**课程名称：　　编译原理与技术**

**实验名称：　　词法分析程序的设计与实现**

# 一、实验目的

学习如何使用程序编写词法分析器。

# 二、实验内容

编写C++词法分析程序，使之可以实现以下功能：

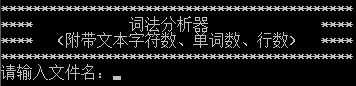
1. 识别出C++语言编写的源程序中每个单词符号，并以记号的形式输出每个单词符号。
2. 识别并读取源程序中的注释。
3. 添加符号表。
4. 统计源程序中的语句行数、单词个数和字符个数，其中标点和空格不计算为单词，并输出统计结果。

# 三、实验环境

Visual Studio 2010

# 四、实验结果

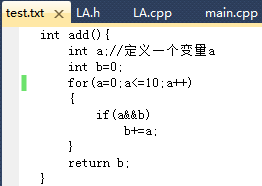
**界面：**



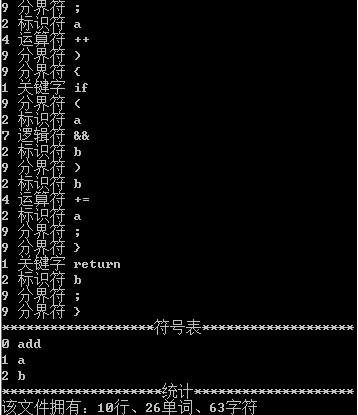
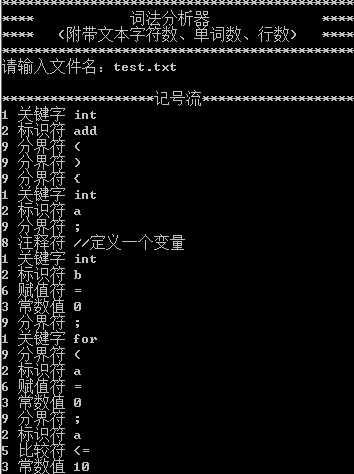
输入文件名后，文件可能打开成功或失败：

1. **成功：**

**①、文件：**



**②、结果：**



**③、功能：**

**记号流：**



1. 关键字：
2. 标识符：
3. 常数值：
4. 运算符：
5. 比较符：
6. 赋值符：
7. 逻辑符：
8. 注释符：
9. 分界符：

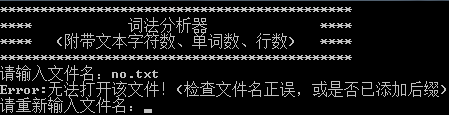
**符号表：**



**统计：**



1. **失败（重新输入）：**



# 五、附录

**（一）、分析思路：**

Case 1: （字母或“\_”开头为关键字或变量）

关键字判断得出关键字或变量。

Case 2: （数字开头为数值）

Default: （各种符号）

判断该字符或加上下一个字符为哪种符号。

**（二）、调试心得：**

1. 由于不同操作系统的文本存储对换行符的处理不同，可能为“\n”或“\r\n”等，所以为避免漏洞，本程序不采用逐一字符读取，而是逐行读取，再对该行进行字符、单词分析。
2. 由分析可知，若单词以字母或“\_”开头，则该关键字为关键字或变量，这里可由关键字判断区分；若单词以数字开头，则该单词为数值；其他情况则是各种符号，这里可进一步判断具体符号。
3. 空格可区分单词与单词，而有时单词与符号之间没有空格，因此空格不能作为单词的判断条件。在此我通过顺序循环读取至非字母或数字的方法取出关键字、变量或数值。
4. 当单词以字母、“\_”或数字开头时，要继续提取后面的字符，应该注意变量非开头部分可以是数字，数值间可有小数点。
5. 关键字的判断通过单词与已给出的关键字比较得出，而各种符号的判断因为符号的复杂性与符号位置的复杂性不好直接取出比较，因此通过直接判断是什么字符得出。
6. 程序除了应该忽略空格外，应该注意“\t”也要忽略。
7. 除了空格和“\t”，其他字符都将被统计到总字符数中。而单词的统计除了关键字、变量和数值外，各种符号也作为单词统计，且将区分类似于“+”还是“++”的格式，不是单纯一个符号一个单词。
8. 在对单词的提取过程中，由于单词大小不定，导致存储该单词的变量的内存空间不定，且对注释的提取可能发现需要较大内存空间，因此采用动态分配空间的方法。这里使用realloc函数，它可以为变量重新分配内存空间，并复制原本内容，把原先内存释放。值得注意的是存储的值后面不知为何多出了“驱驱驱驱”，因此把最后指针指向NULL，防止出现乱码。
9. 个人使用vector存储符号表，若它存储某函数内某变量时，若这个变量值在这个函数内变化，则vector存储的值也会变化。