

# 编译原理课程实验报告

## 实验 1：词法分析

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |         |    |        |          |            |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|----|--------|----------|------------|--|
| 姓名                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 张泽宇     | 院系 | 计算机学院  | 学号       | 1163300620 |  |
| 任课教师                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 辛明影     |    | 指导教师   | 辛明影      |            |  |
| 实验地点                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 格物 213  |    | 实验时间   | 4 月 14 日 |            |  |
| 实验课表现                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 出勤、表现得分 |    | 实验报告得分 |          | 实验总分       |  |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 操作结果得分  |    |        |          |            |  |
| 一、需求分析                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |         |    |        |          | 得分         |  |
| 词法分析系统所要完成的功能：<br>1. 能识别以下几类单词：<br>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）；<br>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）；<br>运算符（①算术运算符；②关系运算符；③逻辑运算）；<br>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“.”）；<br>常数（无符号整数和浮点数，包括科学计数法，字符串常数等）；<br>注释（/*.....*/形式）；<br>八进制数和十六进制数。<br>2. 能通过文件导入 FA 转换表以及测试用例。<br>3. 在用户界面显示并编辑测试用例。<br>4. 在用户界面显示翻译结果以及错误提示。                                                                                                                                                                                                                                                                         |         |    |        |          |            |  |
| 二、文法设计                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |         |    |        |          | 得分         |  |
| (1) 给出各类单词的词法规则描述<br>1.标识符：<br>(A B ... Z a b ... z _)(A B ... Z a b ... z _0 1 2 ... 9)*<br>2.关键字：<br>char   long   short   float   double   const   Boolean  void  null  false  true  enum  int <br>do  while  if  else for  then   break   continue   class   static  final  extends new  return  <br>signed  struct   union  unsigned  goto   switch case  default  auto  extern   register  sizeof  <br>typedef   volatile<br>3.运算符：<br>> >= < <= == !=   &    && ! ^ + - * / % ++ -- += -= *= /= <br>4.界符：<br>, = : [] () { } . '"<br>5.常数：<br>(0 1 2 ... 9)(0 1 2 ... 9)*((.(0 1 2 ... 9)(0 1 2 ... 9)*) ε)<br>((E(+ -)ε)((0 1 2 ... 9)(0 1 2 ... 9)*) ε)<br>6.注释：<br>/*(注释文本)*/ |         |    |        |          |            |  |

7.八进制数:

$0(1|2|3|4|5|6|7)(0|1|2|3|4|5|6|7)^*$

8.十六进制数:

