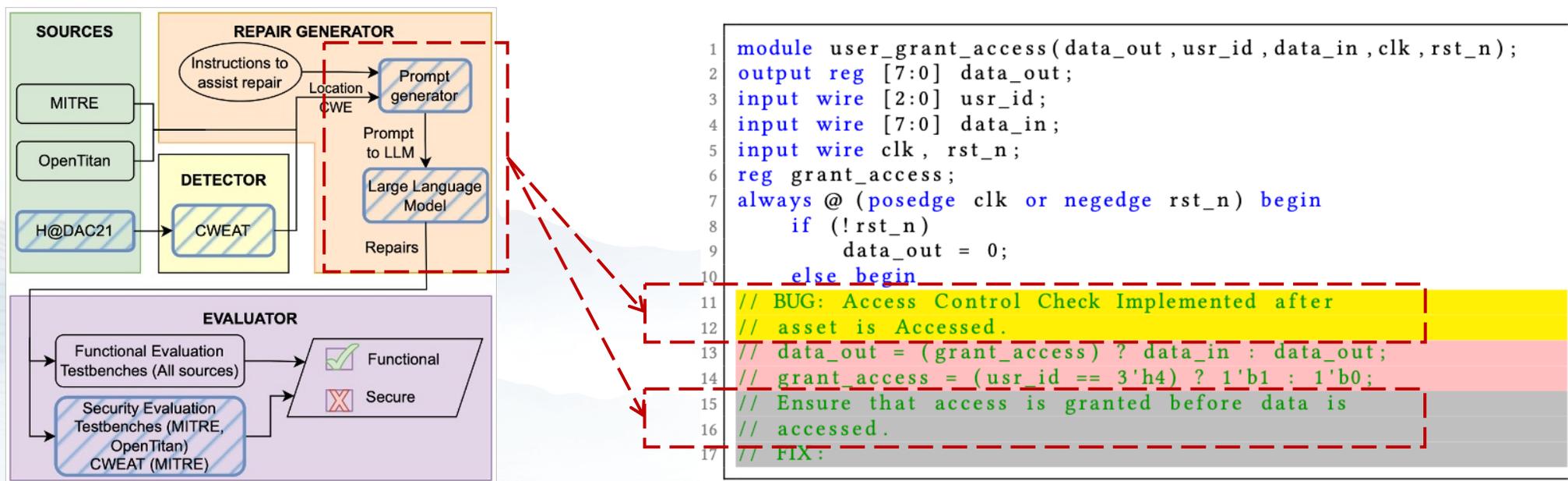
基于大语言模型的断言生成

- 三个定制的大语言模型：提取结构规范、映射信号定义、生成断言
- 提供一个评估断言生成能力的开源基准框架
- 在语法和功能上验证生成断言的准确性

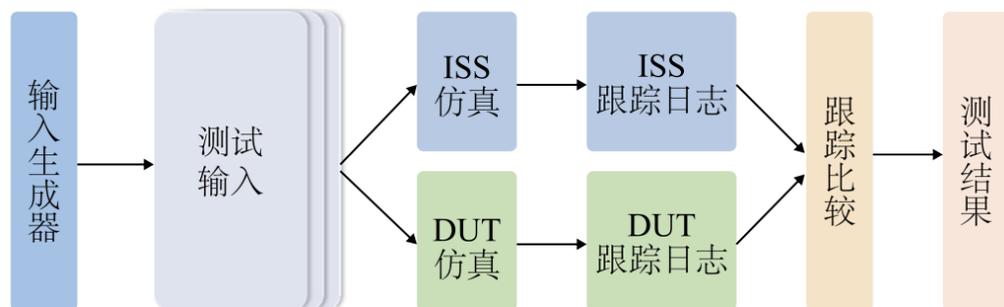


基于大语言模型的芯片设计程序缺陷修复

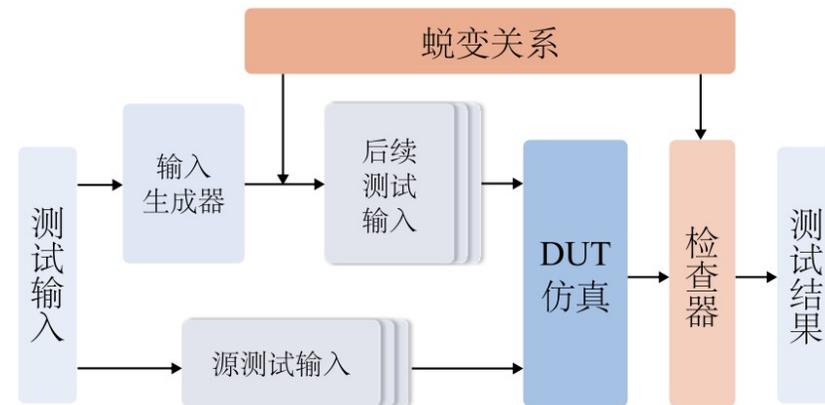
- 收集10种来自OpenTitan等CWE信息
- 设计缺陷指令和修复指令，帮助大语言模型进行修复
- 在功能和安全上评估修复结果



