**编译原理实验报告**

基于有限自动机的词法分析器构造

141250179 袁阳阳

# 目标

本次实验的主要目的是对自定义的程序语言进行词法分析器程序构造，从而更好地理解词法分析原理。程序读取一个java程序文件，并对其中的内容进行词法分析，可识别保留字（为简化程序选取一部分有代表性的保留字进行识别）、变量名、数字、操作符（为简化程序选取一部分有代表性的操作符进行识别）、注释符（保留字优先于变量名），并输出格式为<tokenType, str>的token序列，并对未识别字符、整数过大等异常情况进行报错处理。

# 内容概述

本报告主要描述了一个简单的词法分析器构造过程，包括实现过程中的理论推导，具体的核心算法和数据结构的描述，最终成品的功能概要和个人的感悟体会。

# 假设与依赖

## 3.1实验环境

操作系统：windows

编程语言：java

## 3.2语言定义

### 3.2.1保留字

保留字识别在语义上没有太大的区别，为简化程序复杂度，增强可读性，本程序仅选取一部分具有代表性的java保留字进行识别，包括 public, static, int, void, if, else, while.

### 3.2.2特殊符号

同样为了程序的简化，运算符只选取具有代表性的运算符子集

|  |  |
| --- | --- |
| 符号表达 | 符号类别 |
| +, -, \*, /，+=，++ | 算数运算符 |
| <,>,<=,==， |  |
|  |  |
|  |  |
|  |  |

# 主要思路和方法

# 相关的有限自动机描述

# 主要数据结构描述

# 核心算法描述

# 测试用例

# 困难与解决方案

# 10.总结与收获