目录

[Python编程基础 2](#_Toc477790637)

[变量 2](#_Toc477790638)

[字符串 2](#_Toc477790639)

[find 查找字符串是否含有子串 2](#_Toc477790640)

[replace 字符串替换 3](#_Toc477790641)

[列表 3](#_Toc477790642)

[遍历列表时获取每个元素的index 3](#_Toc477790643)

[字典 3](#_Toc477790644)

[元组 3](#_Toc477790645)

[函数式编程 3](#_Toc477790646)

[map 3](#_Toc477790647)

[filter 4](#_Toc477790648)

[reduce 4](#_Toc477790649)

[lambda 4](#_Toc477790650)

[闭包 5](#_Toc477790651)

[概念 5](#_Toc477790652)

[存在意义 5](#_Toc477790653)

[与普通函数不同 5](#_Toc477790654)

[使用场景 5](#_Toc477790655)

[面向对象 6](#_Toc477790656)

[类 6](#_Toc477790657)

[对象 6](#_Toc477790658)

[classmethod VS staticmethod 7](#_Toc477790659)

[内省 7](#_Toc477790660)

[inspect 7](#_Toc477790661)

[getmembers 7](#_Toc477790662)

[设计模式 8](#_Toc477790663)

[单例模式 8](#_Toc477790664)

[装饰器实现 8](#_Toc477790665)

[多线程 8](#_Toc477790666)

[threading.Thread 模块 8](#_Toc477790667)

[队列 9](#_Toc477790668)

[task\_done(): 9](#_Toc477790669)

[join(): 9](#_Toc477790670)

[内置库 9](#_Toc477790671)

[sys 9](#_Toc477790672)

[argv 9](#_Toc477790673)

[exc\_info() 9](#_Toc477790674)

[os 10](#_Toc477790675)

[argparse 10](#_Toc477790676)

[定义一个解析器 10](#_Toc477790677)

[创建定位参数(positional arguments) 10](#_Toc477790678)

[创建可选参数(optional arguments) 10](#_Toc477790679)

[第三方库 12](#_Toc477790680)

[日志记录 12](#_Toc477790681)

[日志等级 12](#_Toc477790682)

[控制器 14](#_Toc477790683)

[模块级别函数（重要） 15](#_Toc477790684)

[配置信息 15](#_Toc477790685)

[通过ini 15](#_Toc477790686)

[分析工具 16](#_Toc477790687)

[profilehook 16](#_Toc477790688)

[异步队列 17](#_Toc477790689)

[Celery 17](#_Toc477790690)

[安装: 17](#_Toc477790691)

[配置并运行 17](#_Toc477790692)

作者：梁毅强

邮箱：398504533@qq.com

# Python编程基础

## 变量

## 字符串

### find 查找字符串是否含有子串

s = “this is python string”

print s.find(‘this’)

#如果存在则返回子串在字符串的开始位置，这里是0，如果不存在则返回-1。

print ‘this’ in s

#如果存在则返回True,否则返回False

### replace 字符串替换

字符串替换

s = “a.cn,b.com” #将逗号替换为空格

new\_s = s.replace(“,”, “ ”)

### lower() 和 upper()

针对英文字母，进行大小转换

s = “abdE”

new\_str = s.lower() # abde

new\_str = s.upper() #ABDE

### format 格式化字符串

#### 语法

通过 {} 代替%

限定符模式语法 {xx:xx}

#### 映射

##### 通过位置

In [1]: '{0},{1}'.format('kzc',18)

Out[1]: 'kzc,18'

In [2]: '{},{}'.format('kzc',18)

Out[2]: 'kzc,18'

In [3]: '{1},{0},{1}'.format('kzc',18)

Out[3]: '18,kzc,18'

##### 通过关键词参数

In [5]: '{name},{age}'.format(age=18,name='kzc')

Out[5]: 'kzc,18'

##### 通过对象属性

class Person:

def \_\_init\_\_(self,name,age):

self.name,self.age = name,age

def \_\_str\_\_(self):

return 'This guy is {self.name},is {self.age} old'.format(self=self)

##### 通过下标

In [7]: p=['kzc',18]

In [8]: '{0[0]},{0[1]}'.format(p)

Out[8]: 'kzc,18'

#### 格式限定符

##### 填充与对齐

填充常跟对齐一起使用

^、<、>分别是居中、左对齐、右对齐，后面带宽度

:号后面带填充的字符，只能是一个字符，不指定的话默认是用空格填充

比如

In [15]: '{:>8}'.format('189')

Out[15]: ' 189'

In [16]: '{:0>8}'.format('189')

Out[16]: '00000189'

In [17]: '{:a>8}'.format('189')

Out[17]: 'aaaaa189'

##### 精度与类型f

精度常跟类型f一起使用

In [44]: '{:.2f}'.format(321.33345)

Out[44]: '321.33'

##### 其他类型

主要就是进制了，b、d、o、x分别是二进制、十进制、八进制、十六进制。

In [54]: '{:b}'.format(17)

Out[54]: '10001'

In [55]: '{:d}'.format(17)

Out[55]: '17'

In [56]: '{:o}'.format(17)

Out[56]: '21'

In [57]: '{:x}'.format(17)

Out[57]: '11'

用，号还能用来做金额的千位分隔符。

In [47]: '{:,}'.format(1234567890)

Out[47]: '1,234,567,890'

