### python3学习手册

**简介**：

Python由Guido van Rossum于1989年底发明，于1991年发行第一版。

Python是一种解释型、面向对象、动态数据类型、可交互的语言

Python源代码遵循GPL协议

python 2.0于2000年10月16日发布，支持Unicode

python 3.0于2008年12月3日发布

官方于2020年1月1日停止更新2.x版本，Python 2.7为最后一个py 2.x版本

Python官网： https://www.python.org

PyCharm官网： https://www.jetbrains.com/pycharm/download/

python安装路径：

|  |  |  |
| --- | --- | --- |
| 操作系统 | python版本 | 安装路径 |
| windows | 2.7 | C:\Python27 |
| windows | 3.10 | C:\Users\用户名\AppData\Local\Programs\Python\Python310 |
| centos7 | 2.7.5 | /usr/bin/python2 --> /usr/bin/python2.7 (可执行文件) |

作者：李茂福

2022-03-26～

**0、linux执行python脚本的方式**

①**进入交互模式**

# python #输入python回车

>>> 执行python代码，类似shell脚本

>>> exit() #退出交互模式

②**执行一次性代码**

# python -c "要执行的代码，语句用;隔开"

③**执行脚本**

# chmod +x test.py

# ./test.py

或

# python test.py

**一、centos7源码安装python3**

centos7默认自带python2.7.5版本，无pip

先去官网下载源码包：https://www.python.org/downloads/source/

或者直接 wget https://www.python.org/ftp/python/3.7.12/Python-3.7.12.tgz



下载目标版本源码包，并上传到目标linux服务器上

安装依赖：

# yum install -y zlib-devel bzip2-devel openssl-devel ncurses-devel sqlite-devel \

readline-devel tk-devel gcc gcc-c++ make patch libffi-devel python-devel \

gdbm-devel xz-devel db4-devel libpcap-devel

# tar -xvf Python-3.x.x.tgz

# cd Python-3.x.x

# ./configure prefix=/usr/local/python3

# make && make install

创建软链接：

# ln -s /usr/local/python3/bin/python3 /usr/bin/python3

# ln -s /usr/local/python3/bin/pip3 /usr/bin/pip3

# python2.7 -c "import sys; print sys.path"

['', #程序当前路径

'/usr/lib64/python27.zip',

'/usr/lib64/python2.7',

'/usr/lib64/python2.7/plat-linux2',

'/usr/lib64/python2.7/lib-tk',

'/usr/lib64/python2.7/lib-old',

'/usr/lib64/python2.7/lib-dynload',

'/usr/lib64/python2.7/site-packages',

'/usr/lib/python2.7/site-packages']

# python3 -c "import sys; print(sys.path)"

['', #程序当前路径

'/usr/local/python3/lib/python37.zip',

'/usr/local/python3/lib/python3.7',

'/usr/local/python3/lib/python3.7/lib-dynload',

'/usr/local/python3/lib/python3.7/site-packages']

# python --version #查看版本

# python3 -V #查看版本

**二、基础语法**

**1.保留字（关键字）**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| True | False | None | as | nonlocal |  |
| and | exec | not | assert | finally | or |
| break | for | pass | class | from | print |
| continue | global | raise | def | if | return |
| del | import | try | elif | in | while |
| else | is | with | except | lambda | yiel |

**2.标准数据类型**

|  |  |
| --- | --- |
| Number | int bool. float complex ( 1+2j ) （py3无long长整型） |
| String | 有序strx='xxxx' stry="xxx" 不区别单引号与双引号，单个字符也是string |
| List | 有序 列表，数组 |
| Tuple | 有序 元组 |
| Set | 无序 不重复（自动删除重复项） 集合 |
| Dictionary | 无序 字典 {键值对 集合} 关键字必须互不相同 |
| bytes | 有序 字节数组，其元素值不可变，其元素类型为int（0~255） |
| bytearray | 有序 字节数组，其元素值可变，其元素类型为int（0~255） |

不可变类型： Number String Tuple bytes

可变类型： List Dictionary Set bytearray

set 集合是由若干个元素组成的，基本功能是 进行成员关系测试 和 删除重复元素

set可以进行数学上的集合运算（差集，并集，交集等）

可用 { ... } 和 set() 函数创建集合，创建空集合必须使用 set()而不能用 {}

例：

a=set("fldsaj")

b=set("fdslj")

a - b #a和b的差集

a | b #a和b的并集

a & b #a和b的交集

a ^ b #a和b中不同时存在的元素

**3.判断变量类型**

type(x) #返回x的类型，如 <class 'int'>

isinstance(x, int) #返回True或False

type()不会 认为子类是父类的一种类型

isinstance(x,int)会认为子类是父类的一种类型

**4.类型转换**

①将字符串**list转**为int型list

strx=["23","15","6","44"]

strnn=list(map(int, strx))

② int(x) int(x,base=n) 将x输出为10进制的整型

③ chr() 输入0～255内的整数，返回对应ascii字符，输入可为10进制，也可为16进制

④ unichr() 输入任意unicode编码值，返回对应unicode字符

⑤ ord() 输入为一个字符，ascii及unicode都行，输出为对应的int编码值

⑥ hex() 将一个整型数字转为16进制字符串，0xnn

⑦ oct() 将一个整型数字转为8进制字符串，0onn

**5.字符串大小写转换及替换过滤**

strx="what cAnd Wfds sfjdsD"

strx.lower() #全转为小写

strx.upper() #全转为大写

strx.capitalize() #句首字母大写

strx.title() #将每单词首字母大写

strx.swapcase() #全部大小写翻转，互换

strx.replace("old", "new") #把strx中的old替换成new

strx.rstrip() #过滤掉换行符及行尾的所有空白符

strx.lstrip() #过滤掉行首的所有空白符

strx.strip() #过滤位于行首以及行尾的所有空白字符以及换行符

strx.split() #将strx拆分，默认以空白符（含换行符）为拆分符

strx.split(" ") #将strx拆分，""之间指定拆分符

strx.split(" ", 1) #将strx拆分，""之间指定拆分符，且只拆分一次，即分成2部分

**6. print输出带颜色的字**

此种方法仅linux命令行下有效

print( "\033[0m 带颜色的文字 \033[0m" )