$0x(1|...|9|a|...|f|A|...|F)(0|...|9|a|...|f|A|...|F)^*$

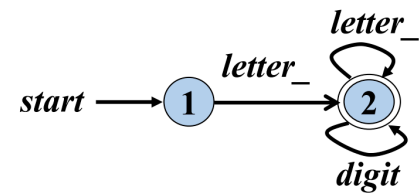
(2) 各类单词的转换图

1.标识符、关键字

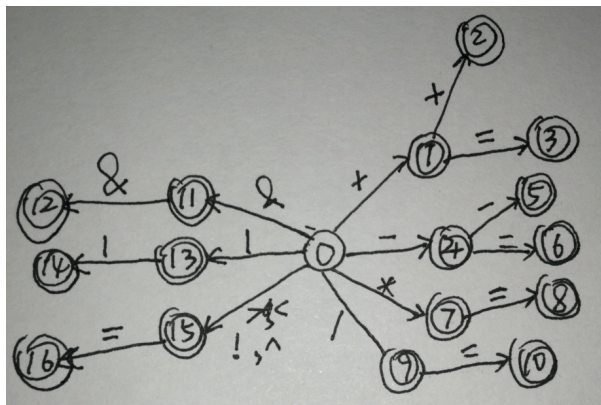
➤  $digit \rightarrow 0|1|2|...|9$

➤  $letter\_ \rightarrow A|B|...|Z|a|b|...|z|_$

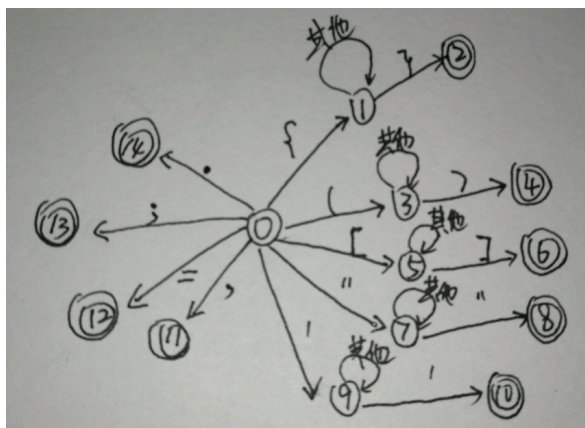
➤  $id \rightarrow letter\_ (letter\_ | digit)^*$



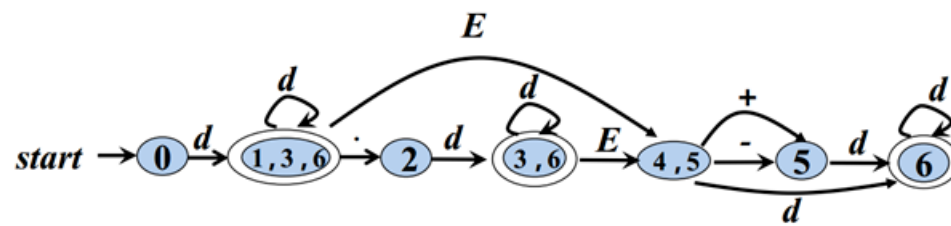
2.运算符



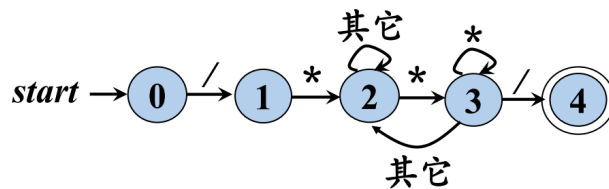
3.界符



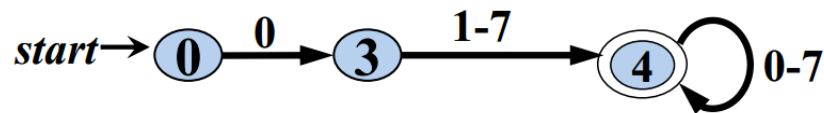
#### 4.常数



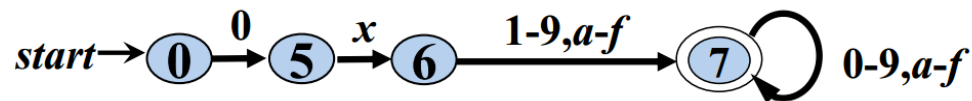
#### 5.注释:



#### 6.八进制



#### 7.十六进制



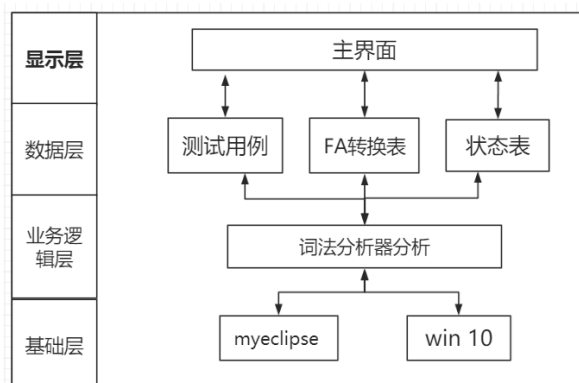
### 三、系统设计

得分

要求：分为系统概要设计和系统详细设计。

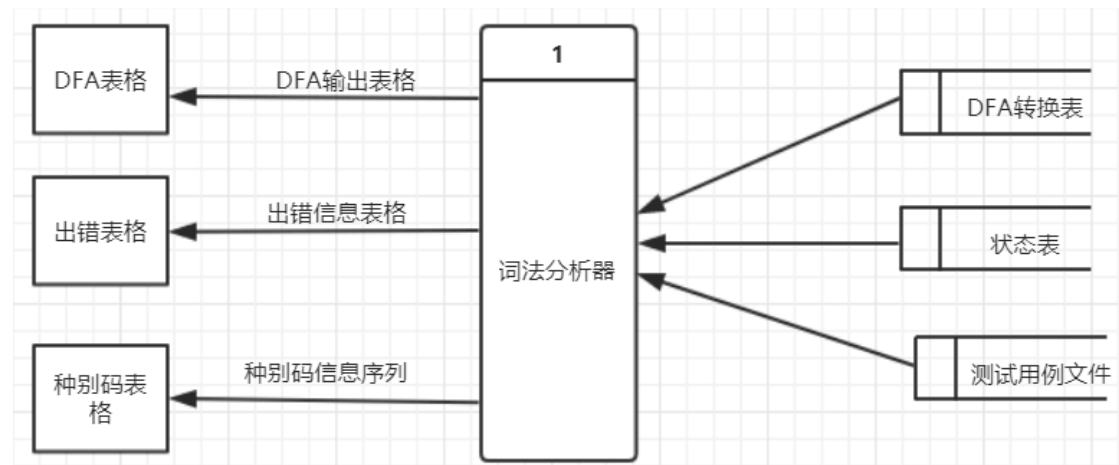
(1) 系统概要设计：给出必要的系统宏观层面设计图，如系统框架图、数据流图、功能模块图等以及相应的文字说明。

#### 1.系统框架图：



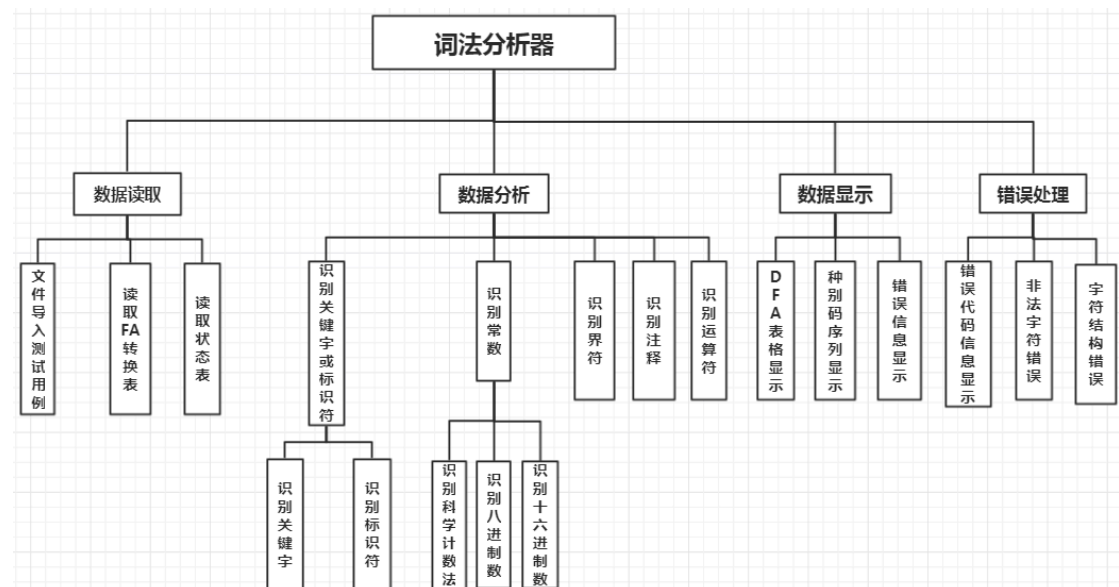
其中，显示层是用户界面的显示，数据层对应的是测试用例以及 FA 转换表的读取，逻辑层是对数据进行分析的核心部分，基础层表示硬件条件。

## 2.数据流图:



通过文件导入 FA 转换表和状态的信息列表，词法分析器分析，根据导入的测试用例开始分析，相关结果在用户界面中显示出来。

## 3.功能模块图:



该词法分析器主要有四大功能：对数据的读取、分析、显示和错误检测。

(2) 系统详细设计：对如下工作进行展开描述

✓ 核心数据结构的设计

DFA Table: 对 FA 表的读入

有以下属性: `int state`: 当前状态

`String[] input`: 输入符号集合

`int nextState`: 下一状态

作用：将从 FA 转换表中读到的内容存入 DFATable 对象中，之后用到直接调 get set 方法

DFATableState：对状态表的读入

有以下属性：**int state**：当前状态

**String type**：该状态对应的应识别出的类型

**boolean isFinish**：是否是终止状态标识

作用：将从状态表中读到的内容存入 DFATableState 对象中，之后用到直接调 get set 方法，用于判断当前字符串的类型和出错检测。

#### ✓ 主要功能函数说明

1.getData(File file, int ignoreRows) : String[][]