参考：<http://blog.csdn.net/handsomekang/article/details/9183303>

### 字符串类型

#### str类型

str本质上是一串二进制字节序列，也叫byte string

#### unicode类型

字符（符号）

### 字符串编码

python2默认编码是ascii,python3统一用unicode

#### encode

将字符（符号）转换为二进制数据的过程

#### decode

将二进制字符序列转换为字符（符号）

#### 最佳实践

1.所有 text string 都应该是 unicode 类型，而不是 str，如果你在操作 text，而类型却是 str，那就是在制造 bug。

2.在需要转换的时候，显式转换。从字节解码成文本，用 var.decode(encoding)，从文本编码成字节，用 var.encode(encoding)。

3.从外部读取数据时，默认它是字节，然后 decode 成需要的文本；同样的，当需要向外部发送文本时，encode 成字节再发送。

## 列表

### 常用方法

list.append(x)

在列表最末尾添加元素; 等于a[len(a):] = [x].

list.extend(L)

将一个列表的所有元素添加到另一个列表中， 等于a[len(a):] = L.

list.insert(i, x)

在给定的位置插入一个元素，i是位置，x是元素，最常用的有a.insert(0, x)，从开头插入，a.inert(len(a),x) 相当于a.append(x)

list.remove(x)

从列表中移除一个元素，如果这个元素的值是x,如果没有这个元素，则会产生一个ValueError错误。

list.pop([i])

以给定的位置移除列表中的一个元素，并返回这个元素，如果不给定位置，则a.pop()移除最后一个元素。

list.index(x)

返回列表中以x为值的元素位置，如果没有，则抛出ValueError异常

list.count(x)

统计列表里x的个数

list.sort(cmp=None, key=None, reverse=False)

元素排序

示例：

ls = [(1, ‘F’), (2, ‘B’), (3, ‘D’), (0, ‘E’), (4, ‘A’)]

ls.sort(key=lambda x:x[1]) 以每个元素的第二个值排序

list.reverse()

返转排序

### del

有两种用法

1. 删除这个列表对象变量

ls = [1,2,3,4]

del ls #已经没有ls这个变量，后面不能使用

1. 删除列表里的元素

ls = [1,2,3,4]

del ls[:] # ls == []

### 列表复制

1. 赋值式指向

指向原来的内存地址

2.[:]

复制内容并指向新的内存地址

示例：

>>> ls = [1,2,3,4]

>>> l1 = ls

>>> id(l1)

46020848

>>> l2 = ls[:]

>>> id(l2)

46032976

>>> ls.append(33)

>>> l2

[1, 2, 3, 4]

>>> l1

[1, 2, 3, 4, 33]

### 像Stack堆栈一样使用list

stack 特性：last-in first-out

>>> stack = [3, 4, 5]

>>> stack.append(6)

>>> stack.append(7)

>>> stack

[3, 4, 5, 6, 7]

>>> stack.pop()

7

>>> stack

[3, 4, 5, 6]

>>> stack.pop()

6

>>> stack.pop()

5

>>> stack

[3, 4]

### 像Queue队列一样使用list

队列特性：first-in first-out

>>> from collections import deque

>>> queue = deque(["Eric", "John", "Michael"])

>>> queue.append("Terry") # Terry arrives

>>> queue.append("Graham") # Graham arrives

>>> queue.popleft() # The first to arrive now leaves

'Eric'

>>> queue.popleft() # The second to arrive now leaves

'John'

>>> queue # Remaining queue in order of arrival

deque(['Michael', 'Terry', 'Graham'])

### 遍历列表时获取每个元素的index

例：

jk = [1,2,3]

sni = ['aa','','bb']

pt = ['cc','dd','']

def get\_total():

totals = []

for idx,val in enumerate(jk):

#print 'idx:',idx, ' val:',val

totals.append({'jk':val, 'sni':sni[idx], 'pt':pt[idx]})

return totals

print get\_total()

[{'sni': 'aa', 'pt': 'cc', 'jk': 1}, {'sni': '', 'pt': 'dd', 'jk': 2}, {'sni': 'bb', 'pt': '', 'jk': 3}]

更为pythonic的写法

def get\_total2():

return [{'jk':val, 'sni':sni[idx], 'pt':pt[idx]} for idx, val in enumerate(jk)]

## 列表生成器

### 作用域

列表生成器有自己的作用域，与闭包不同，其作用域与外界是隔离的。

class A(object):

x = 1

gen = (x for \_ in xrange(10)) # gen=(x for \_ in range(10))

if \_\_name\_\_ == "\_\_main\_\_":

print(list(A.gen))

会抛NameError,因为gen有自己的作用域

## 字典

### iteritems

### items

## 元组

## 迭代器

迭代器（generator）是遵循迭代协议的对象，使用iter()从任何序列对象中得到迭代器（如list,tuple,set,dict等）

### iter

it = iter([1,2,3,4])

print it # <listiterator object at 0x03884390>

### generator

def my\_generator(n):

for i in range(n):

yield i

g = my\_generator(4)

print g

<generator object my\_generator at 0x03837DF0>

### 判断是否可迭代

form collections import Iterable

isinstance(‘abc’, Iterable) #True

isinstance(111, Iterable) #False

## 函数式编程

### map

用一个函数处理列表里的每个元素,这个函数是处理函数

例子：

sq = [1,2,4,5,6]

def add10(ls):

return map(lambda x:x+10, ls)

def add100(ls):

return map(lambda x:x+100, ls)

ls\_10 = add10(sq)

ls\_100 = add100(sq)

[11, 12, 14, 15, 16]

[101, 102, 104, 105, 106]

上例里 lambda x:x+10 和 lambda x:x+100就是处理函数

### filter

用一个函数对列表的元素进行过滤，这个函数只返回True,False

例子：

sq = range(20)

def is\_odd(x): #判断x是否是奇数

return x % 2 != 0

rs = filter(is\_odd, sq)