# \033[和m之间的数字为显示效果及颜色的 数字代码

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数字代码 | 效果 | 颜色 | 前景色代码 | 背景色代码 |
| 0 | 默认，灰白 | 黑 | 30 | 40 |
| 1 | 粗体 | 红 | 31 | 41 |
| 4 | 下划线 | 绿 | 32 | 42 |
| 5 | 闪烁 | 黄 | 33 | 43 |
| 7 | 反白显示 | 蓝 | 34 | 44 |
|  |  | 紫 | 35 | 45 |
|  |  | 青 | 36 | 46 |
|  |  | 灰白，同0默认 | 37 | 47 |

效果+颜色用;分号隔开

print("\033[4;31m这是带下划线红色的字\033[0m")



**7.二维list**

直接定义：

matrix = [ [0,0,0], [0,0,0], [0,0,0] ]

间接定义：

matrix = [ [0 for i in range(3)] for i in range(4) ] #4个小list，每小list为3元素

三维类推：

matrix = [ [ [0 for i in range(3)] for i in range(4) ] for i in range(2) ]

**8.运算符**

①**算术运算符**

+ - \* /

% 取模，取余数

\*\* 幂，如 A \*\* 4 表示A的4次幂

// 取整除，取商的整数部分，向下取整，去尾法，或运算数为float，则结果为n.0

②**比较运算符**

== 等于 != 不等于

> 大于 >= 大于等于

< 小于 <= 小于等于

③**逻辑运算**

not and or

④**成员运算**

in # a in Listx a在Listx中则返回True

not in

⑤**身份运算符**

is # a is b 判断a和b是否引用自同一个对象，是则返回True

is not

# a is b 同 id(a) == id(b) # id(x)用于获取对象x的内存地址

⑥位运算符

& 与运算 ^ 异或

| 或 << 左位移

~ 取反 >> 右位移

\*优先级：

指数\*\* > 位运算 > 乘除 > 加减 > 位移 > 比较 > 赋值 > 身份 > 成员 > 逻辑

**9.条件语句（python不支持switch）**

if 判断句 :

语句体

elif 判断句 :

语句体

else:

语句体

**10.循环**

① **for循环**

for i in xx :

break #跳出当前循环，跳出for

使用索引：

for i in range(10) :

# To Do

②**while循环**

while 判断 :

循环体

continue #跳过剩下的循环体，进入下一轮loop

**11.整除**

在python 2中，**/** 如果参与运算的数为int，就是整除，只取整数

如果 /参与运算的数为float，则返回float，不再是整数

在python 3中，/ 不论参与运算的数为int还是float，不是整除，都返回float

若参与运算的数只有int，则可用 **//** 运算符，取整

也可用 import math ; math.trunc(x)来取整

**12.小数取整**

①内置函数round()为**四舍五入**，n.5则靠偶数取整 ROUND\_HALF\_EVEN

round(2.5)返回2

round(3.5)返回4

②math.ceil() **进一法**

import math

math.ceil(3.1)返回4

③math.floor() **去尾法**

import math

math.floor(3.99) 返回3

**13.遍历dictionary**

①遍历key

dictx = {"k1": "v1", "k2": "v2", "k3": "v3"}  
for key in dictx.keys():  
 print(key, "-------->", dictx[key])

②遍历value

dictx = {"k1": "v1", "k2": "v2", "k3": "v3"}  
for value in dictx.values():  
 print(value)

③遍历item

dictx = {"k1": "v1", "k2": "v2", "k3": "v3"}  
for k, v in dictx.items():  
 print(k, "---->", v)

**14.序列集合字典等 数据 增删改查操作**

①list数据操作（有序，数值可重复）

listxx.append(x) # 把元素x追加到列表的末尾  
listxx.insert(i, x) # 在索引i指定位置插入元素x，原listxx[i]后移一位  
listxx.extend(listyy) # 把listyy追加到列表的末尾  
listxx.remove(x) # 删除列表中第一个值为x的元素，若无这样的元素则返回一个error  
listxx.pop(i) # 从列表的指定索引位置移除元素，并返回此元素值，如果不指定i，则移除并返回最后一个元素

listxx.popleft() # 移除列表头元素并返回其值  
listxx.index(x) # 返回列表中第一个值为x的元素的索引  
listxx.count(x) # 返回x在列表中出现的次数  
listxx.sort() # 对列表中的元素排序，改变列表本身，返回值为None或error，不返回新list，默认升序，从小到大  
listxx.sort(reverse=True) # 排序，降序，从大到小

