

小楼

读书 学习 生活 进步

[首页](#)[新随笔](#)[联系](#)[订阅](#)[管理](#)

公告

昵称：小楼

园龄：8年3个月

粉丝：61

关注：24

<	2017年12月						>
日	一	二	三	四	五	六	
26	27	28	29	30	1	2	
3	4	5	6	7	8	9	
10	11	12	13	14	15	16	
17	18	19	20	21	22	23	
24	25	26	27	28	29	30	
31	1	2	3	4	5	6	

搜索

找找看

谷歌搜索

最新随笔

1. 《结网：改变世界的互联网产品经理》
2. <[你在荒废时间的时候别人都在拼命！]>
3. 《时间的玫瑰》阅读笔记
4. 翻石头价值投资手册-科技行业

随笔-300 文章-0 评论-45

UnicodeEncodeError: 'ascii' codec can't encode characters in position 0-1: ordinal not in range(128)

在使用selenium处理中文网页或者网页标题是中文的时候，出现UnicodeEncodeError: 'ascii' codec can't encode characters in position 0-1: ordinal not in range(128)，



```
1 from selenium import webdriver
2 import sys
3
4 print sys.getdefaultencoding()
5
6 driver = webdriver.PhantomJS()
7 driver.get("http://www.douban.com")
8
9 print driver.title
```



字符串在Python内部的表示是unicode编码，因此，在做编码转换时，通常需要以unicode作为中间编码，即将其他编码的字符串解码（decode）成unicode，再从unicode编码（encode）成另一种编码。

decode的作用是将其他编码的字符串转换成unicode编码，如str1.decode('gb2312')，表示将gb2312编码的字符串str1转换成unicode编码。

encode的作用是将unicode编码转换成其他编码的字符串，如str2.encode('gb2312')，表示将unicode编码的

- 5. No module named flask.ext.sqlalchemy.SQLALchemy
- 6. 《寻找伟大的企业》
- 7. <《基金经理投资笔记丛书4-1：投资是一种生活方式》>
- 8. 超级有用的三个原则
- 9. <欧奈尔制胜法则—如何在股市中赚钱>读书笔记
- 10. <《巴菲特之道 (第三版)》>读书笔记

我的标签

读书笔记(129)
Python(84)
Django(18)
Selenium(17)
Ruby(10)
Linux(9)
Java(8)
AWK(8)
比特币(5)
书籍下载(4)
更多

随笔档案(300)

2016年2月 (5)
2016年1月 (3)
2015年12月 (2)
2015年11月 (1)
2015年10月 (4)
2015年7月 (1)

字符串str2转换成gb2312编码。

因此，转码的时候一定要先搞明白，字符串str是什么编码，然后decode成unicode，然后再encode成其他编码

代码中字符串的默认编码与代码文件本身的编码一致。

douban网站的标题编码是gb2312，
所以使用decode就可以解决这个问题

```
print driver.title.encode("gb2312")
```

为什么会报错“UnicodeEncodeError: 'ascii' codec can't encode characters in position 0-1: ordinal not in range(128)”？本文就来研究一下这个问题。

字符串在Python内部的表示是unicode编码，因此，在做编码转换时，通常需要以unicode作为中间编码，即先将其他编码的字符串解码（decode）成unicode，再从unicode编码（encode）成另一种编码。

decode的作用是将其他编码的字符串转换成unicode编码，如str1.decode('gb2312')，表示将gb2312编码的字符串str1转换成unicode编码。

encode的作用是将unicode编码转换成其他编码的字符串，如str2.encode('gb2312')，表示将unicode编码的字符串str2转换成gb2312编码。

因此，转码的时候一定要先搞明白，字符串str是什么编码，然后decode成unicode，然后再encode成其他编码

代码中字符串的默认编码与代码文件本身的编码一致。

如：s='中文'

- 2015年5月 (6)
- 2015年4月 (4)
- 2015年3月 (3)
- 2015年2月 (6)
- 2015年1月 (8)
- 2014年11月 (4)
- 2014年8月 (2)
- 2014年7月 (1)
- 2014年6月 (1)
- 2014年5月 (2)
- 2014年4月 (4)
- 2014年3月 (7)
- 2014年2月 (6)
- 2014年1月 (5)
- 2013年12月 (1)
- 2013年11月 (1)
- 2013年10月 (1)
- 2013年6月 (8)
- 2013年5月 (2)
- 2013年4月 (5)
- 2013年3月 (9)
- 2013年2月 (2)
- 2013年1月 (4)
- 2012年12月 (2)
- 2012年11月 (9)
- 2012年10月 (2)
- 2012年9月 (7)
- 2012年8月 (10)
- 2012年7月 (8)

如果是在utf8的文件中，该字符串就是utf8编码，如果是在gb2312的文件中，则其编码为gb2312。这种情况下，要进行编码转换，都需要先用decode方法将其转换成unicode编码，再使用encode方法将其转换成其他编码。通常，在没有指定特定的编码方式时，都是使用的系统默认编码创建的代码文件。

标签: Python


好文要顶

关注我

收藏该文







小楼

关注 - 24

粉丝 - 61

+加关注

2

0

« 上一篇：[\[驱动力\]读书笔记](#)

» 下一篇：[反思之一](#)

posted @ 2014-03-11 16:59 小楼 阅读(57491) 评论(1) 编辑 收藏

评论列表

#1楼 2016-06-06 13:49 Lumia1020

mark 一下

支持(0) 反对(0)

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请[登录](#)或[注册](#)，[访问网站首页](#)。

- 【推荐】50万行VC++源码: 大型组态工控、电力仿真CAD与GIS源码库
- 【促销】腾讯云技术升级10大核心产品年终让利
- 【推荐】高性能云服务器2折起，0.73元/日节省80%运维成本
- 【新闻】H3 BPM体验平台全面上线

2012年6月 (8)
2012年5月 (25)
2012年4月 (12)
2012年3月 (13)
2012年2月 (3)
2012年1月 (1)
2011年7月 (1)
2010年11月 (2)
2010年10月 (2)
2010年9月 (1)
2010年6月 (4)
2010年5月 (12)
2010年4月 (2)
2010年3月 (6)
2010年1月 (10)
2009年12月 (13)
2009年11月 (6)
2009年10月 (10)
2009年9月 (18)
2009年8月 (5)

相册(26)

articles(2)

国家地理(24)

个人

Python文档

积分与排名

积分 - 177997



最新IT新闻:

- 人人网二手车的收入超过了公司收入的三分之二，人人车：？？？
 - 被吴恩达的Landing.ai刷屏后，我们看到了三个有趣细节
 - 挪威成为第一个关闭FM广播的国家
 - 腾讯一年一度的产品大奖，得奖的都是谁？
 - 一场事先张扬的失败之后，映客该何去何从
- » 更多新闻...



最新知识库文章:

- 以操作系统的角度述说线程与进程
 - 软件测试转型之路
 - 门内门外看招聘
 - 大道至简，职场上做人做事做管理
 - 关于编程，你的练习是不是有效的？
- » 更多知识库文章...

排名 - 1329

最新评论

1. Re:No module named django.core

不是django-admin.py startproject mysite
而是django-admin startproject mysite那个目录下有两个django-admin 一个是py.....

--warlock.lv

2. Re:解决ImportError: cannot import name webdriver

将from selenium import webdriver改为import selenium，之后调用函数采用selenium.函数名的方式可以解决这个问题，不需要删除

--yinzm

3. Re:svn错误：更新源码出现校验和不匹配问题

1.进入出现问题的文件所在的目录2.执行命令svn update --set-depth empty3.执行命令svn update --set-depth infinityps：如果有文件改动，先把.....

--envoy

4. Re:解决ImportError: cannot import name webdriver

删除了还是会出错，亲

--xunmengzhi

5. Re:UnicodeEncodeError: 'ascii' codec can't encode characters in position 0-1: ordinal not in range(128)

mark 一下

--Lumia1020

阅读排行榜

1. UnicodeEncodeError: 'ascii' codec can't encode characters in position 0-1: ordinal not in range(128)(57491)
2. Python之lxml库学习笔记一(41327)
3. 解决ImportError: cannot import name webdriver(39308)
4. Python模块之 __future__(26227)
5. Python的自增运算符(15890)

评论排行榜

1. [脚本收集]提取国家地理图片(4)
2. 看不完的风景，走不完的路(3)
3. Django: AttributeError: 'str' object has no attribute 'resolve'(3)
4. Python之压缩文件(3)
5. 提取国家地理图片总结(2)

推荐排行榜

1. Python之lxml库学习笔记一(2)
2. [脚本收集]提取国家地理图片(2)
3. Selenium不支持firefox3的解决办法(2)
4. Python获得操作系统信息(2)
5. ImportError: No module named urllib2(2)