#[1, 3, 5, 7, 9, 11, 13, 15, 17, 19]

比较pythonic的做法：

filter(lambda x:x%2 != 0, range(20)) #一句话搞定

重点是lambda 语法

### reduce

用一个处理函数每次处理序列的前两个元素

例子：

ls = [1,2,3,4,5,6]

reduce(lambda x,y:x+y, ls) #21

处理方式：(((((1+2)+3)+4)+5)+6)

### lambda

函数原型：lambda [arguments]: expression

lambda 可以定义一个匿名简单函数

例子：

取出列表中是负数的数

sq = [1,3,-4,5,66,10,-8]

print filter(lambda x:x < 0, sq)

lambda x:x<0 转换成函数为

def isNegative(x):

if x < 0:return x

lambda 的表达式包含了返回值。

## 闭包

### 概念

在一些语言中，在函数中可以（嵌套）定义另一个函数时，如果内部的函数引用了外部的函数的变量，则可能产生闭包。闭包可以用来在一个函数与一组“私有”变量之间创建关联关系。在给定函数被多次调用的过程中，这些私有变量能够保持其持久性。

### 存在意义

同一个函数，根据参数的不同而带了不同的外部变量，实现不同的功能，可以理解为轻量级的接口封装。

### 与普通函数不同

拥有\_\_closure\_\_ 属性，里面定义了一个元组以保存所有的cell对象，每个cell对象保存了这个闭包中所有的外部变量。

def make\_printer(msg1, msg2):

def printer():

print msg1, '----',msg2

return printer

printer = make\_printer("foo", "bar")

printer.\_\_closure\_\_

printer.\_\_closure\_\_[0].cell\_contents

printer.\_\_closure\_\_[1].cell\_contents

### 使用场景

1. 装饰器
2. 对函数进行提前赋值

用functools.parial侧可，也可以用闭包实现

## 面向对象

### 类

新式类：

class MyClass(object):pass

继承自object

旧式类

无继承

class MyClass():pass

新式类与旧式类\_\_dict\_\_上的不同

**class MyOldTest**()**:pass  
  
class MyTest**(object)**:pass  
  
print** MyTest.\_\_dict\_\_  
**print** MyOldTest.\_\_dict\_\_

{'\_\_dict\_\_': <attribute '\_\_dict\_\_' of 'MyTest' objects>, '\_\_module\_\_': '\_\_main\_\_', '\_\_weakref\_\_': <attribute '\_\_weakref\_\_' of 'MyTest' objects>, '\_\_doc\_\_': None}

{'\_\_module\_\_': '\_\_main\_\_', '\_\_doc\_\_': None}

\_\_dict\_\_在这里是类属性，打印的是类变量

#### 类方法

##### \_\_new\_\_(cls[,...])

静态方法，用于创建类cls的新实例，其返回值是新的对象的实例

主要用来继承不可变类型(int、str、tuple)以定制化实例的创建

##### \_\_init\_\_(self[,...])

在实例(\_\_new\_\_())创建完成时调用

注：不可以返回任何值，否则会抛出TypeError

##### \_\_del\_\_(self)

在实例即将销毁时调用，也叫析构器

注：del x不会直接调用x.\_\_del\_\_() ----前者为减少一个x的引用，后者只是在x的引用达到零时调用

##### 多元比较操作

\_\_lt\_\_(self, other)

\_\_le\_\_(self, other)

\_\_eq\_\_(self, other)

\_\_ne\_\_(self, other)

\_\_gt\_\_(self, other)

\_\_ge\_\_(self, other)

以上为“多元比较”方法，优先于\_\_cmp\_\_()

##### \_\_cmp\_\_(self, other)

如果没有定义多元比较，则调用比较操作

##### \_\_hash\_\_(self)

由内建函数hash()调用和用于哈希后的集合成员操作，包括set、frozenset、dict。返回一个整数，如果一个类没有定义\_\_cmp\_\_()和\_\_eq\_\_()方法，那也不应该定义\_\_hash\_\_()

##### \_\_bool\_\_(self)

当对象进行bool(x)或 if x操作时，会调用\_\_nonzero\_\_方法，python3之后改为\_\_bool\_\_

class MyClass(object):

def \_\_init\_\_(self, x=0, y=0):

self.x = x

self.y = y

def \_\_nonzero\_\_(self):

print “\_\_nonzero\_\_ is call”

return False

##### \_\_slots\_\_

（非常重要的一个方法）

在新式类内部中定义的可以节省大量对象占用的内存空间。

示例：

class SlotsTest(object):

\_\_slots\_\_ = ['x', 'y', 'z']

def \_\_init\_\_(self, x, y, z):

self.x = x

self.y = y

self.z = z

def \_\_str\_\_(self):

return "{}-{}-{}".format(self.x, self.y, self.z)

s = SlotsTest(1,2,3)

print dir(s)

['\_\_class\_\_', '\_\_delattr\_\_', '\_\_doc\_\_', '\_\_format\_\_', '\_\_getattribute\_\_', '\_\_hash\_\_', '\_\_init\_\_', '\_\_module\_\_', '\_\_new\_\_', '\_\_reduce\_\_', '\_\_reduce\_ex\_\_', '\_\_repr\_\_', '\_\_setattr\_\_', '\_\_sizeof\_\_', '\_\_slots\_\_', '\_\_str\_\_', '\_\_subclasshook\_\_', 'x', 'y', 'z']