listxx.reverse() # 反转列表，头变尾，尾变头

del listxx[1:3] # 切割操作，删除列表的这一段，含头不含尾

newlist = sorted(listxx) # 返回新列表，且排好序了，默认升序，原listxx不变

②set数据操作（无序，数值不重复）

★若在创建set时，有多个重复的元素，则只会存储其一，元素只能是不可变对象

t=("wh","fdjs","fkdsl")  
setyy=frozenset(t) # 创建不可变set，无add remove update操作

setxx.add("xxx") # 新增元素

setxx.remove("bb") # 删除元素，若元素不存在则返回KeyError，存在则返回None

setxx.update(iterablexx) #添加多个元素，这些元素是iterablexx里的，iterablexx可为list,set,tuple

③dictionary数据操作（无序，key不重复）

dictxx["newkey"] = "newvalue" # 新增键值对，如果已存在则更新值

dictxx.update(lx) #添加多个键值对，lx为另一dictionary

value2 = dictxx["k2"] # 取指定key对应的值

k,v=dictxx.popitem() # 随机返回并删除一个键值对

del dictxx["k2"] # 删除指定key的这对键值对

**15.string和bytes转换**

strxx = "这是w" # 类型为 str 长度 3  
bytesxx = strxx.encode("utf8") # 将str编码为utf8，返回类型bytes，长度 7

print(type(bytesxx[0])) #返回int, bytesxx本身为bytes类型，其元素则为 int

★string的元素仍为str

stryy=bytesxx.decode("utf8") #将bytes类型解码为字符，可指定相应的编码

**16.bytes和bytearray**

bytes字节数组各元素值不可变，bytearray各元素值可变

bytes和bytearray的各元素值都是int类型，取值范围0~255

**①创建bytes数组**

b1 = b"abcd\xff" # 在字符前加修饰符b直接创建，只可以是ASCII字符  
b2 = bytes(3) # 表示创建长度为3的字节数组，使用0填充  
b3 = bytes(listx) # 将可遍历的对象转为字节数组，该对象各元素皆为int，且0~255

b4=bytes("strxx",encoding="utf8") #将str转为bytes，等价于 str.encode("utf8")

print(b3.hex()) # 输出16进制数字符串，默认是b'\x11\x16c\xfe'这种

**②创建bytearray数组**

ba1 = bytearray(3) # 表示创建长度为3的字节数组，使用0填充  
ba2 = bytearray(listx) # 将可遍历的对象转为字节数组，该对象各元素皆为int，且0~255  
ba3 = bytearray("strxx", encoding="utf8") # 将str转为bytearray, str.encode("utf8")默认返回的是bytes

print(ba2.hex()) # 输出16进制数字符串，默认是bytearray(b'\x11\x16c\xfe')

★ bytes和bytearray都可以使用 .decode("utf8") 方法转为string

**三、python推导式**

从一个数据序列构建成另一个新的数据序列

① **list推导式**

newlist=[表达式 for 变量 in 源列表 if 条件] #直接返回list类型

例：

names=["abc", "xdfjl", "fdklsaj"]

new\_names=[name.upper() for name in names if len(name) >3 ]

② **dictionary推导式**

newdic={ key表达式: value表达式 for 变量 in 列表 if 条件} #直接返回dictionary

③ **set集合推导式**

newset={表达式 for 变量 in 列表/元组} #直接返回set

④ **tuple元组推导式**

newtuple=(表达式 for 变量 in 列表) #直接返回的是生成器对象，不是tuple

例：

xx=tuple( x for x in range(1,10)) # tuple()可得元组

**四、函数def func**

**1.定义函数**

def fun\_name(参数):

"描述"

语句

return xx #默认返回 None

**2.参数传递**

①**不定长参数**

def func( \*var ): #参数以tuple形式传入，var为tuple

#或者传参数时直接传入多个值，以逗号隔开

#若传入的变量为tuple时，不带\*则表示只有一个参数，就是tuple本身，带\*则表示把tuple解开，变成多个参数，每个元素为一个参数

②**定长参数加不定长参数**

def func( var1, var2, \*var\_args\_tuple)

#前面2个为固定参数，后面可传入若干个可选参数

③**默认参数**

def func(var1, var2=10, var3=40)

#调用时若未传入参数2,3则使用默认值

**④传参是否可变**

不可变类型： number string tuple bytes 传递的是值

可变类型： list dictionary set bytearray 传递的是引用

将可变类型参数传入函数后，函数里面对参数的修改，会改变函数外的变量值

将不可变类型参数传入函数后，函数里面对参数的修改，不会改变函数外的变量值

**五、python3内置函数**

内置函数是指不用import导入任何模块就能直接使用的函数

①算术相关

abs(x) #返回x的绝对值

min(x1,x2,x3) #返回最小值

hex(x) #将x转换成16进制数，带0x

oct(x) #将x转换成8进制数，带0o

bin(x) #将x转换成2进制数，带0b

int(x) #将x转换成10进制数，默认认为x为10进制数/str，base=16指定x字符串为16进制数

divmod() 函数接收两个数字类型（非复数）参数，返回一个包含商和余数的元组(a // b, a % b)

②生成器

dict() #用于生成一个dictionary

enumerate() 函数用于将一个可遍历的数据对象(如列表、元组或字符串)组合为一个索引序列，同时列出下标和数据，一般用在 for 循环当中

slice() 函数实现切片对象，主要用在切片操作函数里的参数传递

sorted() 函数对所有可迭代的对象进行排序操作，默认升序，reverse=True则降序，返回新的list；list 的 sort 方法返回的是对已经存在的列表进行操作

③查看帮助类

help() 查看目标对象的帮助信息

dir() 函数不带参数时，返回当前范围内的变量、方法和定义的类型列表；带参数时，返回参数的属性、方法列表。如果参数包含方法\_\_dir\_\_()，该方法将被调用。如果参数不包含\_\_dir\_\_()，该方法将最大限度地收集参数信息

④设置，获取

setatrr() 设置对象的属性

getattr() 获取对象的属性

a = A()

