

文字信息统计

杨琦

西安交通大学计算机教学实验中心

【例】文字信息统计

用户输入一段文本（英文），统计其字符总个数，大写字母个数、小写字母个数、数字个数及其他字符个数。

【运行结果】

C++ Programming

字符串总长度:15

大写字母:2

小写字母:10

数字个数:0

其他字符:3



【问题分析】



- ▶ ①输入字符串。

cin >> <字符数组名>，只能输入单词；

因为“>>”以空格、tab键、回车为分隔符，遇到空格认为是一项数据的结束。

- ▶ cin是输入流对象，它有一个成员函数getline()可以读取带空格的一行字符串。基本使用格式为：

cin.getline(<字符数组名>, <字符长度>);

例如：char sentence[100]; // 字符数组大小100

cin.getline(sentence, 99);

// 可以输入最长99（回车结束）的字符串，中间可以有空格。

【问题分析】



▶ ②统计计数。

输入的文本看做字符串，统计字符个数，就是从字符串的第1个字符开始，逐个计数，直到结束。

注意，在C++中，字符串的结束是 '`\0`'。统计各类字符个数只要在逐个计数的过程中再对各类字符计数。

计数就是逐个数，用一个变量表示，每遇到一个字符，该变量增加1。

【算法描述】



设字符串用`str[100]`表示，`str[i]`表示第 $i+1$ 个字符（从1开始）。
用`len`表示字符串长度，`capital`表示大写字母个数，`smallletter`表示小写字母个数，`digit`表示数字个数，`others`表示其他字母个数，
初始时它们的值均为0。

- ①输入字符串；
- ② $i=0$ ；
- ③如果`str[i]='\0'`，转⑥； 否则执行④；
- ④ $len=len+1$
 如果`str[i]`为大写字母
 $capital++$ ；

【算法描述】



否则

如果str[i]为小写字母

smallletter++;

否则

如果str[i]为数字

digit++;

否则

others++

⑤i=i+1, 转③;

⑥输出统计数据。

是否为大写、小写或数字，需要比较字符的ASCII值所在的区间。

【源程序】



```
#include<iostream>           //包含输入输出头文件
using namespace std;         //指定名字空间
int main()    {               //主函数
    const int N=101;          //定义常量，表示问题规模
    char str[N];              //定义字符数组，存放字符串，N是常量
    //定义变量并初始化为0
    int len=0, capital=0, smallletter=0, digit=0, others=0;
    int i;                    //循环变量
    cin.getline (str,N);      //输入
    //处理
    i=0;                      //从字符串的第1个字符开始
```



```
while(str[i]!='\0') {           //不是结束符时，循环
    len++;                      //长度加1
    if(str[i]<='Z' && str[i]>='A') { //大写字母
        capital++;
    }
    else if(str[i]<='z' && str[i]>='a') { //小写字母
        smallletter++;
    }
    else if(str[i]<='9' && str[i]>='0') { //数字
        digit++;
    }
    else {                      //其他字符
```




```
        others++;  
    }  
    i++;           //字符下标加1，指向下一个字符  
}  
  
//输出结果  
cout<<"字符串总长度:"<<len<<endl;  
cout<<" 大写字母:"<<capital<<endl;  
cout<<" 小写字母:"<<smallletter<<endl;  
cout<<" 数字个数:"<<digit<<endl;  
cout<<" 其他字符:"<<others<<endl;  
return 0;         //函数返回  
}
```

【编译和链接】



a.cpp

```
1 //example3-7字符统计
2 #include<iostream>           //包含输入输出头文件
3 using namespace std;         //指定名字空间
4 int main()                   //主函数
5 {
6     const int N=101;          //定义常量, 表示问题规模
7     char str[N];              //定义字符数组, 存放字符串, N是常量
8     int len=0, capital=0, smallletter=0, digit=0, others=0; //定义变量并初始化
9     int i;                    //循环变量
10    cin.getline (str,N);       //输入
11    //处理
12    i=0;                       //从字符串的第1个字符开始
13    while(str[i]!='\0') //不是结束符时, 循环
14    {
15        len++;                 //长度加1
16        if(str[i]<='Z' && str[i]>='A') //大写字母
17        {
18            capital++;
19        }
20        else if(str[i]<='z' && str[i]>='a') //小写字母
21        {
22            smallletter++;
23        }
24        digit += (str[i]>='0' && str[i]<='9');
25        others++;
26        i++;
27    }
```

【运行结果】

C++ Programming

字符串总长度:15

大写字母:2

小写字母:10

数字个数:0

其他字符:3



【程序分析】



当<else块>只有一条if语句时，可以这样写，基本格式是：

```
if (<条件1>)
```

```
{ ... }
```

```
else if (<条件2>)
```

```
{ ... }
```

```
else if (<条件3>)
```

```
...
```

```
else
```

```
{...}
```

注意，每一个else是前面最近的if的else,而不是第一个if的else。

【程序分析1】



字符串的操作还有很多，象大小写转换、查找、替换、插入、删除、反转、去掉末尾空格，去掉前导空格等。

不论哪种操作，都是逐个字符去比较、移动个复制，而“是否处理完”的标志就是是否到达末尾的'\0'字符。字符串末尾的'\0'是其结束标记，一定要牢记。

▶ 在VS2008及以上版本中调用strcpy、strcat等函数时由于安全原因会提示警告

```
#define _CRT_SECURE_NO_WARNINGS
```

【程序分析2】



C++提供了一些字符串处理的库函数，方便字符串操作。
常见的有：

`int strlen(char *s);` //求字符串s的长度

//将字符串source复制到destin中

`char *strcpy(char *destin, char *source);`

`int strcmp(char *string1, char *string2);` //比较string1和string2

//将source连接到destin末尾

`char *strcat(char *destin, char *source);`

`char *strlwr(char *string);` //string转换为小写

`char *strupr(char *string);` //stringz转换为大写