第一部分 词法分析（实验一必作）

实验一 词法分析程序设计与实现

一、实验目的

加深对词法分析器的工作过程的理解；加强对词法分析方法的掌握；能够采用一种编程语言实现简单的词法分析程序；能够使用自己编写的分析程序对简单的程序段进行词法分析。

二、实验内容

自定义一种程序设计语言，或者选择已有的一种高级语言，编制它的词法分析程序。词法分析程序的实现可以采用任何一种编程语言和编程工具。

程序能够从左到右一个字符一个字符地读入源程序，并对构成的源程序的字符流进行扫描和分解，从而识别出一个个单词（也称单词符号或符号）。并给出单词的值和属性。识别出各个具有独立意义的单词，即关键字、标识符、常数、运算符、界符(状态转换图见书P93)。并依次输出各个单词的内部编码及单词符号自身值。（遇到错误时可显示“Error”，然后跳过错误部分继续显示）

三、实验要求：

1. 对单词的构词规则有明确的定义；
2. 编写的分析程序能够正确识别源程序中的单词符号；
3. 识别出的单词以<种别码，值>的形式保存在符号表中，正确设计和维护符号表；
4. 对于源程序中的词法错误，能够做出简单的错误处理，给出简单的错误提示，保证顺利完成整个源程序的词法分析；

四、实验步骤

1. 定义目标语言的可用符号表和构词规则；
2. 依次读入源程序符号，对源程序进行单词切分和识别，直到源程序结束；
3. 对正确的单词，按照它的种别以<种别码，值>的形式保存在符号表中；
4. 对不正确的单词，做出错误处理。

五、实验报告要求

详细说明你的程序的设计思路和实现过程。用有限自动机或者文法的形式对词法定义做出详细说明，说明词法分析程序的工作过程，说明错误处理的实现，写出实验心得，给出程序实现的代码。

六、参考资料

1、以下面一段程序为例

main()

{

int a,b;

a = 10;

b := a + 20;

}

2、需要识别的词

1. 关键字：if、int、for、while、do、return、break、continue；单词种别码为1。
2. 标识符；单词种别码为2。
3. 常数为无符号整形数；单词种别码为3。
4. 运算符包括：+、-、\*、/、\*\*、=、:=、<=、<>、<、>=、>单词种别码为4。
5. 分隔符包括：,、； 单词种别码为5。

3、程序输出形式

要求输出下面的形式：

（2，”main”）

（5，”（“）

（5，”）“）

（5，”{“）

（1，”int”）

（2，”a”）

（5，”,”）

（2，”b”）

（5，”;”）

（2，”a”）

（4，”=”）

（3，”10”）

（5，”;”）

（2，”b”）

（4，”:=”）

（2，”a”）

（4，”+”）

（3，”20”）

（5，”;”）

（5，”}“）