

处理

此包将所有文本字符串显示为Python unicode对象。从Excel 97开始，Excel电子表格中的文本已存储为**UTF-16LE**（16位Unicode转换格式）。较旧的文件（Excel 95及更早版本）不保留Unicode中的字符串；一个**CODEPAGE**记录提供了一个代码页号（例如，1252），其被使用的xlrd导出的编码（对于同样的例子：“CP1252”），其用于转换为Unicode。

如果**CODEPAGE**记录丢失（如果文件是由第三方软件创建的，则可能），**xlrd**将假设编码为ascii，并继续。如果实际编码不是ascii，**UnicodeDecodeError**则会引发异常，您需要自己确定编码，并告诉xlrd：

```
book = xlrd.open_workbook(..., encoding_override="cp1252")
```

如果**CODEPAGE**记录存在但是错误（例如，代码页号是1251，但字符串实际上是用koi8_r编码的），则可以使用相同的机制覆盖它。

提供的**runxlrd.py**具有相应的命令行参数，可用于实验：

```
runxlrd.py -e koi8_r 3rows myfile.xls
```

寻找编码的第一个地方，即“编解码器名称”，是 [Python文档](#)。