>>> getattr(a, 'bar') # 获取属性 bar 值

1

>>> setattr(a, 'bar', 5) # 设置属性 bar 值

>>> a.bar

5

next() 返回迭代器的下一个项目。

next() 函数要和生成迭代器的 iter() 函数一起使用

id() 函数返回对象的唯一标识符，标识符是一个整数。CPython 中 id() 函数用于获取对象的内存地址

input() 函数接受一个标准输入数据，返回为 string 类型

staticmethod() 返回函数的静态方法

⑤判断

all() 函数用于判断给定的可迭代参数 iterable 中的所有元素是否都为 TRUE，如果是返回 True，否则返回 False（元素除了是 0、空、None、False 外都算 True）

any() 函数用于判断给定的可迭代参数 iterable 是否全部为 False，则返回 False，如果有一个为 True，则返回 True。元素除了是 0、空、FALSE 外都算 TRUE

⑥其他

eval() 函数用来执行一个字符串表达式（为python表达式），并返回表达式的值

open() 函数用于打开一个文件，并返回文件对象

ascii() 函数类似 repr() 函数, 返回一个表示对象的字符串, 但是对于字符串中的非 ASCII 字符则返回通过 repr() 函数使用 \x, \u 或 \U 编码的字符

str() 函数将对象转化为适于人阅读的形式



https://www.runoob.com/python3/python3-built-in-functions.html

**六、输入输出**

**1.标准输入input()**

标准输入，执行程序时会暂停住，让用户从console输入，返回string

stdin = input()

stdin2 = input("提示：")

在linux命令行执行程序脚本时，可使用 | 管道符输入到标准输入，这时就不会暂停住了

# echo "inputarg" | ./test.py

**2.命令行输入sys.argv**

import sys

print("以下为命令行参数：", sys.argv)

#sys.argv #为命令行输入的参数list

在linux命令行执行程序脚本时，可直接跟在脚本后的参数为 命令行参数

# ./test.py arg1 arg2

# echo "arg1" | xargs ./test.py

**3.标准输出print()**

**①不指定格式**

print( arg1, arg2, arg3 ) #输出时3个参数之间默认使用空格隔开

print( arg1, arg2, sep="" ) #输出时不使用任何字符隔开，也空在""内指定分隔符

print()默认是有换行的，即默认结尾符为\n，可用end=""指定结尾符

**②print()旧式字符串格式化**

同C语言的sprintf()的格式化字符串用法，用 % 操作符

print( "这是第%2d号，体重%5.2f公折" % (num, wei) )

%-6.3f 表示左对齐，共6位宽，3位小数，浮点型（默认左对齐）

|  |  |  |  |
| --- | --- | --- | --- |
| 类型码 | 含义 | 类型码 | 含义 |
| %c | 字符 | %s | 字符串 |
| %d | 十进制整数 | %u | 无符号整数 |
| %o | 八进制数 | %x | 小写十六进制数 |
| %X | 大写十六进制数 | %f | 浮点数 |
| %e | 科学计数法，小写e | %g | 自动采用最短的%f或%e |
| %F | 同%f，会把inf,nan写成INF,NAN | %E | 同%e,大写E |
| %G | 同%g，会大写 | %b | 二进制数，print()不支持 |

**4.字符串.format()格式化函数**

建议使用"字符串".format(argxx) 的新式字符串格式化函数，返回格式化后的字符串

.format()通过将strxx字符串中的{}来识别替换字段

{字段名!转换字段:格式说明}

print( "我叫{}，年龄{}岁".format("李", 20) )

**①字段名**

\*可省，就写{}，默认从左到右对应后面.format()里的arg

\*写数字，0，1，2 可重复用右边的字段，如0，1，1，1

\*写变量名 xxx

.format()里**传入tuple**参数时，得带上\*号，如：

cof = ("李", 20)

print("我叫{0}，年龄{1}".format( \*cof) )

**传入dictionary**时，得带上\*\*双星号

**传入类的对象**时，不带星号，在{}里取对象的属性

a = classx("李", 20)

print("我叫{0.name}，今年{0.age}".format(a))

**传入list,tuple,dictionary时**，若不想按顺序取值，可用.[i]按下标取值

a=[1,2,3]

print("我{0.[1]}，在{0.[0]}里".format(a))

**②转换字段**

!s 把参数str()一下，即调用str()去格参数

!r repr()

!a 返回ascii()，或\unnnn \xnnnn

**③格式说明 :**

前面有:冒号

顺序为 :填充 对齐方式 正负号 # 宽度 .精度 类型码

**对齐方式**

< 左对齐

> 右对齐

^ 居中

后面必须指定宽度

print("我在{:<6}号".format( n ))

**填充**

在:后，对齐符前 写指定符号，如用0填充

print("我在{:0>6}号".format( n ))

**正负号**

让正数带上+号

print("我在{:0>+6}号".format( n )) #0填充默认在+号与后面数字之间了，这不是我们想要的，这时得用=指定对齐方式，默认右对齐

print("我在{:0=+6}号".format( n ))

**数字形式带前缀**

#b 输出带0b前缀

#o 输出带0o前缀

#x 输出带0x前缀

**,数字分组分隔**

让数字以3个一组带上逗号分隔符

,d 仅对应,d，不可用,f ,x之类

\_d \_o \_x \_b 可对应这4个，以每4个字符为一组带\_下划线输出

**%百分比形式**

{:%} 自动将数字乘以100带上%号，浮点数表示了，默认6位小数

{:.2%} 指定2位小数精度

**类型码**

最后是类型码，同print()旧式字符串格式化里的类型码

d 十进制整数

**七、面向对象**

**类是对数据的构成 及数据能做什么(行为) 的描述**

**①创建类**

class 类名称:

"类的帮助信息，描述"

类的内容

class 类名称(基类1, 基类2): #派生类

类的内容

**②类的方法**

类的方法即类里的函数，python的类方法必须有一个额外的第一个参数，类方法第一个参数名为self，表示类的实例(对象)，而非类本身

定义时有self，但在使用类方法时，不用再写这个参数

对象：通过类定义的数据结构实例

**③构造函数 \_\_init\_\_(self)**

class MyClassA:  
 xxx = 4 # 类变量，它的值在这个类的所有实例间共享，可在内部/外部使用它  
 def \_\_init\_\_(self, name, age): # 构造函数，初始化方法  
 self.name = name  
 self.age = age

**④创建对象，对象属性**

类的实例化

对象名=类名(x) # x参数是传给\_\_init\_\_()方法的

对象可直接添加/删除/修改 属性

对象名.newAttr = value

del 对象名.xxAttr

使用函数来访问属性

getattr( obj, attrName ) #获取对象的属性

hasattr( obj, attrName ) #检查对象是否存在这个属性

setattr( obj, attrName, value) #设置属性的值，若属性不存在，则创建

delattr( obj, attrName) #删除属性

**⑤内置类属性**

\_\_dict\_\_ #类的属性，一个字典，由类的数据属性组成  
\_\_doc\_\_ #类的文档字符串  
\_\_name\_\_ #类名  
\_\_module\_\_ #类定义所在的模块  
\_\_bases\_\_ #类的所有父类，一个元组

⑥析构函数

\_\_del\_\_(self) 析构方法，在对象销毁时被调用

⑦特殊属性

\_protedAttr #以单下划线开头表示protected变量，只允许其本身与子类进行访问

\_\_privateAttr #以双下划线开头表示private变量，只允许这个类本身访问

\_\_privateMethod #以双下划线开头的方法为private方法

\_\_xxx\_\_() #特殊方法，以双下划线开头及双下划线结尾

对象不可访问类的私有变量（以双下划线开头），但可用

对象名.\_类名\_\_私有属性名 来访问

**八、异常处理**

异常是python对象，表示一个错误，当程序发生异常时，需要捕获处理它，否则程序会终止执行。

**①捕获异常可用try, except语句**

try:  
 #要执行的语句  
except xx类: #可以不匹配具体的错误类  
 #要执行的语句  
except yy类 as 数据:  
 #要执行的语句  
else:  
 #如果没有异常发生时 执行的语句  
finally:  
 #不管有无异常最终都要执行的语句

**②触发异常**

可以使用raise语句 自己 触发异常

def myFunc(varx):  
 if varx < 10:  
 raise Exception("提示信息", varx)  
 # 触发异常后，后面的代码不会再执行  
  
# 使用  
try:  
 myFunc(n)  
except Exception as err:  
# 要执行的语句，err就是抛出的varx

自定义异常类，创建一个新类，继承自Exception类

class MyError(Exception): #创建自定义类，继承自Exception  
 def \_\_init\_\_(self, arg):  
 self.arg = arg  
 # 其他语句  
  
# 捕获异常  
try:  
 raise MyError("xx") # 抛出异常  
except MyError as err:  
# to do

**九、文件操作**

使用内置函数open()打开一个文件，返回file对象，再调用此对象的相关方法进行读写操作

**①文件打开模式**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 模式 | 文本 | 二进制 | 指针位置 | 是否覆盖 | 含义 |
| 只读 | r | rb | 开头 | 否 | 只读；若文件不存在则报错 |
| 只写 | w | wb | 开头 | 是 | 全覆盖；打开文件时清空文件；不存在则新建 |
| 只写 | a | ab | 末尾 | 否 | 追加；不存在则新建 |
| 读写 | w+ | wb+ | 开头 | 是 | 读写，打开文件时就清空文件；若文件不存在则创建新的文件 |
| 读写 | r+ | rb+ | 开头 | 否 | 读写，写时从头插入，不覆盖；若文件不存在则报错 |
| 读写 | a+ | ab+ | 末尾 | 否 | 读写，写时追加；若文件不存在则新建 |

例：

fileobj = open("test.txt", "r")  
fileobj.close() # 关闭文件

**②处理打开文件时的异常**

try:  
 fileobj = open("test.txt", "r")  
 # to do  
except Exception as err:  
 print(err)  
finally:  
 if fileobj:  
 fileobj.close()

也可直接使用with去打开文件，就不就手动关闭文件了

with open('test.txt', 'r') as fileobj:  
 # to do

**③打开文件时指定编码**

打开文本文件时默认是跟随系统本地语言的编码的，最好是要给open()函数传入encoding参数，例如，读取GBK编码的文件：（操作文本文件一定要指定编码）

fileobj = open('gbk.txt', 'r', encoding='gbk')

**④文本文件操作**

**.write()**方法写入文件，打开文件默认是跟随系统本地语言编码，写入文件也是，得在打开文件时指定文本编码；写入内容时，默认不带换行符，得自己加

# coding=utf-8  
with open("test.txt","w+",encoding="utf8") as fileobj:  
 fileobj.write("默认是不换行的")  
 fileobj.write("，支持使用转义加n去换行\n这是第二行")

效果：



**.flush()** 方法刷新文件内部缓冲，把内部缓冲区的数据立刻写入文件

**.read(n)** 读方法，一次读n个字符（换行也算一个字符）

# coding=utf-8  
with open("test.txt","r+",encoding="utf8") as fileobj:  
 stx=fileobj.read(10)  
 print(stx)

**.readline()** 一次读一行，返回的字符串是带有换行符的

**.readlines()** 读取所有行，返回的是由每一行组成的list（带有换行符）

**⑤二进制文件操作，struct模块**

在打开文件时，操作符添加b就行，读写仍然使用.read()和.write()方法，读取的二进制内容是存储在bytes里，可以使用struct模块去解析成对应的数据类型

\*打开文本文件与打开二进制文件的区别：

**打开文本文件时**：遇到EOF(0x1A)则认为文件结束了，不管文件里是\r\n还是\n换行，读取到string里只有\n，写入时也会转换\n为相应的系统换行符

fp.read()返回的是string类型

**打开二进制文件**则会读取到文件结尾，不转换\r\n等字符，且不可指定encoding="编码"

fp.read()返回的是bytes类型

struct模块中最重要的三个函数是pack(), unpack(), calcsize()

struct.pack(fmt, v1, v2, ...) 按照给定的格式(fmt)，把数据封装成bytes(实际上是类似于c结构体的字节流)

struct.unpack(fmt, bytesxx) 按照给定的格式(fmt)解析bytes，返回解析后的tuple

struct.calcsize(fmt) 计算给定的格式(fmt)占用多少字节的内存

struct中支持的格式

|  |  |  |  |
| --- | --- | --- | --- |
| format | 对应的C类型 | python类型 | 字节数 |
| c | char | string of length 1 | 1 |
| s | char[] | string | 1 |
| b | signed char | int | 1 |
| B | unsigned char | int | 1 |
| h | short | int | 2 |
| H | unsigned short | int | 2 |
| i l | int long | int | 4 |
| I L | unsiged int/long | int | 4 |
| q | long long | ~~long~~ int | 8 |
| Q | unsigned long long | ~~long~~ int | 8 |
| f | float | float | 4 |
| d | double | float | 8 |
| ? | \_Bool | bool | 1 |

format指定原始数据的字节序：

|  |  |  |
| --- | --- | --- |
| 符号 | byte order | 字节对齐方式 |
| @ | native | 凑够4个字节 |
| = | native | 按原字节数 |
| < | little-endian | 按原字节数，小端字节序 |
| > | big-endian | 按原字节数，大端字节序，逻辑上是啥顺序就是啥 |
| ! | network = big-endian | 按原字节数，网络字节序 |

使用方法是在format的第一个位置指示即可

struct解析二进制串示例：

import struct  
with open("test.txt","rb+") as fileobj:  
 strb=fileobj.read(10)  
 a,b,c=struct.unpack("<2ih",strb)  
 print(a,b,c)

<2ih 表示使用小端字节序，2个int赋值给a和b两个变量，1个short赋值给c

import struct  
with open("test.txt","rb+") as fileobj:  
 strb=fileobj.read(10)  
 a,b,c=struct.unpack("<ih4s",strb)

<ih4s 表示小端字节序，一个i赋值给a，一个h赋值给b，4s表示4字节的string赋值给c

★struct.unpack()返回的是tuple，如果只有一个元素，则左边要写a,=struct.unpack()，带有个逗号，否则a就是tuple的类型了

struct将数据转换成二进制串示例

a = 5  
b = 89.4  
strb = struct.pack("if", a, b)  
print(len(strb)) #结果是8

★python3在将str="xxx"转换成3s时，会报错：

struct.error: argument for 's' must be a bytes object

需要在传入的字符参数前加b

import struct  
a=b'hello'  
b=b'world!'  
bytes=struct.pack('5s6s',a,b)

或者把str转为对应编码的bytes

strxx.encode("utf8")

★struct.pack()返回的是bytes类型二进制串

**十、模块，包**

包是一个分层次的文件目录结构，它定义了一个由模块及子包等组成的python应用环境

包是一种管理python模块命名空间的形式

包是一个文件夹，文件夹下必须存在\_\_init\_\_.py文件，该文件可以为空。

（\_\_init\_\_.py文件用于标识当前文件夹是一个包）包下的其他\*.py文件为模块

**①导入模块**

import 包名1.子包名.模块名  
# 使用模块里的变量/函数时得使用完整名称  
包名1.子包名.模块名.funxx()

from 包名1.子包名 import 模块名x  
# 这时可直接使用 模块加变量/函数的方式  
模块名x.funxx()

from .模块名x  
# 相对路径导入表示 模块名x 在本模块同目录下， ..2个点表示在父目录下的  
模块名x.funxx()

from 包名x import \*  
#这时，只会导入 包名x 下的\_\_init\_\_.py里的 \_\_all\_\_=[这里列出的函数及变量]

一个模块被另一程序第一次导入时，其主程序将运行，若不想让它运行，可做个判断：

if \_\_name\_\_=="\_\_main\_\_":  
 #被导入模块的主程序写这里，只要它自己执行时才运行这里的代码  
else:  
 #来自另一程序的调用

★一个模块被导入时，解释器默认会将它编译成可执行的字节码.pyc

在模块所处包下生成一个名为\_\_pycache\_\_的目录，里面就是各模块的字节码文件

linux下可设置环境变量不生成字节码文件：

export PYTHONDONTWRITEBYTECODE=1

import sys  
sys.dont\_write\_bytecode = True #在导入时指定不生成字节码

**②python解释器查找模块顺序**

python解释器会去哪里找要导入的模块呢？

\*首先查找sys.modules字典里的模块（这是以前导入的所有模块的缓存）

\*然后匹配内置的模块sys.builtin\_module\_names

\*最后根据sys.path路径下去找【pip安装的包在site-packages目录下】

都找不到就报错ModuleNotFoundError

# python2.7 -c "import sys; print sys.path"  
['', #程序当前路径  
'/usr/lib64/python27.zip',   
'/usr/lib64/python2.7',   
'/usr/lib64/python2.7/plat-linux2',   
'/usr/lib64/python2.7/lib-tk',   
'/usr/lib64/python2.7/lib-old',   
'/usr/lib64/python2.7/lib-dynload',   
'/usr/lib64/python2.7/site-packages',   
'/usr/lib/python2.7/site-packages']  
  
  
# python3 -c "import sys; print(sys.path)"  
['', #程序当前路径  
'/usr/local/python3/lib/python37.zip',   
'/usr/local/python3/lib/python3.7',   
'/usr/local/python3/lib/python3.7/lib-dynload',   
'/usr/local/python3/lib/python3.7/site-packages']

★判断当前环境是否有某个模块：

try:  
 import 模块名  
except Exception as e:  
 print(e)

如果没有指定的模块，则报错



**十一、pip操作**

pip是一个 Python 包安装与管理工具

Python 2.7.9 + 或 Python 3.4+ 以上版本都自带 pip 工具

pip安装的包默认在 site-packages目录下

# pip freeze #查看用pip安装的包

# pip list #查看所有pip包

# pip list -o #同 --outdated 查看所有包及可更新的版本

#pip install 包名 #安装包

#pip install 包名==x.x.x #安装包时指定相应版本号

#pip install -r 含包名列表及版本号的txt文件 #一次安装文件里的所有包

#pip install -U 包名 #更新包，同--upgrade

#升级pip本身的版本

#python -m pip install --upgrade pip

#只下载包，不安装

#pip download -d /下载目录 包名

#pip默认包下载地址为 https://pypi.org/simple

可指定成其他的pip源地址

**①命令行指定**

#pip install 包名 -i https://xxx.com/simple --trusted-host xxx.com

**②在当前用户家目录下创建.pip目录**，其下创建pip.conf文件（windows为家目录下的pip目录下创建pip.ini文件，内容如下

[global]

index-url=https://xxx.com/simple

trusted-host=xxx.com

国内的pip源地址：

http://mirrors.aliyun.com/pypi/simple

https://pypi.tuna.tsinghua.edu.cn/simple

**十二、python常用模块**

**1.decimal精确计算模块**

import decimal  
nn=decimal.Decimal(3.141592653589793)  
nn/=decimal.Decimal(2.2)  
print(nn)

Decimal数只能和Decimal数进行算术操作

import decimal  
print(decimal.getcontext()) #获取上下文，查看精度等默认参数，默认prec=28  
decimal.getcontext().prec=60 #设置为60位精度

**2.struct二进制数据处理模块**

请见第九章 文件操作 ⑤二进制文件操作，struct模块

**3.time时间日期模块**

时间是以秒为单位的浮点小数

时间戳表示从1970年1月1日0时到现在经过的秒数，传统4字节表示秒数只支持1970到2038年的时间

**①获取时间戳**

import time  
print("当前时间戳",time.time())  
print(type(time.time()))



**②时间结构体，有9个字段**

import time  
localtime=time.localtime(time.time())  
print(localtime)  
print(type(localtime)) #结果如下

time.struct\_time(tm\_year=2022, tm\_mon=4, tm\_mday=5, tm\_hour=22, tm\_min=19, tm\_sec=7, tm\_wday=1, tm\_yday=95, tm\_isdst=0)  
<class 'time.struct\_time'>

**③格式化时间**

import time  
print(time.strftime("%Y-%m-%d %H:%M:%S",time.localtime()))  
print(time.strftime("%a %b %d %H:%M:%S %Y",time.localtime())) #%a表示周几

 #日期有05

import time  
print(time.ctime(time.time()))  
print(time.asctime(time.localtime())) #同 time.ctime()

 #日期就一个 5

**④时间字符串转为时间结构体**

import time  
timestr="Tue Apr 05 22:22:15 2022"  
t=time.strptime(timestr,"%a %b %d %H:%M:%S %Y")  
print(t)  
print(type(t))

**⑤时间结构体转为时间戳**

import time  
timestr="Tue Apr 05 22:22:15 2022"  
t=time.strptime(timestr,"%a %b %d %H:%M:%S %Y")  
tstamp=time.mktime(t) #时间结构体转为时间戳  
print(tstamp) #结果 1649168535.0

**⑥暂停线程的运行**

import time  
time.sleep(5) #单位为 秒

**4.os操作系统接口 模块**

import os  
cdir = os.getcwd() # 返回当前工作目录str  
os.chdir("D:\\") # 切换当前目录

ret = os.system("dir") # 执行系统命令，成功执行返回0，失败返回非0数值

**5.sys模块**

import sys  
print(sys.argv) # sys.argv为标准输入参数列表，第0个元素为程序名本身

**6.math模块**

import math  
x = 2  
math.sqrt(x) # 返回x的平方根，float  
math.pi # 常量 pi=3.141592653589793

math.e # 常量 e=2.718281828459045  
math.log(x) # 返回log x的值，以e为底  
math.log10(x) # 返回log10 x的值，以10为底  
math.sin(x) # 返回sin x的值，x为弧度，360度为2π弧度  
math.cos(x)  
math.tan(x)

math.pow(x,y) #返回x的y次幂，同 x \*\* y

**7.random模块**

import random  
random.random() # 随机生成一个[0,1)范围内的实数，float  
random.uniform(0, 5) # 随机生成一个[0,5]范围内的实数，float  
random.randint(0, 8) # 随机生成一个[0,8]范围内的整数，int  
random.randrange(1, 100, 2) # 从[1到100]随机选一个整数，step默认为1，这里指定为2，即从1，3，5，7...选一个整数

**8.glob文件通配符模块**

import glob  
print(glob.glob("\*.py")) # 返回由 当前目录下匹配上的文件名 组成的list

9.re正则匹配模块

**10.pickle序列化模块**

序列化是指将内存中的对象 存储到文件中（二进制文件）

反序列化将文件内容读出并生成对象（恢复成对象）

import pickle  
datax = {"k1": "v1", "k2": "v2", "k3": "v3"}  
datay = ["fdas", "fkldsj", 1323]  
with open("pk.dat", "wb") as fp:  
 pickle.dump(datax, fp) # 一次存入一个对象，对象类型可不相同，但读出时得按相应顺序读出  
 pickle.dump(datay, fp)

import pickle  
with open("pk.dat", "rb") as fp:  
 datax = pickle.load(fp) # 一次读取一个对象，对象类型可不相同  
 datay = pickle.load(fp)

**十三、多线程**

Python3 通过两个标准库 \_thread 和 threading 提供对线程的支持

**①简单线程，函数式**

import \_thread  
import time  
  
def my\_thread\_func(thread\_name, count): # 定义线程函数  
 for i in range(count):  
 time.sleep(1)  
 print("线程名：{},第{}次输出".format(thread\_name, i))  
  
try:  
 \_thread.start\_new\_thread(my\_thread\_func, ("thread1", 5)) #创建子线程  
 \_thread.start\_new\_thread(my\_thread\_func, ("thread2", 5))  
except Exception as e:  
 print(e)  
  
for i in range(8): #主线程要是退出了，则子线程也结束了，要确保在子线程结束前，主线程仍在运行  
 time.sleep(1)  
 print("this is main thread")

**②threading模块Thread类**

提供复杂的方法

import threading  
import time  
  
class myThread(threading.Thread):  
 def \_\_init\_\_(self, name, delay): # 重写init初始化方法  
 threading.Thread.\_\_init\_\_(self)  
 self.name = name  
 self.delay = delay  
  
 def run(self): # 重写run方法，用户线程要执行的内容在这里  
 time.sleep(self.delay)  
 print("%s: %s" % (self.name, time.ctime(time.time())))  
 print("退出线程：" + self.name)  
  
thread1 = myThread("Thread-1", 4) # 创建子线程  
thread2 = myThread("Thread-2", 3)  
  
thread1.start() # 开始线程  
thread2.start()  
thread1.join() # 等待线程退出  
thread2.join()  
print("退出主线程")

**十四、网络通信**

**①Tcp-server**

import socket  
host = "0.0.0.0"  
port = 1234  
sock1 = socket.socket(socket.AF\_INET, socket.SOCK\_STREAM) # 创建socket，ipv4,tcp  
sock1.bind((host, port)) # 监听的服务端ip,port，为一个元组  
sock1.listen(5) # 等待客户端连接  
while True:  
 con, addr = sock1.accept()  
 print("客户端的地址为：", addr)  
 con.send("发给客户端的信息".encode("utf8")) #只能发送bytes  
 con.close()  
sock1.close() # 关闭socket

**②Tcp-client**

import socket  
host = "127.0.0.1"  
port = 1234  
sock1 = socket.socket(socket.AF\_INET, socket.SOCK\_STREAM) # 创建socket，ipv4,tcp  
sock1.connect((host, port)) # 连接服务端ip,port，为一个元组  
recv = sock1.recv(1024) # 接收对端发来的数据，最多收1024字节，返回bytes  
sock1.close() # 关闭socket

**③Udp-server**

import socket  
addr\_ser=("0.0.0.0",1234)  
sock1 = socket.socket(socket.AF\_INET, socket.SOCK\_DGRAM) # 创建socket，ipv4,udp  
sock1.bind(addr\_ser) # 监听的服务端ip,port，为一个元组  
while True:  
 data, addr = sock1.recvfrom(1024) # 接收客户端发来的数据，最多1024字节  
 print("客户端的地址为：", addr)  
 print("客户端发来的数据：", data)  
 sock1.sendto("发给客户端的信息".encode("utf8"), addr)  
sock1.close() # 关闭socket

**④Udp-client**

import socket  
addr\_ser = ("127.0.0.1", 1234)  
sock1 = socket.socket(socket.AF\_INET, socket.SOCK\_DGRAM) # 创建socket，ipv4,tcp  
sock1.sendto("发送给server的数据".encode("utf8"), addr\_ser)  
recv = sock1.recv(1024) # 接收对端发来的数据，最多收1024字节，返回bytes  
sock1.close() # 关闭socket

**⑤tcp-server与多线程**

import \_thread  
import socket  
  
def my\_thread\_func(con, addr): # 定义线程函数  
 print("本服务端为：",con.getsockname())  
 print("客户端的地址为：",addr)  
 con.send("发给客户端的信息".encode("utf8"))  
 con.close()  
  
host = "0.0.0.0"  
port = 1234  
sock1 = socket.socket(socket.AF\_INET, socket.SOCK\_STREAM) # 创建socket，ipv4,tcp  
sock1.bind((host, port)) # 监听的服务端ip,port，为一个元组  
sock1.listen(5) # 等待客户端连接  
while True:  
 con, addr = sock1.accept()  
 try:  
 \_thread.start\_new\_thread(my\_thread\_func, (con, addr))  
 except Exception as e:  
 print(e)