#对不在\_\_slots\_\_里的属性赋值

s.o = 22 #AttributeError: 'SlotsTest' object has no attribute 'o'

注：少了\_\_weakref\_\_

##### \_\_weakref\_\_

##### \_\_contains\_\_

包含关系测试，就像 a in mylist 一样，如果自定义的类定义\_\_contains\_\_,在实例里就可以通过 a in myobject这样的语法判断a是否在对象myobject中

示例：

class Tain(object):

def \_\_init\_\_(self, ls):

self.ls = ls

def \_\_contains\_\_(self, item):

return item in self.ls

ls = [11,22,33,44]

obj = Tain(ls)

print 11 in obj

### 对象

\_\_dict\_\_ 对象的这个属性可以打印实例变量，显示在\_\_init\_\_里定义的属性和动态定义的属性

**class MyClass**(object)**:** ivar **=** '22'  
 **def \_\_init\_\_**(self)**:** self.name **=** 'cat'  
  
v **=** MyClass()  
**print** v.\_\_dict\_\_  
v.age **=** 20 #动态定义的属性  
**print** v.\_\_dict\_\_

输出

{'name': 'cat'}

{'age': 20, 'name': 'cat'}

#### classmethod VS staticmethod

### 循环引用

class LeakTest(object):

def \_\_init\_\_(self):

print 'Object with id %d born here.' % id(self)

def \_\_del\_\_(self):

print 'Object with id %d dead here.' % id(self)

def foo():

A = LeakTest()

B = LeakTest()

A.b = B

B.a = A

#调用

foo()

#结果

Object with id 57713520 born here.

Object with id 57715312 born here.

没有显示dead here，所以说循环引用了，没有析构，长时间如此会引起内存泄漏

### 内省

#### inspect

##### getmembers

*# -\*- coding=utf-8 -\*-***class** MyCar(object):  
 **def** \_\_init\_\_(self, name, age):  
 self.name = name  
 self.age = age  
  
 **def** get\_name(self):  
 **return** self.name  
  
 **def** get\_age(self):  
 **return** self.age  
  
  
**import** inspect  
  
car = MyCar(**'aaa'**, 22)  
methods = inspect.getmembers(car, predicate=inspect.ismethod)  
**for** method **in** methods:  
 **if** method[0].startswith(**'get'**):  
 **print** method[0], **' '**, method[1]()

get\_age 22

get\_name aaa

## 设计模式

### 单例模式

#### 装饰器实现

**def singleton**(*cls*)**:** instances **=** {}  
  
 **def wrapper**(*\*args*, *\*\*kwargs*)**:  
 if not** instances.has\_key(cls)**:** instances[cls] **=** cls(**\****args*, **\*\****kwargs*)  
 **return** instances[cls]  
  
 **return** wrapper  
  
**@singleton  
class MyClass**(object)**:  
 def \_\_init\_\_**(self, *n*)**:** *#装饰器作用在init之前，执行装饰器的wrapper方法*

self.n **=** *n* **print** 'n:',*n*p **=** MyClass(10)  
  
p2 **=** MyClass(11)

## 多线程

### threading.Thread 模块

start

getName()

setName()

isDaemon()

setDaemon()

join(timeout)

run()

join等待直到线程结束

## 命令行操作

1.像shell脚本一样运行python脚本

如果脚本并不知道python解析器地址，要运行python脚本

1. 在脚本开头加上#!/usr/bin/env python
2. chmod +x script.py 赋执行权限
3. ./script.py 运行脚本

2.指定源文件编码

默认会以ASCII编码保存源文件要改变编码，需要在#!/usr/bin/env python下一行加上

# -\*- coding: encoding -\*-

3.-m 参数的用法

像运行脚本一样运行一个库模块

python -m SimpleHTTPServer 8080

再看SimpleHTTPServer 的代码

def test(HandlerClass = SimpleHTTPRequestHandler,

ServerClass = BaseHTTPServer.HTTPServer):

BaseHTTPServer.test(HandlerClass, ServerClass)

if \_\_name\_\_ == '\_\_main\_\_':

test()

文档提示，需要有\_\_main\_\_ 才能让模块运行于-m 参数下，也就是上述是运行了test()函数

## 队列

### task\_done():

指示前面get出来队列元素已经处理完成，

### join():

等待队列为空再执行别的动作

# 内置函数

## int

int(x, base=10)

默认是10进制，如果想计算一个2进制的数等于10进制的多少时，可以用base=2

int(“110”, base=2)

指定二进制数为110,计算出十进制为6

## abs(x)

返回一个数的绝对值。x可以是普通的整数，长整数或浮点数，如果参数是复数，则返回它的模

## basestring()

是str和unicode的超类，不能被调用或实例化，但可以用来测试一个对象是否是str或者unicode的实例。isinstance(obj, basestring) 等同于 isinstance(obj, (str, unicode))

## bin(x)

将一个整数转化成一个二进制字符串，结果是一个合法的Python表达式。

## locals()

局部变量，不可以修改

def f(arg):

x = 1

print locals()

f(333)

结果：{'x': 1, 'arg': 333}

## globals()

## all(iterable)

1. 如果iterable里的所有元素都是True(非空、非0)的，那么all返回True,否则只要有一个元素是假的，就返回False.
2. 如果iterable为空的，也会返回空.

