

Tai-e 介绍与入门

孔启皓

Peking University

2026 年 10 月 11 日

Tai-e 介绍

- 代码: <https://github.com/pascal-lab/Tai-e>
- 文档: <https://tai-e.pascal-lab.net/docs/0.5.4/reference/en/index.html>
- API: <https://tai-e.pascal-lab.net/docs/0.5.4/api/index.html>

Tai-e 下载与配置

项目数据包中已经包含了 Tai-e 程序，如果你计划在项目模板基础上进行，不用进行额外的下载操作。

配置过程可以参考 Tai-e 框架（教学版）配置指南。南大推荐使用 IntelliJ IDEA 作为开发工具，该工具可以直接配置相关环境并运行。如果你计划使用该工具，可以直接参考配置指南打开 `handout/Tai-e/` 文件夹。

（助教个人感受 IntelliJ IDEA 太臃肿了，不如 vscode 插件好用）

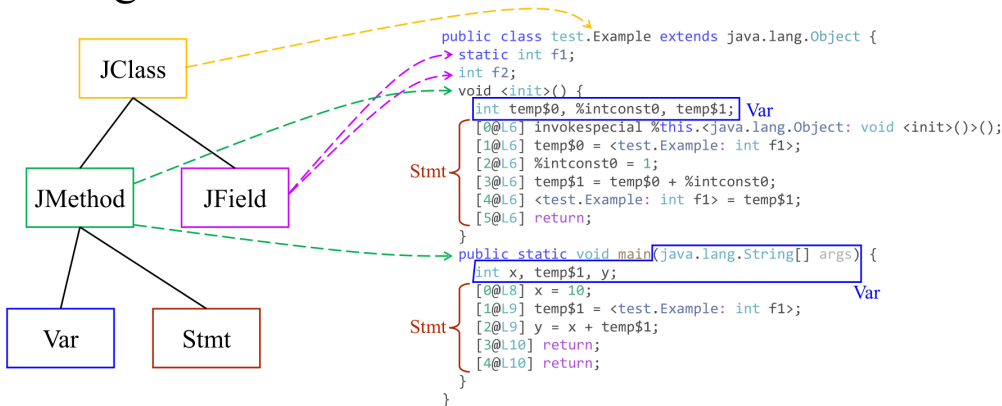
Tai-e 运行

以 windows 系统为例。在 Tai-e 目录下运行 `./gradlew run --args="-cp src/test/pku -m hello -a ir-dumper"` 可以看到生成的 ir 文件（默认生成在 `output/result.txt`）。

- `ir-dumper` 是 Tai-e 内置的一个工具，展示 IR 的文本表示。后续我们的所有操作都是基于 IR 进行的。

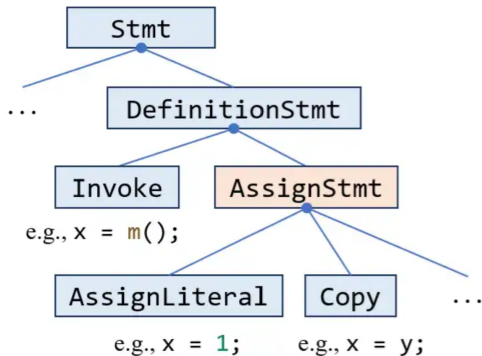
总览

Program Abstraction and Tai-e IR



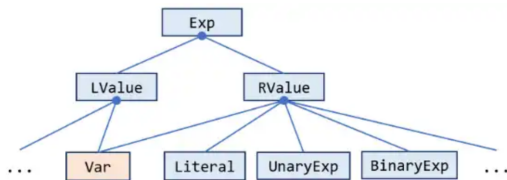
API 调用

当然我们不会直接解析上面的文本形式 IR。Tai-e 提供了一个类 `pascal.taie.ir.IR`。通过方法 `ir.getStmts()`，你可以获取 IR 的详细语句信息。Stmt 的部分类继承信息如下：



IR 中的表达式

Tai-e 转化后的 IR 中每个 Stmt 是一个形如 $x = y \text{ op } z$ 的形式。Stmt 右侧的部分视为一个表达式，在 Tai-e 中以 Exp 类型表示。关于 Exp 的部分继承信息如下：



如何让 Tai-e 运行新的分析

Tai-e 的分析分为三种层级：Method, Class, Program。分别通过继承 `MethodAnalysis`, `ClassAnalysis`, `ProgramAnalysis` 来实现。分别表示该分析作用在 IR、JClass 和全局上。

三种分析方式的调用方法存在差异。具体方法可见 <https://tai-e.pascal-lab.net/docs/current/reference/en/develop-new-analysis.html>。

让 Tai-e 运行新的分析需要在 `src/main/resources/tai-e-analyses.yml` 中注册该分析，填入类和 ID。ID 需要和你编写的类的 ID 一致。

如何让 Tai-e 运行新的分析

在本项目中，由于你的目标是运行一个全局的分析工具，所以应该直接继承 ProgramAnalysis 来实现。

目前模板中已经将 CourseAnalysis.java 注册为 pku-course 分析。你可以修改这个入口，也可以直接在这个文件中编写。

如何编写新的分析

可以仔细阅读南大项目的前三次作业内容 (<https://tai-e.pascal-lab.net/pa1.html> ~ <https://tai-e.pascal-lab.net/pa3.html>)

调用 `World.get().getMainMethod().getIR()`; 可以直接获取主函数的 IR, 你可以在这个基础上进行, 也可以调用 Tai-e 的 `cfg` 来简化分析流程 (也不一定会简化)。

如何编写新的分析

演示：编写一个线性的符号分析工具。

一些提醒

- ① 虽然 Lab-1 的前两个小 lab 保证不会出现类和数组，但由于 Java 规定 `main` 函数必须以 `String[]` 参数，所以实际还是会出现 `String[] args` 以及 `args.length`。你需要对该输入进行特殊判断。
- ② 如果你不需要更改 docker 配置，可以不提交该文件，助教会支持模板依赖的环境下运行项目。同时不要压缩多余的文件，期望你提交的压缩文件不应该超过 10MB。
- ③ 你不应该调用任何直接或间接调用指针分析结果的工具。例如 Tai-e 内置的 ICFG 实际上调用了类之间的包含关系，间接调用了指针分析 PTA，所以不应该被调用。
- ④ 多看文档。懒得翻文档也可以让 AI 帮你看。