

# 编译原理课程实验报告

## 实验 1：词法分析

姓名	马宇斌	院系	软件学院	学号	161110407
任课教师	韩希先	指导教师	韩希先		
实验地点	研究院中 517	实验时间	2018.10.20		
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、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。
4. 在掌握基本知识后，通过编写程序一步步测试来实现以上目的

### 二、实验内容

要求：对如下工作进行展开描述

#### 1. 给出语言的词法规则描述

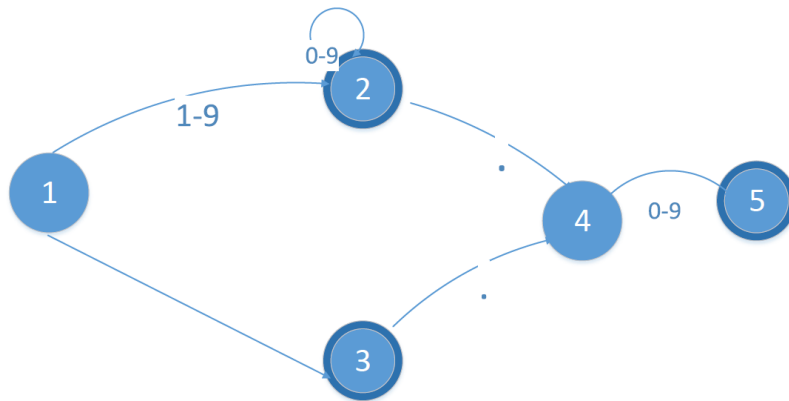
- 标识符、关键字、整常数、字符常数、浮点常数
- 单界符：+，-，×，÷，...
- 双界符：/\*，:=，...
- 注释

根据所学知识得，单词是语言中具有独立意义最小单位，所以我们根据输入的源程序，按照一系列的构词规则来对单词进行分析检测。我们将单词分为：关键字、标识符、运算符、界符以及常量等。

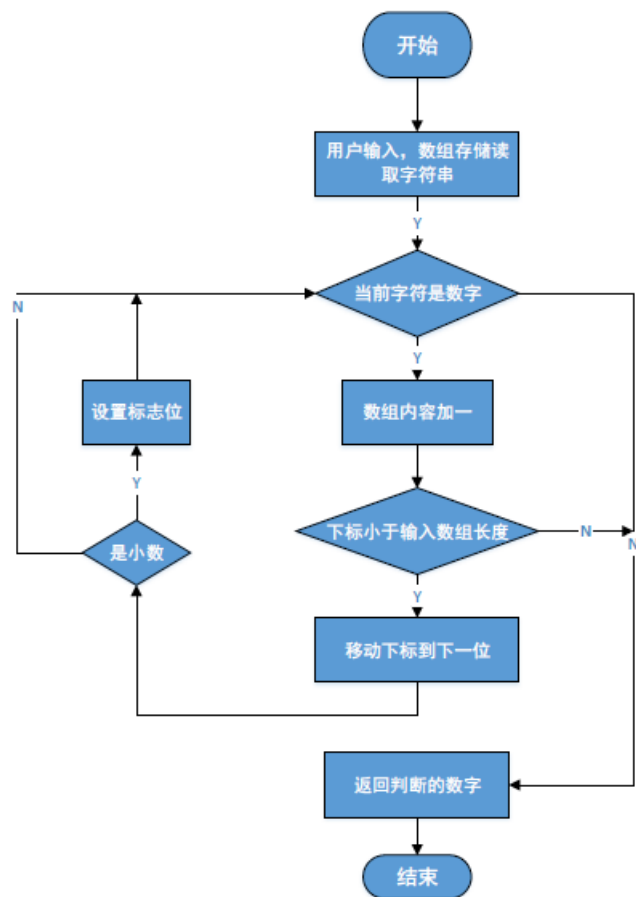
- ① 标识符：以字母开头，有字母、数字或‘\_’构成
- ② 关键字：构成方式同标志符，为{if, else, do, while, then}之一
- ③ 常整数：常数的类型一般有整数、浮点数等。例如，100, 3.14。
- ④ 字符常数：字符常量是用单引号括起来的单个字符构成
- ⑤ 字符串常数：字符串常量是一对双引号括起来的字符序列。例如："Hello"
- ⑥ 运算符：+、-、\*、/等
- ⑦ 界符：逗号，分号，括号等