相当于下面的代码：

def all(iterable):

for element in iterable:

if not element:

return False

return True

示例：

ls = [1,2,3]

all(ls) #True

ls2 = [1,None,33,2233]

all(ls2) #False 因为有None元素

ls3 = [1,0,332]

all(ls3) #False 因为有0

ls4 = []

all(ls4) #True 空列表也返回True

## compile()

语法：compile(source, filename, mode[, flags[, dont\_inherit]])

作用：将source编译为代码或AST对象

source 源码字符串

filename 通过这个参数可以导入模块文件

mode 有三种，分别是exec eval single

示例：

test\_module.py

def dd(x):

print “dd module say:”,x

>>> codeobj = compile('from test\_module import dd\nx=2\ndd(x)', "test\_module.py", "exec")

>>> exec(codeobj)

dd module say:2

通过filename导致了test\_module模块，source 里用了导入模块里的dd函数。

## any(iterable)

1.如果iterable里有一个真的元素，则返回True,否则返回False

2.如果iterable为空，则返回False

### type

#### type(object)

返回object的类型

class MyClass(object):

pass

print type(MyClass) # type

my = MyClass()

print type(my) # \_\_main\_\_.MyClass,新式类返回值，如果旧式类，则返回instance

#### type(name, bases, dict)

通过type函数创建类

def fn(self, name):

return 'name is:',name

NewClass = type("NewClass", (object,), dict(name=fn))

n = NewClass()

print n.name("leon") #name is leon

print type(n) # <class '\_\_main\_\_.NewClass'>

## sorted()

语法：sorted(iterable[, cmp[, key[, reverse]]])

示例1.

>>> sorted([5, 2, 3, 1, 4])

[1, 2, 3, 4, 5]

注：返回值为已经排好序的新列表

如果参数是列表，则列表自带sort()方法

>>> a = [5, 2, 3, 1, 4]

>>> a.sort()

>>> a

[1, 2, 3, 4, 5]

注：列表的排序已经被改变

### key 参数的用法

key 接收一个函数作为参数，用这个函数进行排序

示例：

student\_tuples = [

('john', 'A', 15),

('jane', 'B', 12),

('dave', 'B', 10),

]

sorted(student\_tuples, key=lambda x:x[2])

结果：[('dave', 'B', 10), ('jane', 'B', 12), ('john', 'A', 15)]

# 内置库

## copy

### copy

返回一个影子克隆

### deepcopy

返回一个深度克隆

## weakref

弱引用的主要用途是实现缓存或映射一个大对象,可以避免出现引用循环导致的内存泄漏现像，弱引用不会被引用计数器计算。

### ref

语法：(object[,callback])

创建一个弱引用对象，object是被引用的对象，callback为回调函数(当弱引用被删除时，会调用此函数)

class Foo(object):

pass

>>> foo = Foo()

>>> ref = weakref.ref(foo)

>>> print ref()

<\_\_main\_\_.Foo object at 0x101823490>

>>> del foo

>>> print ref()

## math

### 数学理论函数

#### fabs(x)

返回x的绝对值（浮点型），如果不需要浮点值，可以直接用内置abs函数

#### floor(x)

返回只保存小数点左边的浮点数，小数点右边的值全变为0

### 三角函数

#### hypot(x, y)

与sqrt(x\*x+y\*y) 相等，功能是求向量中从原点到point(x,y)的距离

#### sin(x)

#### tan(x)

#### acos(x)

#### cos(x)

#### asin(x)

#### atan(x)

### 幂和对数函数

#### pow(x, y)

相当于xy

#### sqrt(x)

相等于x\*\*x 开平方

#### exp(x)

返回e\*\*x

#### log(x[,base])

如果只有一个参数，则返回x的自然对数 (base=e)

如果两个参数，则等价于log(x)/log(base)

### 常量

#### pi

pi = 3.141592...

#### e

e = 2.718281...

## functools

### partial

修改函数签名，使用一些参数固化，以生成一个更为简单的函数

示例：

1. 生成接受一个二进制字符串参数的函数，计算出十进制的值

>>>basetwo = partial(int, base=2)

>>>basetwo("11111") #31

### wraps

作用：对原函数的\_\_doc\_\_、\_\_name\_\_、\_\_module\_\_、\_\_dict\_\_ 重新设定

源码：

def wraps(wrapped,

assigned = WRAPPER\_ASSIGNMENTS,

updated = WRAPPER\_UPDATES):

return partial(update\_wrapper, wrapped=wrapped,

assigned=assigned, updated=updated)

示例：

from functools import wraps

def my\_decorator(f):

@wraps(f)

def wrapper(\*args, \*\*kwargs):

print "Calling decorated function"

return f(\*args, \*\*kwargs)

return wrapper

@my\_decorator

def foo():

"""Docstring"""

print 'Called foo function'

foo()

print foo.\_\_name\_\_

print foo.\_\_doc\_\_

如果不在装饰器上添加wraps,则取不到原函数的\_\_doc\_\_, \_\_name\_\_也变为修饰器的wrapper

## sys

### argv

argv[0] 脚本名称，是否显示全路径要看操作系统，根据验证，win7和centos6.5只显示文件名，不显示全路径，要显示全路径可以用os.path.realpath(argv[0])

argv[1..n] 脚本的参数，因为是通过标准输入，所以所传递的类型全都是str。

### exc\_info()

当程序产生异常时，此函数能返回含有三个值的元组，(type, value, traceback)

type: 异常对象类型

value:传递给异常对象的参数

traceback: traceback object

### path

sys.path 这个是模块搜索路径，默认是根据PYTHONPATH环境变量初始化，PYTHONPATH是一个路径列表

getrefcount

语法：getrefcount(object)

返回对象的引用计数，要注意的是，调用这个函数后，对象的引用计数要加1

## os

### environ

以mapping对象来表示字符串环境变量

>>>type(os.environ)

