

# 编译原理作业 (1)

姓名: 魏恒峰 学号: hfwei@nju.edu.cn

评分: \_\_\_\_\_ 评阅: \_\_\_\_\_

2024 年 3 月 10 日

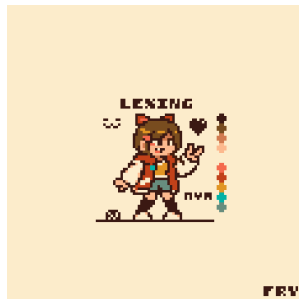
请独立完成作业, 不得抄袭。

若得到他人帮助, 请致谢。

若参考了其它资料, 请给出引用。

鼓励讨论, 但需独立书写解题过程。

**允许并鼓励使用 ChatGPT 等工具, 但需明确说明使用方式。**



## 1 作业 (必做部分)

### 题目 1 (C 语言中的 ANTLR 4 词法规约)

阅读 [C 语言 .g4 文件](#), 完成以下任务 (均可用图示辅助解释)。

- (1) 定位并解释其中的“字符串” (StringLiteral) 词法规则。
- (2) 定义并解释其中的“常量” (IntegerConstant、FloatingConstant、CharacterConstant) 词法规则。
- (3) 使用 `lexer grammar` <sup>①</sup> 在 ANTLR 4 工具中测试 `C.g4` 中的词法单元
  - 考虑如何设计测试用例覆盖尽可能多的情况?
  - 检查你对词法规则的理解是否与 ANTLR 4 的输出一致。
- (4) 其它: 请自行挖掘有趣的内容。

<sup>①</sup> 关于 `lexer grammar` 的用法:

- 见《ANTLR 4 权威指南》第 4.1 节
- 注意: Gradle ANTLR 插件在需要将 `lexer grammar` 导入到更大的 `grammar` 文件中时有一个尚未修复的“幺蛾子” (bug), 参见 [build.gradle 文件](#)。

解答:

---

### 题目 2 (词法分析器代码分析)

查看 [Clang Lexer 文档](#), 阅读 [Clang 词法分析器源码 Lexer.cpp](#), 完成以下任务 (均可用图示辅助解释)。

- 整理函数 `Lexer::Lex()` 的主要逻辑。
- 定义到处理 `StringLiteral` 词法单元的代码, 并分析代码的主要逻辑。
- 定义到处理 `IntegerConstant` 与 `FloatingConstant` 词法单元的代码, 并分析代码的主要逻辑。
- 其它: 请自行挖掘有趣的内容。

解答:

---

## 2 作业 (选做部分)

### 题目 1 (手写词法分析器)

- 为 `FloatingConstant` 词法单元手写词法分析器, 通过与 ANTLR 4 的输出进行对比检查正确性。建议画出状态转移图。

解答:

---

## 3 反馈

请在 Zulip 上讨论对作业或者课程的意见。