参考模型（差分测试）



蜕变测试



构造功能等价的参考模型

协同仿真：测试输入同步传输

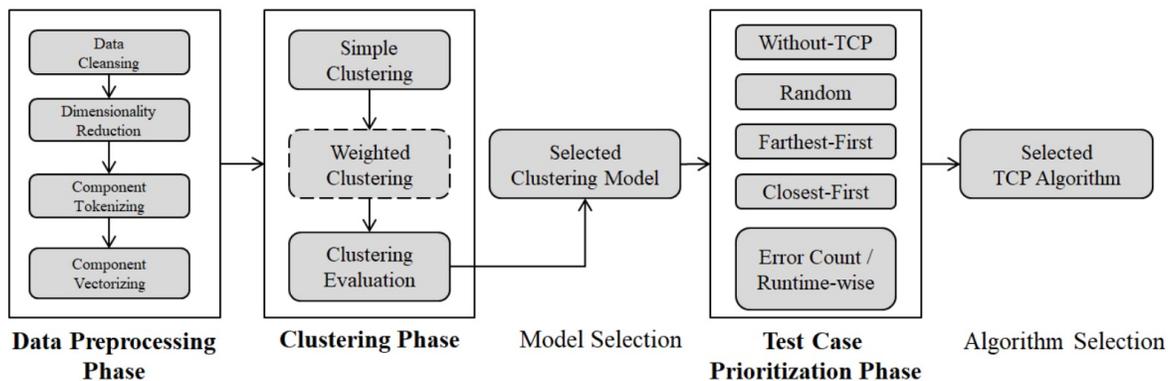
比较状态和行为

参考模型缺失的场景

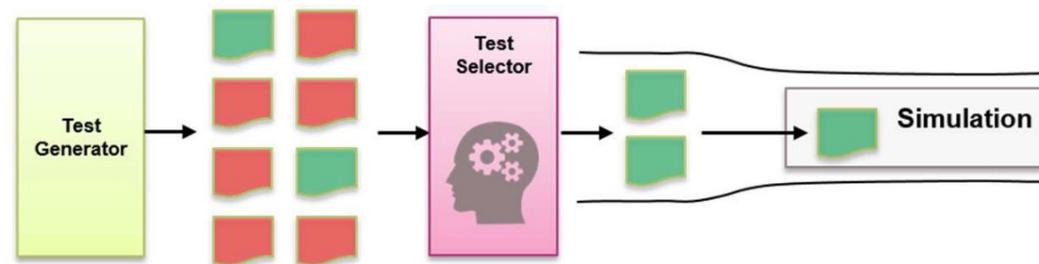
构造蜕变关系，变换测试输入

检测输出变化，判断正确性

测试优先级排序



测试选择



基于聚类的方法

将仿真日志提取为测试用例的特征

根据任务、功能和错误信息进行排序

尽早发现重要错误，加快错误检测

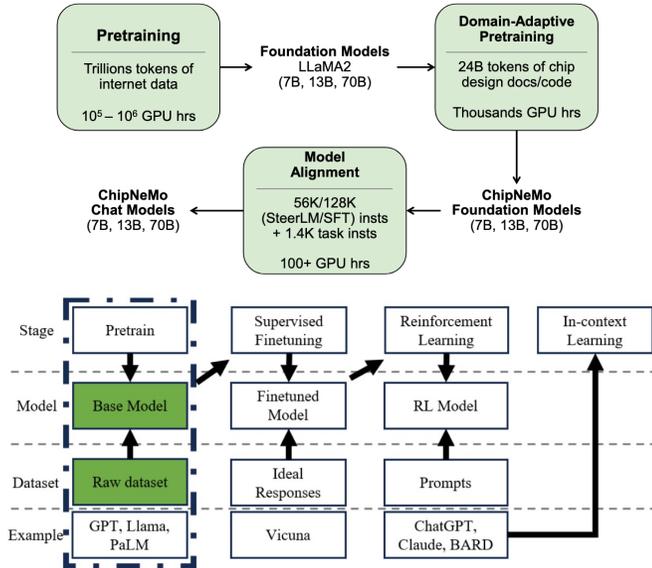
基于无监督学习的方法

对测试输入进行编码和预测

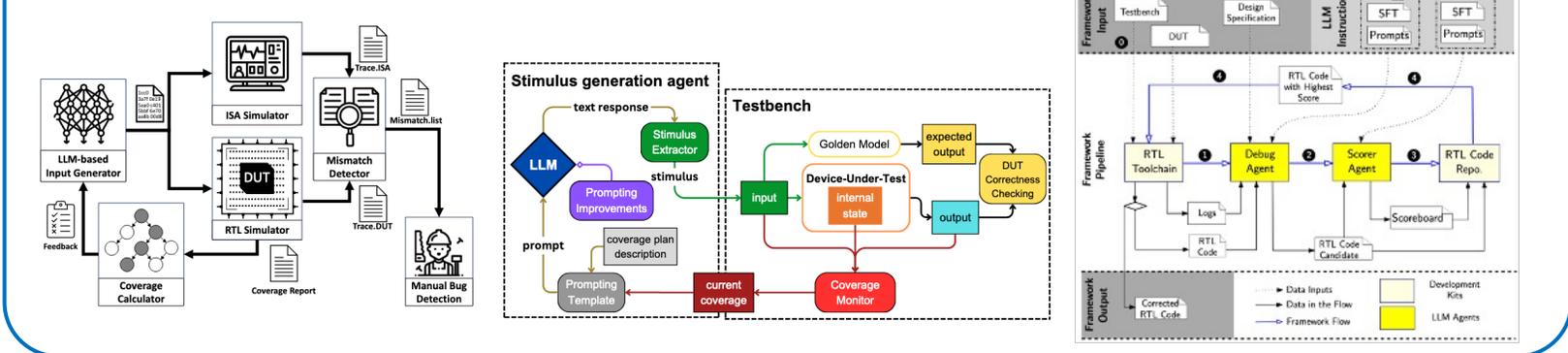
