
编译原理实验指导书

实验 1 词法分析

实验目的

1. 巩固对词法分析的基本功能和原理的认识。
2. 能够应用自动机的知识进行词法分析。
3. 理解并处理词法分析中的异常和错误。

实验内容

设计实现类高级语言的词法分析器，**基本功能**为识别以下几类单词：

- 标识符（由大小写字母、数字以及下划线组成，但必须以字母或者下划线开头）
- 关键字（①类型关键字：整型、浮点型、布尔型、记录型；②分支结构中的if和else；③循环结构中的do和while；④过程声明和调用中的关键字）
- 运算符（①算术运算符；②关系运算符；③逻辑运算）
- 界符（①用于赋值语句的界符，如“=”；②用于句子结尾的界符，如“;”；③用于数组表示的界符，如“[”和“]”；④用于浮点数表示的界符“.”）
- 常数（无符号整数和浮点数，包括科学计数法，字符串常数等）
- 注释（/*.....*/形式）

除此之外，可以实现一些**额外功能**，如

- （1）识别字符常数、八进制和十六进制数。
- （2）能够应用子集构造法从输入的NFA自动生成DFA。

实验要求

- （1）要求**基于DFA技术**设计词法分析器。
- （2）系统的输入形式：要求能够通过文件导入**FA转换表**和**测试用例**，可以

通过用户界面显示并编辑测试用例。测试用例要涵盖“实验内容”中列出的各类单词，并包含各种单词拼写错误。

(3) 系统的输出分为两部分：一部分是打印输出词法分析器的DFA转换表。另一部分是打印输出源程序对应的token序列，格式如下图所示：

```
■ 输入
  ■ while(num!=100){num++;}
■ 输出
while    <WHILE,  _  >
(        < SLP ,  _  >
num      < IDN , num >
!=       < NE ,   _  >
100      <CONST, 100 >
)        < SRP ,  _  >
{        < LP ,   _  >
num      < IDN , num >
++       < INC ,  _  >
;        < SEMI,  _  >
}        < RP ,   _  >
```

(3) 要求实验之前完成实验报告中的预习部分（即需求分析、文法设计和系统设计三个板块）。未按时完成预习报告者将扣除相应分数。

(4) 要求当堂完成实验内容，并进行现场验收。未当堂完成验收者将扣除相应分数。

(5) 要求实验结束后一周内提交实验报告及源程序。未按时提交者将影响成绩评定。

实验评分标准

一、课堂表现（10分）

1. 出勤情况（迟到，早退，缺席）
2. 是否遵守课堂纪律

二、操作表现（50分）

1. 功能齐全，结果正确无误（30 分）
2. 界面美观、人性化，具有良好演示效果（10 分）
3. 当堂按时完成（10 分）

三、实验报告（40分）

1. 需求分析（10 分）
2. 设计（20 分）

3. 系统实现及结果分析（10 分）