#### 2. 针对这种单词的状态转换图和程序框图

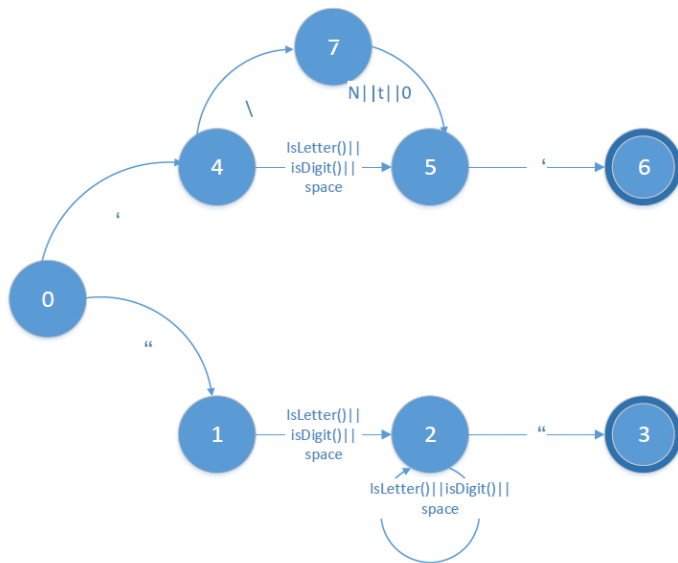
常数的状态转换图：



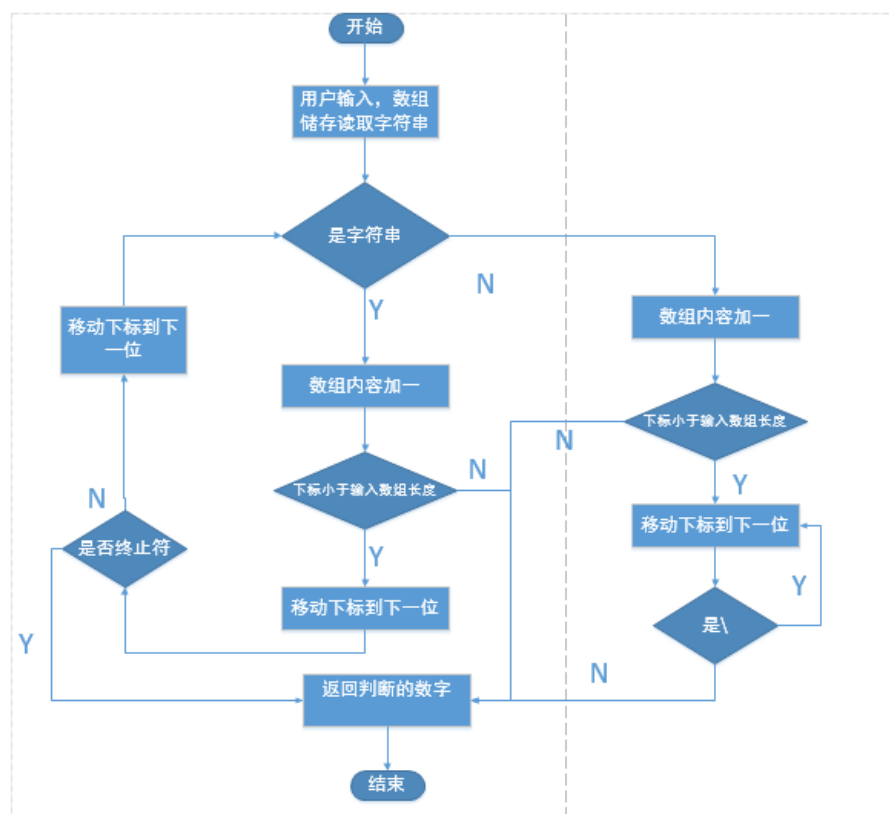
常数的流程图：



字符和字符串的状态转换图：



字符和字符串的流程图：



### 3.核心数据结构的设计

符号表表的设计为如果是标识符，则显示（标识符，标识符名称），如果是标点符号或者关系运算符或者逻辑运算符，则显示（该符号，--）。关键字在词法分析程序中已经保存到一个二维数组中，当匹配时，则输出（关键字，该关键字的值）。如果是常数，则显示（常数，常数的值）。

### 4.错误处理：错误的位置及类型等

如果在进行单词的识别的时候，没有进入任何一种状态，则进行错误处理，记录下出错

的字符，读取下一个字符，如果进入到下一个状态则继续；否则一直出错，知道读到源文件的末尾。

错误的类型主要有非法字符的输入，关键字拼写错误，括号不封闭等

### 三、实验结果

要求：将实验获得的结果进行描述，基本内容包括：

1.针对某测试程序输出其词法分析结果：

测试程序为：

```
int a=0;
while(a>-10)
{
    a++;
};
```

词法分析输出结果：

int	(13, -)	关键字
a	(12, 1)	标识符
=	(14, -)	
0	(13, 1)	
;	(25, -)	
while	(20, -)	关键字
(	(26, -)	
a	(12, 1)	标识符
>	(17, -)	
-	(21, -)	
10	(13, 2)	
)	(27, -)	
{	(28, -)	
a	(12, 1)	标识符
+	(20, -)	
+	(20, -)	
;	(25, -)	
}	(29, -)	
;	(25, -)	

■

### 四、实验中遇到的问题总结

要求：主要阐述两方面的问题

(一) 实验过程中遇到的问题如何解决的？

问题一：标识符和关键字的区分

首先将已有的关键字（自己定义）的保存到一个二维数组 A 中，然后将文件中读到的关键词和标识符都存入到一个二维数组 B 当中，当进行输出单词序列时，在输出之前对 B 进行判断即可，判断 B[i]是否在 A 中，如果在就声明为关键字，否则声明为标识

符。

## （二）思考题的思考与分析

思考题 1：你编写的程序是如何体现自动机的，这样做有什么好处？

通过读取输入的字符串，然后一个字符一个字符的读取，然后根据读取到的字符跳转到自动机中，例如是数字的话就跳转到实数识别自动机中，如果是字母则跳转到关键字的自动机中，同理依次类推。

好处：条理清晰，顺序流畅，遇到错误时，在对应的自动机进行相应的处理。简单易实现。

思考题 2：符号表是怎样处理的，可不可以提出改进方法？

符号表是通过存储在数组中实现的。相对来说每输入一个字符后就遍历一遍字符表数组，效率有些低。对于符号表的改进，我想可以使用哈希映射的方法，每种符号单独建一个哈希映射集，所以的映射集均由一个封装的统一接口访问。这样既能节省空间，又能在符号表中充分体现符号的信息，还提高了访问的效率。

## 五、实验体会

通过词法分析实验，首先我认识到词法分析就是将字符序列转换为单词（Token）序列的过程。词法分析器一般以函数的形式存在，供语法分析器调用。词法分析阶段是编译过程的第一个阶段，这个阶段的任务是从左到右一个字符一个字符地读入源程序，即对构成源程序的字符流进行扫描然后根据构词规则识别单词(也称单词符号或符号)。并且对符号表有了一定的理解。

指导教师评语：

日期：