<type 'instance'> #表示它是一个对象

修改环境变量：os.environ["HOME"] = ""

源代码分析

1.如果没有定义，则定义为一个空字典

try:

environ

except NameError:

environ = {}

### putenv(varname, value)

动态设置环境变量，改变后影响os.system()、popen()、fork()、execv()启动的子进程

但不会直接更新os.environ，所以更倾向于直接修改os.environ

### getpid()

返回当前进程的ID

适用平台:Unix、Windows

### getppid()

返回当前父进程的ID

适用平台：Unix

### path.dirname

获取一个具有完整路径的文件的目录

如 /tmp/myfile.py

os.path.dirname(“/tmp/myfile.py”) # /tmp

### path.abspath

获取一个文件的完整路径

通过用法

当前文件为/tmp/myfile.py

os.path.abspath(\_\_file\_\_) # \_\_file\_\_ 在文件里调用时为当前的文件名

# 得到 /tmp/myfile.py

## argparse

详细用法参考：<http://www.jianshu.com/p/fef2d215b91d>

命令行选项、参数和子命令的解析器

### 定义一个解析器

parser = argparse.ArgumentParser(description='Example with long option names')

### 创建定位参数(positional arguments)

定位参数的意思是,不用带-的参数

简单示例：

保存文件为job.py

import argparse

parser = argparse.ArgumentParser()

parser.add\_argument("echo")

args = parser.parse\_args()

print args.echo #args.echo 为一个变量，即命令行传递过来的第一个参数

如：python job.py foo

则args.echo 的值为foo

### 创建可选参数(optional arguments)

两种方式：

1. 用 – 来指定短参数， 如-h
2. 用 - - 来指定长参数，如 - - help

示例：

import argparse

parser = argparse.ArgumentParser()

parser.add\_argument("-v", "--verbosity", help="increase output verbosity")

args = parser.parse\_args()

if args.verbosity:

print "verbosity turned on"

调用例子：

python prog.py -v 1

如果不给-v指定参数，会报错，如果要像-h这样不用指定参数，就要用action功能

例子二：

import argparse

if \_\_name\_\_ == '\_\_main\_\_':

parser = argparse.ArgumentParser(prog="uop\_jobs")

parser.add\_argument("-v", "--verbose", help="increate output verbose.")

parser.add\_argument("-g", "--group\_type", help="select dba|other sla group.")

#parser.add\_argument('echo')

args = parser.parse\_args()

#print args.echo

print 'verbose:',args.verbose

if args.group\_type == 'dba':

print 'select dba'

elif args.group\_type == 'other':

print 'select other'

else:

raise Exception("select group error.")

命令行调用：

[root@mytest scripts]# python learn\_arg.py -g other

verbose: None

select other

action 的用法

如果option arguments不需要加参数，像-h这样，则可以用action来指定

示例：

import argparse

if \_\_name\_\_ == '\_\_main\_\_':

parser = argparse.ArgumentParser()

parser.add\_argument("-g", "--group", action="store\_true", help="this is group select")

args = parser.parse\_args()

print "select group:",args.group

调用：

[root@mytest scripts]# python learn\_arg2.py

select group: False

[root@mytest scripts]# python learn\_arg2.py -g

select group: True

[root@mytest scripts]#

以上例子用了action=”store\_true”，意思是如果加了-g 则args.group的值为True,否则为False

## operator

### attrgetter(attr[, args...])

### itergetter(item[, args...])

### methodcaller(name[, args...])

## inspect

反射、内省库

### getmembers()

### getmoduleinfo()

### getmodulename()

### getclasstree()

### getargspec()

获取函数参数，返回一个包含四个值的元组(args, varargs,keywords, defaults)

例子：

def foo(name, val, tag=22):

print 'name:', name, ' val:',val, ' tag:',tag

getargspec(foo)

//output ArgSpec(args=['name', 'val', 'tag'], varargs=None, keywords=None, defaults=(22,))

def bar(\*args, \*\*kwargs):

pass

getargspec(bar)

//output ArgSpec(args=[], varargs='args', keywords='kwargs', defaults=None)

### getargvalues()

## itertools

### count

语法：count(start=0, step=1)

无限迭代，从start开始

### cycle

语法：cycle(iterable)

无限迭代

cycle('ABCD') --> A B C D A B C D A B C D ...

repeat

语法：repeat(object[, times])

repeat(10, 3) # 10 10 10

#源码

def repeat(object, times=None):

# repeat(10, 3) --> 10 10 10

if times is None:

while True:

yield object

else:

for i in xrange(times):

yield object

### imap

语法：imap(function, \*iterables)

imap(pow, (2,3,10), (5,2,3)) --> 32 9 1000

### starmap

语法: starmap(function, iterable)

示例：

ls = starmap(pow, [(2,2), (2,3), (2,4)]) #返回一个iter对象

for i in ls:

print i #4 8 16

### product

语法:product(\*itables[, repeat])

生成笛卡尔乘积，repeat为重复次数

示例：

pd = product('ABCD', 'xy', repeat=3) #’ABCD’ ‘xy’ 为两个可迭代对象

## open 文件操作

### readline

一次读取一行，对大文件来说效率高，如果要清除首行，可以这样写，

**def** read\_file\_func():  
 file\_path = **'/proc/net/tcp'** fp = None  
 **try**:  
 **try**:  
 fp = open(file\_path)   
 fp.readline()  
 for line in fp:  
 print 'line:',line  
**except** Exception **as** e:  
 **print 'exception:'**,e  
 **finally**:  
 **if** fp:  
 **print 'close'** fp.close()

