

《编译技术》课程设计文档

《编译技术》课程设计文档

词法分析

编码前设计

编码后修改

词法分析

编码前设计

首先定义 `Token` 类、`Type` 枚举类、`Lexer` 类，`Token` 用于记录每个单词的类型、值、所在行数，`Type` 则是由类别码构成的枚举类、`Lexer` 提供词法分析以及存储结果的功能

在 `Lexer` 类中按字符读取从文件中读入的源码，由 `if-else` 分支语句进行类别判断，然后进入对应分支读取完整单词，得到一个 `Token` 实例，不断重复上述操作直到读到源码末尾

- 对于保留字采用 `HashMap` 进行查寻，保留字为键，类别码为值，当得到一个保留字单词就查询 `HashMap` 获得类别码
- 对于注释空白符跳过不识别，即直接读到注释之后的第一个字符
- 对于行号记录，读到 `\n` 字符行号就加一

本次不需要考虑错误处理的问题

编码后修改

与编码前基本一致，但是考虑到评测系统为 `Linux` 系统于是对于要跳过的空白符需要注意判断 `\r` 的情况