

用C语言实现文本搜索，类似软件中常用的搜索功能

关注我们： 微信公众号 新浪微博 读者QQ交流群：302185285

C语言中文网推出辅导班啦，包括「C语言辅导班、C++辅导班、算法/数据结构辅导班」，全部都是一对一教学：一对一辅导 + 一对一答疑 + 布置作业 + 项目实践 + 永久学习。QQ在线，随时响应！

文本检索是软件最常用的一个功能，下面的代码将教你如果使用C语言统计单词的个数。

```
#include <stdlib.h>
#include <stdio.h>
#include <string.h>

// argc：统计运行程序时命令行参数的个数
// *argv[]：储存每个参数的字符串指针，每一个元素对应一个参数
int main (int argc, char *argv[])
{
    int ch, bian=0, count=0;
    char find[50], temp[1000];
    FILE *pt;    //文件指针

    // 判断是否输入文件
    if (argc != 2){
        printf("请使用格式： %s 文件名", argv[0]);
        exit(1);    //非正常退出
    }

    // 判断能否成功打开文件
    if ((pt = fopen(argv[1], "r")) == NULL){    //将argv[1]赋值给指针pt
        printf("打开文件 %s 失败", argv[1]);
        exit(1);
    }

    printf("请输入要查找的单词:");
    gets(find);
    while((ch=getc(pt))!=EOF){    //EOF 表示文件结束
        if ((ch!=' ') && (ch!='\n')){
            temp[bian]=ch;
            bian++;
        } else {
            temp[bian]='\0';
            if ((strcmp(find, temp))==0) count++;
            bian=0;
        }
    }
}
```

```

    }
}
printf("共在文件%s中查找到字符串\"%s\" %d 个\n",argv[1],find,count);

getch();
return 0;
}

```

代码分析

上面的程序，需要在 dos 命令行下运行。首先 cd 到程序的编译目录，然后输入程序名和文件名，例如，笔者编译后的程序位于 E:\cDemo\Debug\cDemo.exe，那么运行方式如下：

```

Microsoft Windows [版本 6.1.7601]
版权所有 (c) 2009 Microsoft Corporation。保留所有权利。

C:\Users\mozhiyan>e:

E:\>cd cdemo\debug

E:\cDemo\Debug>cdemo.exe demo.txt
请输入要查找的单词:GitHub
共在文件demo.txt中查找到字符串"GitHub" 2 个

```

这里需要简单介绍一下 main() 函数。main() 函数是程序的入口函数，代码内部无法向其传参，只能在程序运行时由系统在向其传参。一般情况下不需要向main()函数传参，但是如果希望程序启动时可以读取有用数据，进行初始化处理，那么可以这样来传参：

```
int main (int argc, char *argv[])
```

argc 表示参数的个数，argv[] 以数组形式保存传递的所有参数。

注意，当前程序会作为第一个参数传递，例如上面输入 cdemo.exe demo.txt，那么 argc = 2，argc[0]=cdemo.exe，argc[1]=demo.txt

还有一种简单的传参方式，就是将文件（图标）拖到程序上，程序就会立即运行，并将该文件作为一个参数，效果与 cdemo.exe demo.txt 相同。

关于main()函数的更多介绍请查看：[C语言main\(\)函数详解](#)

最后说一下实现单词搜索的思路，这是本程序的关键，这部分代码位于while循环内。

用gets读取用户输入的单词并保存到变量find，while((ch=getc(pt))!=EOF) 对文件进行读取，直到文件结束。

用if对读取的进程进行判断：将每次读取到的单个字符ch依次放入字符串数组temp[]中，直到遇到空格符或者换行符，则这个单词结束，添加' \0' 结束标记。用strcmp()函数与用户搜索的单词find进行比较，如果匹配，则返回0，计数器count加1。最后令bian置0。

注意：英语书写时以空格来分隔每个单词，故我们以空格作为单词结束的标志。而之所以要加上换行符，是因为如果

没有检测换行符，当单词在行首（或行末）时，就会和上一行（或下一行）的最后一个单词连在一起，视为1个单词与find比较，导致统计结果缺失。