

Vamei

编程, 数学, 设计

博客园

首页

订阅

管理

随笔-209 文章-1 评论-3802

Python标准库05 存储对象 (pickle包, cPickle包)

作者: Vamei 出处: <http://www.cnblogs.com/vamei> 欢迎转载, 也请保留这段声明。谢谢!

谢谢reverland纠错

在之前对Python对象的介绍中 (面向对象的基本概念, 面向对象的进一步拓展), 我提到过Python“一切皆对象”的哲学, 在Python中, 无论是变量还是函数, 都是一个对象。当Python运行时, 对象存储在内存中, 随时等待系统的调用。然而, 内存里的数据会随着计算机关机和消失, 如何将对象保存到文件, 并储存在硬盘上呢?

计算机的内存中存储的是二进制的序列 (当然, 在Linux眼中, 是文本流)。我们可以直接将某个对象所对应位置的数据抓取下来, 转换成文本流 (这个过程叫做serialize), 然后将文本流存入到文件中。由于Python在创建对象时, 要参考对象的类定义, 所以当我们从文本中读取对象时, 必须在手边要有该对象的类定义, 才能懂得如何去重建这一对象。从文件读取时, 对于Python的内建(built-in)对象 (比如说整数、词典、表等等), 由于其类定义已经载入内存, 所以不需要我们再在程序中定义类。但对于用户自行定义的对象, 就必须要先定义类, 然后才能从文件中载入对象 (比如面向对象的基本概念中的对象那个summer)。

pickle包

对于上述过程, 最常用的工具是Python中的pickle包。

1) 将内存中的对象转换成为文本流:



```
import pickle

# define class
```

```
class Bird(object):
    have_feather = True
    way_of_reproduction = 'egg'

summer = Bird() # construct an object
picklestring = pickle.dumps(summer) # serialize object
```



使用`pickle.dumps()`方法可以将对象`summer`转换成了字符串

`picklestring` (也就是文本流)。随后我们可以用普通文本的存储方法来将该字符串储存在文件 (文本文件的输入输出)。

当然, 我们也可以使用`pickle.dump()`的方法, 将上面两部合二为一:

```
import pickle

# define class
class Bird(object):
    have_feather = True
    way_of_reproduction = 'egg'

summer = Bird() # construct an object
fn = 'a.pkl'
with open(fn, 'w') as f: # open file
    with write-mode
        picklestring = pickle.dump(summer, f) # serialize and save object
```



对象`summer`存储在文件`a.pkl`

2) 重建对象

首先, 我们要从文本中读出文本, 存储到字符串 (文本文件的输入输出)。然后使用`pickle.loads(str)`的方法, 将字符串转换成为对象。要记得, 此时我们的程序中

必须已经有了该对象的类定义。

此外，我们也可以使用`pickle.load()`的方法，将上面步骤合并：



```
import pickle

# define the class before unpickle
class Bird(object):
    have_feather = True
    way_of_reproduction = 'egg'

fn = 'a.pkl'
with open(fn, 'r') as f:
    summer = pickle.load(f)    # read file and build object
```



cPickle包

cPickle包的功能和用法与pickle包几乎完全相同（其存在差别的地方实际上很少用到），不同在于cPickle是基于c语言编写的，速度是pickle包的1000倍。对于上面的例子，如果想使用cPickle包，我们都可以将import语句改为：

```
import cPickle as pickle
```

就不需要再做任何改动了。

总结

对象 -> 文本 -> 文件

`pickle.dump()`, `pickle.load()`, `cPickle`

标签: Python

好文要顶

关注我

收藏该文



Vamei

关注 - 26

粉丝 - 4985

荣誉: 推荐博客

+加关注

3

0

(请您对文章做出评价)

« 上一篇: Python标准库04 文件管理 (部分os包, shutil包)

» 下一篇: 把孩子打造成为码农

posted @ 2012-09-15 11:52 Vamei 阅读(17653) 评论(29) 编辑 收藏

评论列表

#1楼 2012-09-15 14:12 EDM数据

Python标准库不需要深究吧。

支持(0) 反对(0)

#2楼[楼主] 2012-09-15 15:26 Vamei

@ EDM数据

为什么这么说呢?

支持(0) 反对(0)

#3楼 2012-10-07 22:56 ET民工[源自火星]

"当然,我们也可以使用pickle.dump()的方法,将上面两部合二为一:",运行这部分代码时发现两个小错误,

```
with open(fn, 'w') as f: -> with open(fn, 'wb') as f:
```

```
picklestring = pickle.dump(summer, fn) -> picklestring =  
pickle.dump(summer, f)
```

能够运行的代码如下:

```
1 | import pickle  
2 |
```

```

3  # define class
4  class Bird(object):
5      have_feather = True
6      way_of_reproduction = 'egg'
7
8  summer = Bird()                # construct an object
9  fn = 'a.pkl'
10 with open(fn, 'wb') as f:      # open file with write-mode
11     picklestring = pickle.dump(summer, f) # serialize and save object

```

支持(0) 反对(0)

#4楼[楼主] 2012-10-08 05:06 Vamei

@ ET民工[源自火星]

可能是平台的问题。由于UNIX系统不区分wb和w，所以我总写成w。而在windows中，则要写成wb。我在写的时候是用LINUX测试的。

支持(0) 反对(0)

#5楼 2012-10-08 10:54 ET民工[源自火星]

@ Vamei

难怪了，signal一节里面的代码，在windows下也跑不过，原因应该相同吧。

支持(0) 反对(0)

#6楼 2012-11-18 10:51 reverland

不是dump(summer,f)么，原文怎么没改.....

支持(0) 反对(0)

#7楼 2012-11-18 11:04 reverland

每次load的时候都是

```

1  >>> with open('a.pkl','rb') as f:
2      ...     summer = pickle.load(f)
3      ...
4  Traceback (most recent call last):
5      File "<stdin>", line 2, in <module>
6  EOFError

```

支持(0) 反对(0)

#8楼[楼主] 2012-11-18 11:37 Vamei

@ reverland

刚刚改过

支持(0) 反对(0)

#9楼[楼主] 2012-11-18 11:39 Vamei

@ reverland

用我的电脑没有发现你说的情况，可以更具体一些么？你可以成功dump吧？

支持(0) 反对(0)

#10楼 2012-11-18 11:56 reverland

@ Vamei

dump应该没问题吧.....您看看

```
1 lyy@arch ~/Work/py % cat a.pk1
2 ccopy_reg
3 _reconstructor
4 p0
5 (c__main__
6 Bird
7 p1
8 c__builtin__
9 object
10 p2
11 Ntp3
12 Rp4
13 .%
```

按这里来则完全没问题.....

http://woodpecker.org.cn/abyteofpython_cn/chinese/ch12s02.html

支持(0) 反对(0)

#11楼[楼主] 2012-11-18 17:32 Vamei

@ reverland

你是什么操作系统？

支持(0) 反对(0)

#12楼 2012-11-18 18:28 reverland

@ Vamei

linux

python2.7.3和3都这样

支持(0) 反对(0)

#13楼[楼主] 2012-11-18 18:36 Vamei

@ reverland

很奇怪。因为我使用你上面的pickle文件，没有出现EOFError的问题。你先使用那个教程中的方法吧。我们再慢慢发现问题。

支持(0) 反对(0)

#14楼 2012-11-18 19:27 reverland

@ Vamei

刚才看看a.pkl发现竟然是空的.....

然后重新dump/load一点问题没有.....我又cat下还是那样，不知道刚开始干什么了。

可能有个文件句柄忘记close了吧.....

支持(1) 反对(0)

#15楼[楼主] 2012-11-18 20:16 Vamei

@ reverland

:-)，找到原因就好。

支持(0) 反对(0)

#16楼 2013-11-06 23:50 szxiaosong

Python 2.7.6rc1，运行例子没有问题。

支持(0) 反对(0)

#17楼 2013-11-06 23:51 szxiaosong

linux下Python安装方法

1. # wget <http://www.python.org/ftp/python/2.7.6/Python-2.7.6rc1.tgz>
2. # cd Pyth Python linux 安装 on-2.7.6rc1
3. # ./configure
4. # make
5. # make install

支持(0) 反对(0)

#18楼 2013-11-06 23:52 szxiaosong

1. # wget <http://www.python.org/ftp/python/2.7.6/Python-2.7.6rc1.tgz>
2. # cd Python-2.7.6rc1
3. # ./configure
4. # make
5. # make install

支持(0) 反对(0)

#19楼 2013-12-20 10:28 rhythmmer

python3.x中得使用with open(fn,'wb') as f:，而不是with open(fn,'w') as f:

支持(0) 反对(0)

#20楼 2014-06-15 21:21 执虫

@ reverland

确实需要close

支持(0) 反对(0)

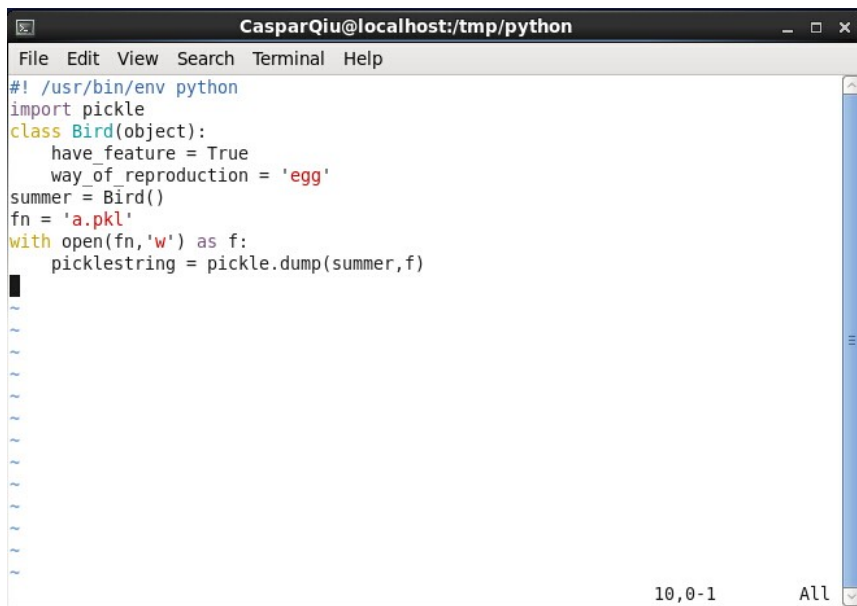
#21楼 2014-08-24 07:40 皓皓^_^

vamei大神，小弟序列化的概念不是太懂，能否指点一下，主要是不太理解为什么要序列化，序列化能解决什么问题，有什么好处，小弟是搞c/c++的，小菜一个，之前就在网络那见过一点序列化，也不太理解具体原因。如果我想把一个结构体保存到文件里，完全可以写进文件里啊，然后再把相应的数据读出来。能给出点没有序列化就实现不了固化数据的例子呢？

支持(0) 反对(0)

#22楼 2014-08-26 14:19 CasparQiu

```
[root@localhost python]# ./pickle.py
Traceback (most recent call last):
  File "./pickle.py", line 2, in <module>
    import pickle
  File "/tmp/python/pickle.py", line 9, in <module>
    picklestring = pickle.dump(summer,f)
AttributeError: 'module' object has no attribute 'dump'
[root@localhost python]#
```

为什么我每次import标准库的时候都会报错，望解答，谢谢。

支持(0) 反对(0)

#23楼[楼主] 2014-08-26 16:03 Vamei

@ CasparQiu

你是什么版本的python?

支持(0) 反对(0)

#24楼[楼主] 2014-08-27 16:15 Vamei

@ 皓皓^_^

Python/Java里面，对象必须经过序列化，才能存储。

C/C++虽然可以存储解构体。但如果读和写不在同一台电脑上，读出来的结果就可能有问题了。

支持(0) 反对(0)

#25楼 2014-08-28 06:35 皓皓^_^

@ Vamei

啊 明白了，就是存在大端小端的问题，谢谢楼主啊

支持(0) 反对(0)

#26楼 2014-08-29 14:56 CasparQiu

@ Vamei

已经解决了，原来是我文件命名的错误，文件名直接命名为pickle.py了

支持(1) 反对(0)

#27楼 2015-05-06 09:57 r3call

我在使用中发现了一点 cPickle 和 pickle 的不同.

即 pickle 在 pickle.loads 的时候,输入可以是 unicode

但是 cPickle 的loads 不行

我是在django中使用的,字符串存储的时候被标识为 unicode.

支持(0) 反对(0)

#28楼 2015-06-10 15:49 pyers

@ CasparQiu

引用

@Vamei

已经解决了, 原来是我文件命名的错误, 文件名直接命名为pickle.py了

嘿嘿, 自己练习的程序名写成pickle.py了。Tks!

支持(1) 反对(0)

#29楼 2015-12-02 21:56 龙漪寒

为什么我安装了cPickle后import cPickle说找不到这个包

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

【推荐】融云即时通讯云一豆果美食、Faceu等亿级APP都在用



GrapeCity.

SpreadJS 表格控件

Excel界面数据处理
可视化、可定制

[了解详情](#)

JPush 极光推送

消息推送领导品牌全面升级

JIGUANG 极光

详情点击

公告

你好，这里是Vamei，一名编程爱好者。我在博客里写了**Python/Linux/网络协议/算法/Java/数据科学**系列文章，从这里开始阅读。非常期待和你的交流。



Vamei 🐦 : 我科研发出#功能性机器人#佳佳，据说以本校5位校花为原型。双双[图]

4月15日 22:33 | 微博



weibo.com/vamei

我的微博

下列教程已经做成电子出版物，内容经过修订，也方便离线阅读：

协议森林

欢迎阅读我写的其他书籍：

现代小城的考古学家

天气与历史的相爱相杀

随手拍光影

昵称：Vamei

园龄：4年1个月

荣誉：推荐博客

粉丝：4985

关注：26

+加关注

常用链接

我的随笔

我的评论

我的参与

最新评论

我的标签

我的标签

Python(61)

Java(42)

[大数据\(22\)](#)

[Linux\(17\)](#)

[网络\(16\)](#)

[算法\(15\)](#)

[文青\(14\)](#)

[技普\(9\)](#)

[系列索引\(6\)](#)

[开发工具\(4\)](#)

[更多](#)

[系列文章](#)

[Java快速教程](#)

[Linux的概念与体系](#)

[Python快速教程](#)

[数据科学](#)

[协议森林](#)

[纸上谈兵：算法与数据结构](#)

[积分与排名](#)

[积分 - 659668](#)

[排名 - 122](#)

[最新评论](#)

[1. Re:Java基础11 对象引用](#)

[受教！](#)

--MissLost

[2. Re:Python快速教程](#)

[看评论区一片喝彩！看来我得在此扎营了！](#)

--测试小蚂蚁

[3. Re:Python进阶06 循环对象](#)

[好好地列表解析变成了表推导](#)

--ashic

[4. Re:“不给力啊，老湿！”：RSA加密与破解](#)

[感谢楼主精彩分享](#)

--worldball

[5. Re:概率论04 随机变量](#)

你写的这一系列太棒了，刚加入博客园就在你这里学到了，我要转载过去学习一下

--yixius

6. Re:Python基础03 序列

挺好的教程、、、

--王小拽的号

7. Re:Python进阶07 函数对象

```
def func(x,y): print x**ydef test(f,a,b): print 'test' print f(a,b)test (func,3,2)
```

输出的内容:tes.....

--M-edea

8. Re:Python进阶02 文本文件的输入输出

@coderXT换行符: \n...

--行者之印

9. Re:数据科学

博主啊，这里是一枚即将大二的计算机新人，大一学了python，java，还有一些算法，数据结构，图论了，感觉我对数学又一些反感，但是听说离散数学对计算机专业的很重要，不知道怎么去学比较好呢，我想像您写.....

--Acokil

10. Re:为什么要写技术博

楼主是用自己自定义的模板吗？在博客园里找不到这种风格的blog模板？

--行者之印

11. Re:来玩Play框架01 简介

挖煤哥,我补充了一下Windows下的搭建play框架,希望有点帮助,谢谢!

--Sungeek

12. Re:来玩Play框架07 静态文件

```
@helper.form(action = routes.Application.upload, 'enctype ->
"multipart/form-data") {--action = rout.....
```

--quxiaozha

13. Re:来玩Play框架07 静态文件

该记录将/assets/下的URL，对应到项目的/public文件夹内的文件。比如在项目的/public/images/test.jpg，就可以通过/assests/images/test.jpg这一.....

--quxiaozha

14. Re:来玩Play框架06 用户验证

支持挖煤哥~~~

--quxiaozha

15. Re:“不给力啊，老湿！”：RSA加密与破解

@maanshancss请你仔细阅读了这个文章再来评价。...

--Vamei

推荐排行榜

1. “不给力啊，老湿！”：RSA加密与破解(218)
2. Python快速教程(140)
3. 野蛮生长又五年(91)
4. Java快速教程(88)
5. 协议森林01 邮差与邮局 (网络协议概观)(79)
6. 为什么要写技术博(71)
7. 编程异闻录(54)
8. 博客一年：心理之旅(49)
9. 协议森林08 不放弃 (TCP协议与流通信)(45)
10. Python快速教程 尾声(43)
11. 协议森林(42)
12. Java基础01 从HelloWorld到面向对象(42)
13. Python基础08 面向对象的基本概念(40)
14. 一天能学会的计算机技术(34)
15. 博客第二年，杂谈(33)

Copyright ©2016 Vamei

05370199