读取 Excel 的内容，第一维数组存储的是一行中格列的值，二维数组存储的是多少个行

file 读取数据的源 Excel ignoreRows 读取数据忽略的行数

2.lexicalAnalysis(String str): void

读入已转换为字符串的测试用例，进行词法分析。为最核心函数。

3.stateChange(char currentChar,int currentState,DFATable[] dfa) : int

读入当前指向的字符，当前所处状态和 DFA 表

遍历 dfa 表，根据当前输入字符和状态寻找下一状态，找到返回下一状态，没找到判断输入是否存在，不存在则输入为非法字符，返回-2，否则返回 -1

4.tokenID(String str): int

用于返回种别码

开始给已存在的关键字，符号一个初始种别码，读入字符串，返回该字符串的种别码，如该字符串不在已存在的种别码序列中，加入进去并返回

5.isKeyword(String str): boolean

判断是否为关键字

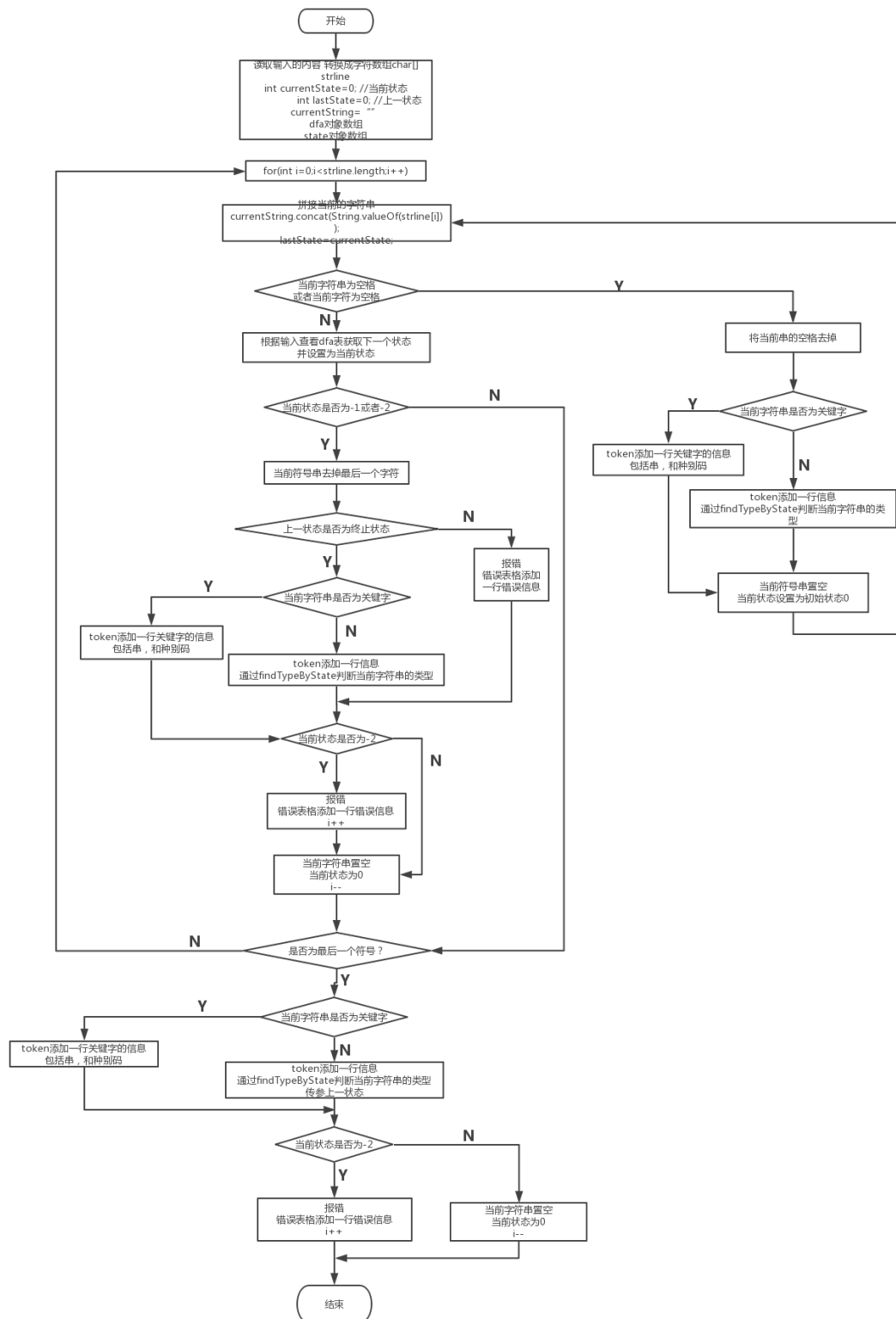
6.findTypeByState(int state,DFATableState[] dfaState): String

