

Little Tools

kevinluo

Contents

1	iconv 编码转换	1
1.1	参考链接	1
1.2	用法	1
1.3	问题	1
1.3.1	“iconv: illegal input sequence at position” 的错误	1

contents

1 iconv 编码转换

1.1 参考链接

iconv: illegal input sequence at position

1.2 用法

使用 iconv 命令

```
iconv -f GBK -t UTF-8 file1 -o file2
```

```
# iconv -f GBK -t UTF-8 $$@.tmp >$$@
# 加入-c, 表示忽略那些不能解释的字符
# iconv -f GBK -t UTF-8 -c $$@.tmp >$$@
```

1.3 问题

1.3.1 “iconv: illegal input sequence at position” 的错误

iconv: illegal input sequence at position

在使用 iconv 转换文件的字符编码时，如果遇到类似 “iconv: illegal input sequence at position” 的错误，原因是需要转换的字符编码没有涵盖文件中的字符，比如，将一个简体中文的 GB2312 的文件转换为 BIG5 的编码，而在繁体编码的 BIG5 里面，不包含很多的简体中文字符，所以在转换的时候就会遇到如上的错误。

顺便提供一个用于查看文件编码的工具 “enca”，我在 everest 0.5 下做的 RPM 包。用法很简单，：

```
# enca filename
```

解决方法，忽略那些不能解释的字符：

```
iconv -f cp936 -t utf-8 -c file1 > file2
```