

Python 中文编码

前面章节中我们已经学会了如何用 Python 输出 "Hello, World!", 英文没有问题, 但是如果你输出中文字符"你好, 世界"就有可能碰到中文编码问题。

Python 文件中如果未指定编码, 在执行过程会出现报错:

```
#!/usr/bin/python
print "你好, 世界";
```

以上程序执行输出结果为:

```
File "test.py", line 2
SyntaxError: Non-ASCII character '\xe4' in file test.py on line 2, but no encoding declared; see http://
www.python.org/peps/pep-0263.html for details
```

Python中默认的编码格式是 ASCII 格式, 在没修改编码格式时无法正确打印汉字, 所以在读取中文时会报错。

解决方法为只要在文件开头加入 `# -*- coding: UTF-8 -*-` 或者 `#coding=utf-8` 就行了

注意: `#coding=utf-8` 的 `=` 号两边不要空格。

实例(Python 2.0+)

```
#!/usr/bin/python
# -*- coding: UTF-8 -*-
print "你好, 世界";
```

运行实例 »

输出结果为:

```
你好, 世界
```

所以如果大家在学习过程中, 代码中包含中文, 就需要在头部指定编码。

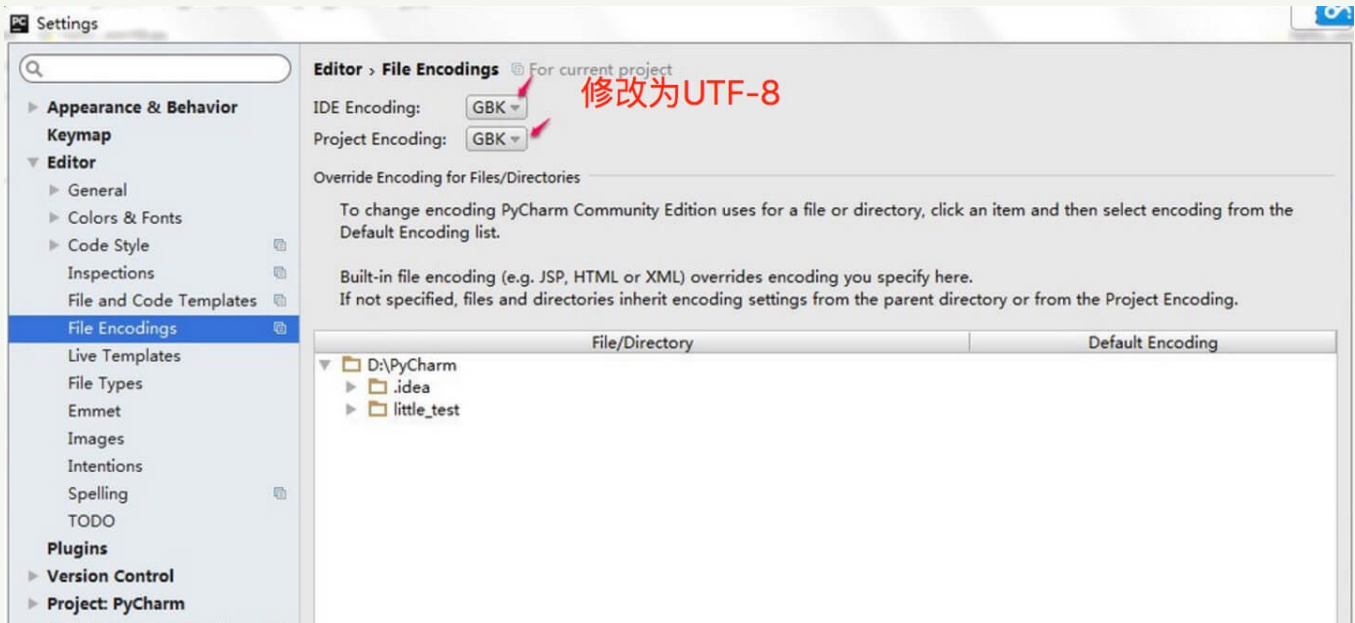
注意: Python3.X 源码文件默认使用utf-8编码, 所以可以正常解析中文, 无需指定 UTF-8 编码。

注意: 如果你使用编辑器, 同时需要设置 py 文件存储的格式为 UTF-8, 否则会出现类似以下错误信息:

```
SyntaxError: (unicode error) 'utf-8' codec can't decode byte 0xc4 in position 0:
invalid continuation byte
```

Pycharm 设置步骤：

- 进入 **file > Settings**，在输入框搜索 **encoding**。
- 找到 **Editor > File encodings**，将 **IDE Encoding** 和 **Project Encoding** 设置为 **utf-8**。

[← Python JSON](#)[Python partition\(\) 方法 →](#)**3 篇笔记** **写笔记**