### readlines

将文件以行的形式保存在一个list中

### read

读取整个文件，将数据放在一个字符串内存变量中，如果文件大于可用内存，则会出错

## collections

### namedtuple

生成

Cards = namedtuple("Cards", ["rank", "suit"])

card = Cards("3", "J")

print card #Cards(rank='3', suit='J')

## subprocess

本模块允许你产生新的进程，连接它们的input/output/error 管道，并且获得返回代码，这个模块用于替代下面几个老函数

~~os.system~~

~~os.spawn\*~~

~~os.popen\*~~

~~popen2.\*~~

~~commands.\*~~

### Popen

语法：

subprocess.Popen(args, bufsize=0, executable=None, stdin=None, stdout=None, stderr=None, preexec\_fn=None, close\_fds=False, shell=False, cwd=None, env=None, universal\_newlines=False, startupinfo=None, creationflags=0)

参数：

cwd 转移到指定的目录下

最要的一个子模块

以下是它的方法

#### poll()

#### wait()

#### communicate(input=None)

进程交互用，用以发送数据到stdin,从stdout和stderr读取数据,一直读到文件结束，等街进程终止，input参数是一个字符串，为了发送给子进程，如果不需要发送数据给子进程，则设置为None

#### send\_signal(signal)

#### terminate()

#### kill()

示例：

1. 将结果重定向到文件

fd = os.open("/tmp/subprocess\_log", os.O\_WRONLY) #读写方式创建一个文件描述符

s = subprocess.Popen(["cat", "/tmp/report\_cron\_traffic.log"], stdout=fd, stderr=subprocess.PIPE)

s.communicate() #这个在这里可用可无

可以查看 /tmp/subprocess\_log文件已经将结果写入

### call

语法：subprocess.call(args, \*, stdin=None, stdout=None, stderr=None, shell=False)

示例：

1. 显示当前目录的列表

>>>subprocess.call([“ls”, “-l”])

....显示目录下的文件信息

0

>>>subprocess.call(“ls -l”, shell=True)

....显示目录下的文件信息

0

shell 默认为False

参数：

args 必须是一个字符串或者序列，提供序列是首选，因为它可以让模块避免执行一些危险操作，和控制限额（如控制执行的命令），如果提交的是字符串，shell参数必须为True

shell 默认为False ,当为True时，会将args当作shell标准的指令来执行，通过这样做是比较危险的，因为

stdin，stdout,stderr 指定执行程序的标准输入输出和标准错误，有效的值可以是PIPE,一个存在的文件描述符，一个存在的文件对象和None.在这个函数里有要用stdout=PIPE和stderr=PIPE

### check\_call

### check\_output

### PIPE

### STDOUT

# 第三方库

## 服务器软件

### uwsgi

#### 重要的参数

socket:

http:

### nginx（前端代理）

## 日志记录

### 日志等级

| **Level** |  |  | **Numeric value** |  |
| --- | --- | --- | --- | --- |
| CRITICAL |  |  | 50 | 特别糟糕的事情发生时，如内存耗尽、磁盘满 |
| ERROR |  |  | 40 | 有错误发生时，  def get\_user\_by\_id(user\_id):      user = db.read\_user(user\_id)      if user is None:          logger.error('Cannot find user with user\_id=%s', user\_id)          return user      return user |
| WARNING |  |  | 30 | 有重要的事件发生时，但并不是错误  def authenticate(user\_name, password, ip\_address):      if user\_name != USER\_NAME and password != PASSWORD:          logger.warn('Login attempt to %s from IP %s', user\_name, ip\_address)          return False      # do authentication here |
| INFO |  |  | 20 | 在处理请求或者服务器状态变化等日常事务中，我会使用 INFO 等级  def handle\_request(request):      logger.info('Handling request %s', request)      # handle request here        result = 'result'      logger.info('Return result: %s', result)    def start\_service():      logger.info('Starting service at port %s ...', port)      service.start()      logger.info('Service is started') |
| DEBUG |  |  | 10 | 大多数的情况下，你都不想阅读日志中的太多细节。因此，只有你在调试过程中才会使用 DEBUG 等级。我只使用 DEBUG 获取详细的调试信息，特别是当数据量很大或者频率很高的时候，比如算法内部每个循环的中间状态。  def complex\_algorithm(items):      for i, item in enumerate(items):          # do some complex algorithm computation            logger.debug('%s iteration, item=%s', i, item) |
| NOTSET |  |  | 0 |  |

一般情况最后的notset 不会使用

设置日志等级的方法

**import** logging  
logging.basicConfig(level=logging.DEBUG)

根据等级向上冒泡原则，等级数值越大，权限越高

设定等级的好处：

1. 控制消息级别，过滤掉不重要的信息，只关注重点信息

### 控制器

|  |  |  |
| --- | --- | --- |
| 控制器 | 解释 | 方法 |
| StreamHandler | 流控制器，类似于--stdout --stderr | emit()  flush() |
| FileHandler | 文件控制器，保存到文件进行持久化 | close()  emit() |
| RotatingFileHandler | 放置文件控制器  RotatingFileHandler(filename, mode='a', maxBytes=0, backupCount=0, encoding=None, delay=0)  可控制日志文件大小，保存日志数量 | doRollover()  emit() |
| TimedRotatingFileHandler |  |  |
| SocketHandler | 基于TCP,保存到网络 | close()  emit()  handleError()  makeSocket()  makePickle()  send()  createSocket() |
| DatagramHandler | 继承自Sockethandler,基于UDP | emit()  makeSocket()  send() |
| SysLogHandler | 参数linux的SysLog | close()  emit()  encodePriority()  mapPriority() |

