

Python for Senior Lesson 4

v1.0.0

2016.10 by David.Yi

复习

- python 常用内置函数
- List 列表的10个方法
- List 列表的复制
- 写一个简单的图形程序

本次内容要点

- 文件和目录操作之一
- 有趣的 python 命令
- 思考一下

文件和目录操作之一

Python 的 os 库：有很多和操作系统相关的功能，还有很多和文件、路径和执行系统命令相关的函数。

os 库常用函数

- os.sep 可以取代操作系统特定的路径分割符
- os.name 字符串指示你正在使用的平台。比如对于Windows，它是'nt'，而对于Linux/Unix用户，它是'posix'
- os.getcwd() 函数得到当前工作目录，即当前Python脚本工作的目录路径
- os.chdir(dirname) 改变工作目录到dirname
- os.getenv() 用来读取环境变量
- os.putenv() 用来设置环境变量
- os.listdir() 返回指定目录下的所有文件和目录名
- os.remove() 删除一个文件
- os.system() 运行shell命令
- os.linesep 字符串给出当前平台使用的行终止符。例如，Windows使用'\r\n'，Mac使用'\n'。
- os.mkdir() 建立路径
- os.rmdir() 删除路径

注意：我们在 /Users/用户名 路径下建立一个用来测试的文件 test.txt

不同操作系统在路径和文件处理上有一定差异，这里的举例是基于 macOS，在 windows 下面会有较大差异，但是用法是一致的

关于文件系统的延伸阅读

- 文件系统介绍
<https://zh.wikipedia.org/wiki/%E6%96%87%E4%BB%B6%E7%B3%BB%E7%BB%9F>
(<https://zh.wikipedia.org/wiki/%E6%96%87%E4%BB%B6%E7%B3%BB%E7%BB%9F>)
- windows 文件系统 FAT、FAT32、NTFS 介绍 <https://support.microsoft.com/zh-cn/kb/100108>
(<https://support.microsoft.com/zh-cn/kb/100108>)
- linux 文件系统介绍 http://cn.linux.vbird.org/linux_basic/0230filesystem.php
(http://cn.linux.vbird.org/linux_basic/0230filesystem.php)

In [2]:

```
import os

# 操作系统路径分隔符
print(os.sep)

# 操作系统平台名称
print(os.name)
```

```
/
posix
```

In [3]:

```
# 获取当前路径
os.getcwd()
```

Out[3]:

```
'/Users/yijun/Documents/dev_python/python_beginner/python_senior_slides/lesson4'
```

In [4]:

```
# 切换路径
os.chdir('/Users/yijun')
os.getcwd()
```

Out[4]:

```
'/Users/yijun'
```

In [4]:

```
# 列出当前目录下所有文件
os.listdir()
```

Out[4]:

```
['.ipynb_checkpoints',
 'lesson4_python_senior.ipynb',
 'list_generator.ipynb',
 'python_senior_L4_v1.pptx']
```

In [6]:

```
# 注意返回的数据类型
print(type(os.listdir()))
```

```
<class 'list'>
```

In [3]:

```
# 计算目录下有多少文件
a = os.listdir()
print(len(a))
```

```
4
```

In [10]:

```
# 可以指定路径参数，来列出该目录下所有文件
l = os.listdir('/Users/yijun')

# 可以判断各类情况，比如第一个是大写的 P 字母
for i in l:
    if i[0:1] == 'P':
        print(i)
```

```
PAServer
Pictures
Public
PycharmProjects
```

In [6]:

```
# 操作系统换行符
# 在一些文本文件处理中 useful

os.linesep
```

Out[6]:

```
'\n'
```

In [22]:

```
# 建立路径

os.chdir('/Users/yijun')
os.mkdir('test')
```

```
-----
-----
FileExistsError                                Traceback (most recent call
1 last)
<ipython-input-22-91172d87a831> in <module>()
      2
      3 os.chdir('/Users/yijun')
----> 4 os.mkdir('test')

FileExistsError: [Errno 17] File exists: 'test'
```

os.path 常用函数之一

- `os.path.isdir()` 检查给出的路径是否是一个目录
- `os.path.isfile()` 检查给出的路径是否是一个文件
- `os.path.exists()` 检查给出的路径或者文件是否存在
- `os.path.getsize()` 获得路径或者文件的大小
- `os.path.getatime()` 返回所指向的文件或者目录的最后存取时间
- `os.path.getmtime()` 返回所指向的文件或者目录的最后修改时间