根据状态找到对应的类型

7.isList(String[] arr, char currentChar): boolean

判断是否有该输入

✓



|             |    |  |
|-------------|----|--|
| 四、系统实现及结果分析 | 得分 |  |
|-------------|----|--|

要求：对如下内容展开描述。

- (1) 系统实现过程中遇到的问题；
1. DFA 表的数据包括当前状态，输入，下一状态，JAVA 不方便三元组。  
解决办法：把数据放到对象里，采用对象数组读取数据。
  2. 按照开始那套顺序来遇到非法字符死循环，总是在在判断非法字符  
解决办法：遇到非法字符输出，并且使程序 i+1,continue;
  3. 最后一个字符能检测到但是显示不出来  
解决办法：增加一个判断是否为最后一个字符，是则根据上一状态判断最后一个字符串的类型。

(2) 针对某测试程序输出其词法分析结果：  
导入测试用例如下：

```
int x=0x12AB,z=0;/*0x12Wa*/
chardemo_@='a';
Double y=5.23E+5;
Float k=0.25E+;/*error*/
Char c[5];
Do
{
c[z]=z;
z++;
}while(z<=5);
if(y>5||z<3)
{
demo_1='B';
}
```

输出结果如下：

| TOKEN TABLE |         |     |         |
|-------------|---------|-----|---------|
| 字符串         | 类别      | 种别码 | value   |
| int         | 关键字     | 12  |         |
| x           | 标识符     | 76  | x       |
| =           | 界符      | 65  |         |
| 0x12AB      | 十六进制数   | 77  | 0x12AB  |
| ,           | 界符      | 64  |         |
| z           | 标识符     | 78  | z       |
| =           | 界符      | 65  |         |
| 0           | 整数      | 79  | 0       |
| .           | 界符      | 66  |         |
| /*0x12WA*/  | 注释      | 80  |         |
| char        | 关键字     | 0   |         |
| demo_1      | 标识符     | 81  | demo_1  |
| @           | 非法字符    | 无   |         |
| =           | 界符      | 65  |         |
| '           | 界符      | 75  |         |
| a           | 标识符     | 82  | a       |
| '           | 界符      | 75  |         |
| .           | 界符      | 66  |         |
| double      | 关键字     | 4   |         |
| y           | 标识符     | 83  | y       |
| =           | 界符      | 65  |         |
| 5.23E+5     | 科学计数法常数 | 84  | 5.23E+5 |

| 字符串       | 类别      | 种别码 | value   |
|-----------|---------|-----|---------|
| 5.23E+5   | 科学计数法常数 | 84  | 5.23E+5 |
| :         | 界符      | 66  |         |
| float     | 关键字     | 3   |         |
| k         | 标识符     | 95  | k       |
| =         | 界符      | 65  |         |
| 0.25E+    | 非法字符    | 无   |         |
| :         | 界符      | 66  |         |
| /*error*/ | 注释      | 87  |         |
| char      | 关键字     | 0   |         |
| c         | 标识符     | 88  | c       |
| [         | 界符      | 67  |         |
| 5         | 整数      | 89  | 5       |
| ]         | 界符      | 68  |         |
| :         | 界符      | 66  |         |
| do        | 关键字     | 13  |         |
| {         | 界符      | 71  |         |
| c         | 标识符     | 88  | c       |
| [         | 界符      | 67  |         |
| z         | 标识符     | 78  | z       |
| ]         | 界符      | 68  |         |
| =         | 界符      | 65  |         |
| z         | 标识符     | 78  | z       |

