

文件读写模式

刘宇辉 wolpythonlondon@gmail.com

November 4, 2011

文件分为二进制文件和文本文件，文件的读写模式分为二进制模式和文本模式。文本文件的表示基本上就是字符在计算机系统内部的表示。比如：7896，在文本文件中，就是 7,8,9,6 对应的 ansii 码，而在二进制文件中，则是 7896 这个数字的二进制表示，可以是 0001111011011000，也可能是 00000000000000000001111011011000，具体的要看数据类型。

而二进制模式和文本模式本质上和二进制和文本文件没有关系。也就是说，以二进制模式打开读写的文件不一定是二进制文件，同理以文本模式打开读写的文件也不一定是文本文件。二进制模式指的是，内存里面是什么，就输出什么，比如内存里是 2345，那么输出的时候，就输出 2345。文本模式在 Windows 上是，写操作时，对内存里的 LF(\n)，进行替换，替换为 CR LF 两个字符，那么实际在磁盘上存储的文件就多了一个字符，读操作时，就是就是逆过程。这个过程对于应用程序是看不到的，是 Windows 内核来做的。在 Unix 中，文本模式和二进制模式没有什么区别，内核不会查看数据进行替换。

在 Windows 上的编辑器，比如 notepad++，可以选择“显示所有字符”，这样就可以看到磁盘的文本文件里“原生态”的字符，而不是被 Windows 内核处理过的文件。原理也就是在打开文本文件的时候，选择二进制模式。

我使用跨平台的轻型 IDE Geany 来演示。

```
1 #include <iostream>
2 #include <fstream>
3
4 using namespace std;
5
6 int main()
7 {
8     fstream out;
9     fstream out1;
```

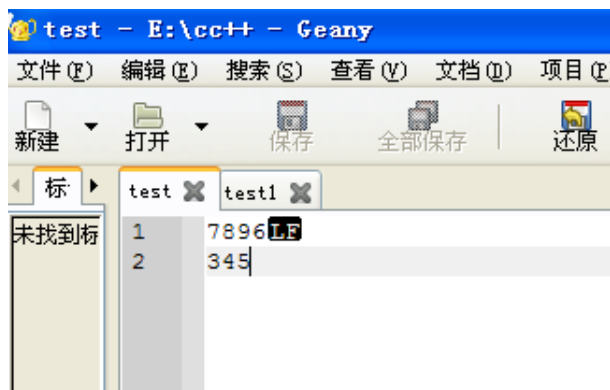


Figure 1: test(binary mode) in Windows

```
10      int a = 7896;
11      int b = 345;
12
13      out.open("test", fstream::binary | fstream::out);
14      out1.open("test1", fstream::out);
15
16      out << a << endl << b;
17      out1 << a << endl << b;
18  }
```

Listing 1: test.cpp

我们向文本文件 `test` 和 `test1` 里面分别用文本模式和二进制模式写入了两个数 7896,345. 然后, 分别用编辑器打开这两个文本文件, 在显示时, 选择显示行尾, 这样我们就能够看到在行末的字符(硬盘上存储的文件, 而不是经过内核处理过存在于内存中的文件)。

我们可以清楚的看到在实际硬盘存储中, Windows 上经过 `text mode` 写入的文件, 在行末都会有一个 `CF` 字符, 这个字符是不可见字符(那为什么我们又看到了呢? 这是一种表示方式而已:-)), 而 `binary mode` 和 Linux 上的两种模式写入的文件都没有多余的字符。我们再来看一下文件的大小, 从而比较一下: 在实际的磁盘上, Windows 上经 `text mode` 写入的文件 `test1`, 比经 `binary mode` 写入的文件 `test` 大了一个字节, 这个字节就是 `CF` 字符。而 Linux 上则是相同的。

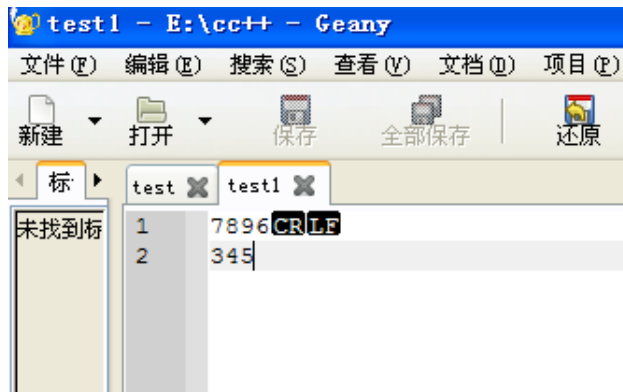


Figure 2: test1(text mode) in Windows

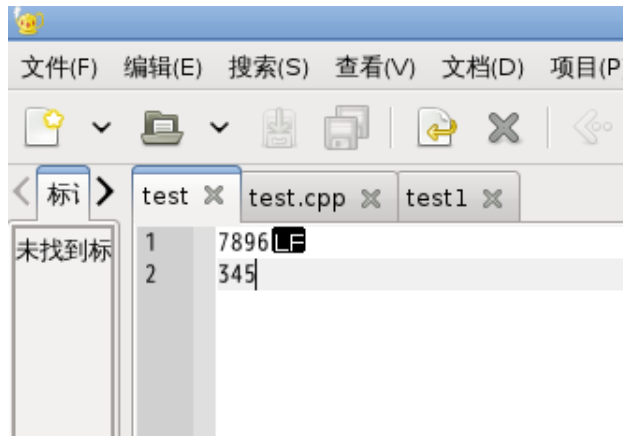


Figure 3: test(binary mode) in Debian

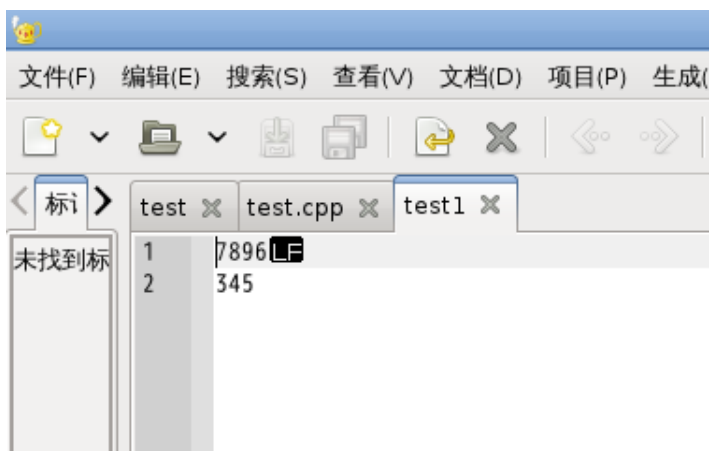


Figure 4: test1(text mode) in Debian

```
Administrator@d0bacdd675b8464 /e/cc++
$ ls -l test test1
-rw-r--r-- 1 Administrator Administrators 8 Nov  3 18:38 test
-rw-r--r-- 1 Administrator Administrators 9 Nov  3 18:38 test1
```

Figure 5: Windows 上的文件大小

```
wolf@debian:/tmp$ ls -l test test1
-rw-r--r-- 1 wolf wolf 8 11月  3 21:33 test
-rw-r--r-- 1 wolf wolf 8 11月  3 21:33 test1
```

Figure 6: Debian 上的文件大小