In [9]:

```
# 检查给出的路径是否是一个目录，存在
os.path.isdir('/Users/yijun')
```

Out[9]:

```
True
```

In [10]:

```
# 检查给出的路径是否是一个目录，不存在
os.path.isdir('/Users/yijunn')
```

Out[10]:

```
False
```

In [12]:

```
# 文件不是路径, 所以返回 False
os.path.isdir('/Users/yijun/test.txt')
```

Out[12]:

False

In [11]:

```
# 检查给出的路径是否一个文件, 存在
os.path.isfile('/Users/yijun/test1.txt')
```

Out[11]:

True

In [12]:

```
# 检查给出的路径是否一个文件, 不存在
os.path.isfile('/Users/yijun/test.txt')
```

Out[12]:

False

In [14]:

```
# 路径不是文件, 所以返回 False
os.path.isfile('/Users/yijun')
```

Out[14]:

False

In [16]:

```
# 对路径和文件都通用的检查方式
print(os.path.exists('/Users/yijun'))
print(os.path.exists('/Users/yijun/test.txt'))
```

True

True

In [47]:

```
# 获得路径或者文件的大小

os.path.getsize('/Users/yijun/test.txt')
```

Out[47]:

19

In [48]:

```
# 获得路径或者文件的大小
```

```
os.path.getsize('/Users/yijun')
```

Out[48]:

2448

In [26]:

```
# 返回所指向的文件或者目录的最后存取时间
```

```
os.path.getatime('/Users/yijun/test.txt')
```

Out[26]:

1476880061.0

In [3]:

```
# 返回所指向的文件或者目录的最后存取时间
```

```
import os  
import time
```

```
# 将日期格式化
```

```
dt = time.localtime(os.path.getatime('/Users/yijun'))
```

```
# print(dt)
```

```
print(time.strftime('%Y-%m-%d %H:%M:%S', dt))
```

2016-10-22 13:22:46

In [28]:

```
# 返回所指向的文件或者目录的最后修改时间
```

```
os.path.getmtime('/Users/yijun/test.txt')
```

Out[28]:

1476879992.0

In [19]:

```
# 返回所指向的文件或者目录的最后修改时间
```

```
# 使用 time.ctime() 方法来格式化日期
```

```
import time, os
```

```
time.ctime(os.path.getmtime('/Users/yijun/test.txt'))
```

Out[19]:

'Wed Oct 19 20:26:32 2016'

os.path 常用函数之二

- os.path.split() 返回一个路径的目录名和文件名
- os.path.abspath() 返回规范化的绝对路径
- os.path.isabs() 如果输入是绝对路径，返回True
- os.path.split() 将路径分割成目录和文件名的二元素元组
- os.path.splitdrive() 返回 (drivename, fpath) 元组
- os.path.dirname() 返回路径的目录，其实就是 os.path.split(path)的第一个元素
- os.path.basename() 返回路径最后的文件名，其实就是 os.path.split(path)的第二个元素
- os.path.splitext() 分离文件名与扩展名，返回(fname,fextension)元组
- os.path.join() 将多个路径组合后返回，第一个绝对路径之前的参数将被忽略
- os.path.commonprefix(list) 返回list中，所有路径共有的最长的路径

In [29]:

```
# 返回一个路径的目录名和文件名
```

```
os.path.split('/Users/yijun/test.txt')
```

Out[29]:

```
(' /Users/yijun', 'test.txt')
```

In [5]:

```
# 返回规范化的绝对路径  
# 会自动补齐完整路径
```

```
os.path.abspath('tt12.txt')
```

Out[5]:

```
 '/Users/yijun/Documents/dev_python/python_beginner/python_senior/les  
son4/tt12.txt'
```

In [34]:

```
# 如果输入是绝对路径，返回True
```

```
print(os.path.isabs('test.txt'))  
print(os.path.isabs('/Users/yijun/test.txt'))
```

```
False
```

```
True
```

In [37]:

```
# 将路径分割成目录和文件名的二元素元组
```

```
s = os.path.split('/Users/yijun/test.txt')  
print(s)  
print(type(s))
```

```
(' /Users/yijun', 'test.txt')  
<class 'tuple'>
```

In [38]:

```
# 返回 (drivename, fpath) 元组
# mac 下 drivename 返回为空
os.path.splitdrive('/Users/yijun/test.txt')
```

Out[38]:

```
('', '/Users/yijun/test.txt')
```

In [42]:

```
# 返回路径的目录，其实就是 os.path.split(path) 的第一个元素
os.path.dirname('/Users/yijun/test.txt')
```

Out[42]:

```
'/Users/yijun'
```

In [41]:

```
# 返回路径最后的文件名，其实就是 os.path.split(path) 的第二个元素
os.path.basename('/Users/yijun/test.txt')
```