| 字符串   | 类别  | 种别码 | value |
|-------|-----|-----|-------|
| :     | 界符  | 66  |       |
| =     | 界符  | 65  |       |
| z     | 标识符 | 78  | z     |
| :     | 界符  | 66  |       |
| z     | 标识符 | 78  | z     |
| ++    | 运算符 | 58  |       |
| :     | 界符  | 66  |       |
| }     | 界符  | 72  |       |
| while | 关键字 | 14  |       |
| (     | 界符  | 69  |       |
| z     | 标识符 | 78  | z     |
| <=    | 运算符 | 44  |       |
| 5     | 整数  | 89  | 5     |
| )     | 界符  | 70  |       |
| :     | 界符  | 66  |       |
| if    | 关键字 | 15  |       |
| (     | 界符  | 69  |       |
| y     | 标识符 | 83  | y     |
| >     | 运算符 | 41  |       |
| 5     | 整数  | 89  | 5     |
|       | 运算符 | 49  |       |
| z     | 标识符 | 78  | z     |

| 字符串    | 类别  | 种别码 | value  |
|--------|-----|-----|--------|
| )      | 界符  | 70  |        |
| :      | 界符  | 66  |        |
| if     | 关键字 | 15  |        |
| (      | 界符  | 69  |        |
| y      | 标识符 | 83  | y      |
| >      | 运算符 | 41  |        |
| 5      | 整数  | 89  | 5      |
|        | 运算符 | 49  |        |
| z      | 标识符 | 78  | z      |
| <      | 运算符 | 43  |        |
| 3      | 整数  | 90  | 3      |
| )      | 界符  | 70  |        |
| {      | 界符  | 71  |        |
|        | 初始  | 91  |        |
|        | 初始  | 91  |        |
| demo_1 | 标识符 | 81  | demo_1 |
| =      | 界符  | 65  |        |
| '      | 界符  | 75  |        |
| B      | 标识符 | 92  | B      |
| '      | 界符  | 75  |        |
| :      | 界符  | 66  |        |
| }      | 界符  | 72  |        |



(3) 输出针对此测试程序对应的词法错误报告;

| ERROR TABLE |       |      |
|-------------|-------|------|
| 出错符号        | 出错地方  | 出错原因 |
| @           | 第38字符 | 非法字符 |
| 0.25E+      | 第75字符 | 非法字符 |

(4) 对实验结果进行分析。

选择文件

测试

TOKEN TABLE

| 字符串    | 类别  | 种别码 | value  |
|--------|-----|-----|--------|
| /      | 齐符  | 70  |        |
| :      | 界符  | 66  |        |
| if     | 关键字 | 15  |        |
| (      | 界符  | 69  |        |
| y      | 标识符 | 83  | y      |
| >      | 运算符 | 41  |        |
| 5      | 整数  | 89  | 5      |
|        | 运算符 | 49  |        |
| z      | 标识符 | 78  | z      |
| <      | 运算符 | 43  |        |
| 3      | 整数  | 90  | 3      |
| )      | 界符  | 70  |        |
| {      | 界符  | 71  |        |
|        | 初始  | 91  |        |
|        | 初始  | 91  |        |
| demo_1 | 标识符 | 81  | demo_1 |
| =      | 界符  | 65  |        |
| '      | 界符  | 75  |        |
| B      | 标识符 | 92  | B      |
| '      | 界符  | 75  |        |
| :      | 界符  | 66  |        |
| )      | 界符  | 72  |        |

ERROR TABLE

| 出错符号   | 出错地方  | 出错原因 |
|--------|-------|------|
| @      | 第38字符 | 非法字符 |
| 0.25E+ | 第75字符 | 非法字符 |

可以看到，在导入测试文件后，能正确显示其词法分析结果，其中 @符号为非法输入，检测出，int char 等关键字，x，z demo\_1 等标识符，识别出，注释，0x12AB 十六进制，5.23E+5 科学计数法，识别出，0,3,5 常数和各种符号都能被准确识别。

指导教师评语：

日期：