选择异常特征值测试用例

重复仿真，加快覆盖率提升

预训练模型

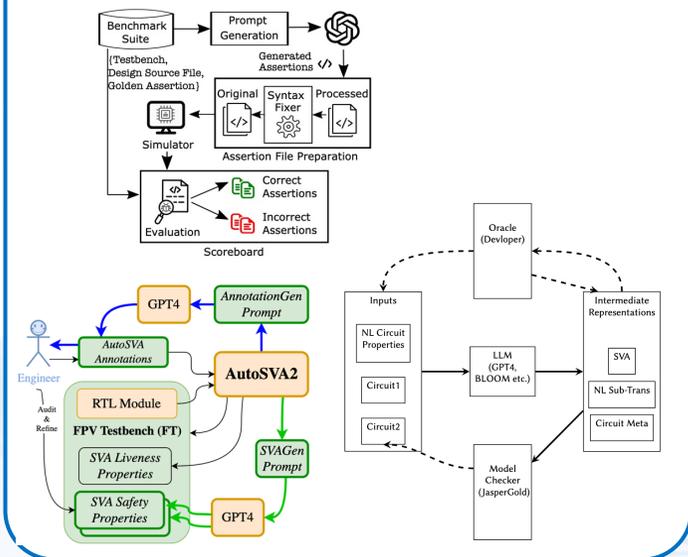


缺陷检测

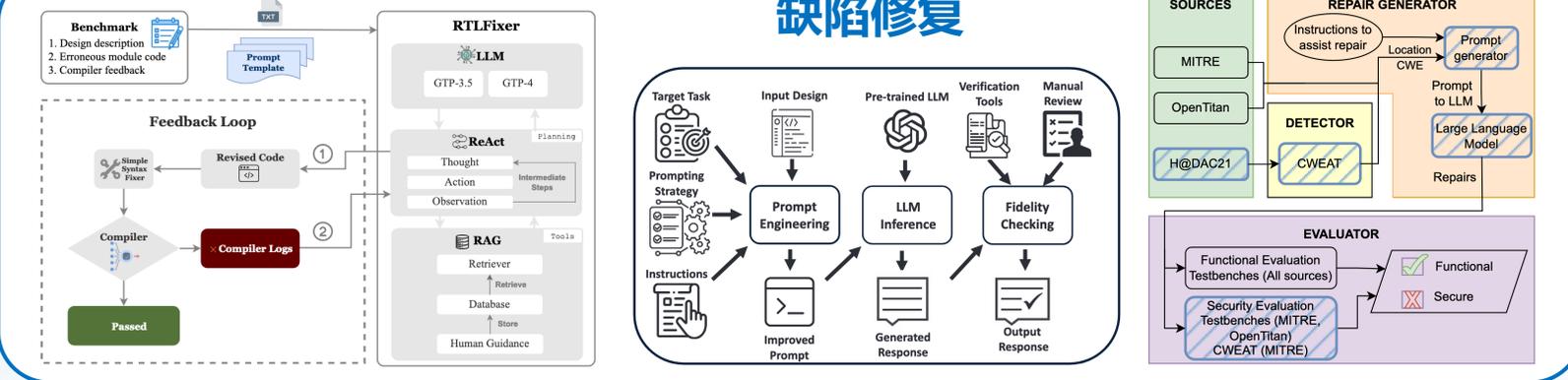


大模型下的 芯片设计程序质量保障

代码 (断言) 生成



缺陷修复





2024 CCF ChinaSoft
中国软件大会

智能化芯片设计程序测试

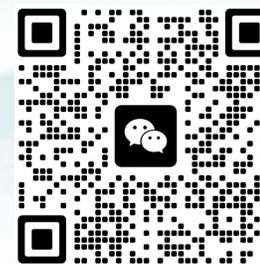
李晓鹏

superpung@tju.edu.cn



天津大学
Tianjin University

<https://xiaopeng.li>



2024年11月16日