可以定制日志的等级、日志保存方式、日志保存格式(format)等

handler = logging.FileHandler(**"./tlog.log"**) #保存在文件里  
handler.setLevel(logging.INFO) #等级为INFO  
logger.addHandler(handler) #添加

### 模块级别函数（重要）

logging.getLogger([name])

如果不设定name,则返回logger层级中的根logger(root logger),一般名称是点分隔的层级名，如”a”, “a.b”, “a.b.c.d”，以相同名字调用的总返回同一个logger

logging.getLoggerClass()

返回标准的Logger类，或者最后一次传给setLoggerClass()的类

logging.setLoggerClass(klass)

logging.debug(msg [,\*args[, \*\*kwargs]])

kwargs是两个关键词参数：

1. exc\_info 默认为False,设置为True后，异常信息会被加入到日志消息中,也可以传递一个tuple(以sys.exc\_info返回的形式)，将使用元组信息。
2. extra 字典形式，用以填充LogRecord的\_\_dict\_\_

### 配置信息

#### 通过ini

**import** logging  
**import** logging.config

*# load the logging configuration*logging.config.fileConfig(**'logging.ini',** disable\_existing\_loggers=False

)

##################配置示例#################

[loggers]  
keys=root  
  
[handlers]  
keys=fileHandler  
  
[formatters]  
keys=simpleFormatter  
  
[logger\_root]  
level=DEBUG  
handlers=fileHandler  
  
[handler\_fileHandler]  
class=FileHandler  
level=INFO  
formatter=simpleFormatter  
args=("/tmp/filelog.log",)  
  
[handler\_consoleHandler]  
class=StreamHandler  
level=DEBUG  
formatter=simpleFormatter  
args=(sys.stdout,)  
  
[formatter\_simpleFormatter]  
format=%(asctime)s - %(name)s - %(levelname)s - %(message)s  
datefmt=

####################################################

重要参数：disable\_existing\_loggers=False

表示如果模块里存在logger,则让其生效，默认值为True,会关闭已经存在的logger

## 分析工具

### profilehook

三个主要的装饰器

profile() 对函数进行性能分析

coverage() 计算函数被调用的次数

timecall() 简单统计函数处理时间

使用例子：

@profile  
**def** check\_fun():  
 **for** i **in** range(100000):  
 **pass**

## 异步队列

### Celery

#### 安装:

pip install –U celery flower

flower是一个专门监控celery性能的前端界面应用

#### 配置并运行

默认使用redis作为broker

示例：

tasks.py

#!/usr/bin/env python

# -\*- coding:utf-8 -\*-

from celery import Celery

app = Celery("tasks", backend="redis://localhost", broker='redis://localhost//')

@app.task()

def add(x, y):

return x + y

命令行下运行：celery -A tasks worker

计算：

>>> from task import add

>>> res = add.delay(1,2)

>>> res.get()

## 版本管理 PEP-440

### Version scheme

#### 公开版本的定义

[N!]N(.N)\*[{a|b|rc}N][.postN][.devN]

五个部分：

Epoch segment: N!

Release segment: N(.N)\*

Pre-release segment: {a|b|rc}N

Post-release segment: .postN

Development release segment: .devN

# 测试驱动开发

## selenium

### 安装selenium

python3 -m pip install --upgrade selenium

### 运行起来

*# -\*- coding=utf-8 -\*-***from** selenium **import** webdriver  
browser = webdriver.Firefox(executable\_path=**r'D:\django\_project\geckodriver.exe'**)  
browser.get(**"http://localhost:8000"**)  
**assert 'Django' in** browser.title

### 用unittest测试

*# -\*- coding=utf-8 -\*-***from** selenium **import** webdriver  
**import** unittest  
  
**class** NewVisitorTest(unittest.TestCase):  
 **def** setUp(self):  
 self.browser = webdriver.Firefox(  
 executable\_path=**r'D:\django\_project\geckodriver.exe'** )  
  
 **def** tearDown(self):  
 self.browser.quit()  
  
 **def** test\_can\_start(self):  
 self.browser.get(**"http://localhost:8000"**)  
 self.assertIn(**'Django'**, self.browser.title)  
 *# self.fail("Finish the test!")***if** \_\_name\_\_ == **'\_\_main\_\_'**:  
 unittest.main(warnings=**'ignore'**)

## unittest

### asserts

| **Method** | **Checks that** | **New in** |
| --- | --- | --- |
| [assertEqual(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertEqual) | a == b |  |
| [assertNotEqual(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertNotEqual) | a != b |  |
| [assertTrue(x)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertTrue) | bool(x) is True |  |
| [assertFalse(x)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertFalse) | bool(x) is False |  |
| [assertIs(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIs) | a is b | 3.1 |
| [assertIsNot(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIsNot) | a is not b | 3.1 |
| [assertIsNone(x)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIsNone) | x is None | 3.1 |
| [assertIsNotNone(x)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIsNotNone) | x is not None | 3.1 |
| [assertIn(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIn) | a in b | 3.1 |
| [assertNotIn(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertNotIn) | a not in b | 3.1 |
| [assertIsInstance(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIsInstance) | isinstance(a, b) | 3.2 |
| [assertNotIsInstance(a, b)](https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertNotIsInstance) | not isinstance(a, b) | 3.2 |

## Celery

### broker

消息传递的中间件，当应用程序调用celery异步任务时，会向broker传递消息，之后celery的worker会取出消息。

### backend

将消息持久化的组件