Out[41]:

```
'test.txt'
```

In [44]:

```
# 分离文件名与扩展名，返回 (fname, fextension) 元组
os.path.splitext('/Users/yijun/test.txt')
```

Out[44]:

```
('/Users/yijun/test', '.txt')
```

In [50]:

```
# 将多个路径组合后返回，第一个绝对路径之前的参数将被忽略
os.path.join('/Users/yijun', 'test.txt')
```

Out[50]:

```
'/Users/yijun/test.txt'
```

In [46]:

```
# 返回list中，所有路径共有的最长的路径
l = ['/Users/yijun/test.txt', '/Users/yijun/test/aaa.txt', '/Users/yijun/bbb.tx
t']
os.path.commonprefix(l)
```

Out[46]:

```
'/Users/yijun/'
```


In [8]:

```
# 遍历一个目录下的所有文件
```

```
import os
```

```
def list_dir(root_dir):  
    for lists in os.listdir(root_dir):  
        path = os.path.join(root_dir, lists)  
        print(path)  
        if os.path.isdir(path):  
            list_dir(path)
```

```
# 注意不要挑选目录下过多文件的，否则会耗费电脑资源
```

```
list_dir('/Users/yijun/Documents/dev_python/dig/doc')
```

```
/Users/yijun/Documents/dev_python/dig/doc/dig_v100_详细设计特性列表.doc  
x  
/Users/yijun/Documents/dev_python/dig/doc/r2c1_v101_详细设计.docx  
/Users/yijun/Documents/dev_python/dig/doc/r2c1_v102_详细设计特性列表.do  
cx  
/Users/yijun/Documents/dev_python/dig/doc/readme.txt  
/Users/yijun/Documents/dev_python/dig/doc/reportx_design.md  
/Users/yijun/Documents/dev_python/dig/doc/reportx_design.txt
```

In [7]:

```
# 遍历一个目录下的所有文件
```

```
# 显示文件的字节数，用 getsize()
```

```
import os
```

```
def list_dir(root_dir):  
    for lists in os.listdir(root_dir):  
        path = os.path.join(root_dir, lists)  
        if lists[0:1] != '.':  
            filesize = os.path.getsize(path)  
            print(path, ' ', filesize)  
            if os.path.isdir(path):  
                list_dir(path)
```

```
# 注意不要挑选目录下过多文件的，否则会耗费电脑资源
```

```
list_dir('/Users/yijun/Documents/dev_python/dig/doc')
```

```
/Users/yijun/Documents/dev_python/dig/doc/dig_v100_详细设计特性列表.doc  
x    65257  
/Users/yijun/Documents/dev_python/dig/doc/r2c1_v101_详细设计.docx    69  
549  
/Users/yijun/Documents/dev_python/dig/doc/r2c1_v102_详细设计特性列表.do  
cx    20853  
/Users/yijun/Documents/dev_python/dig/doc/readme.txt    782  
/Users/yijun/Documents/dev_python/dig/doc/reportx_design.md    1752  
/Users/yijun/Documents/dev_python/dig/doc/reportx_design.txt    1544
```

In [9]:

```
# 遍历一个目录下的所有文件
# 过滤 . 开头的文件，一般是系统文件
# 显示文件的字节数
# 显示 .py 后缀的文件，引入 endswith 用法

import os

def list_dir(root_dir):
    for lists in os.listdir(root_dir):
        path = os.path.join(root_dir, lists)
        if lists[0:1] != '.' and lists.endswith('.py'):
            filesize = os.path.getsize(path)
            print(path, ' ', filesize)
            if os.path.isdir(path):
                list_dir(path)

# 注意不要挑选目录下过多文件的，否则会耗费电脑资源
list_dir('/Users/yijun/Documents/dev_python/dig')
```

```
/Users/yijun/Documents/dev_python/dig/__init__.py    1
/Users/yijun/Documents/dev_python/dig/dig.py         5253
/Users/yijun/Documents/dev_python/dig/dig_main.py    480
```

有趣的 **python** 命令

python -m this

python -m calendar

在终端方式下运行，看看是什么结果

思考一下

写一个可以搜索硬盘上指定路径指定类型的文件

In [13]:

```
import fnmatch
import os

images = ['*.jpg', '*.jpeg', '*.png', '*.tif', '*.tiff']
matches = []

for root, dirnames, filenames in os.walk('/Users/yijun/Documents/dev_python/machinelearninginaction/'):
    for extensions in images:
        for filename in fnmatch.filter(filenames, extensions):
            matches.append(os.path.join(root, filename))

print(matches)
```

```
['/Users/yijun/Documents/dev_python/machinelearninginaction/Ch10/Portland.png']
```

In []: