

# python文件操作

## 1.文件的编码

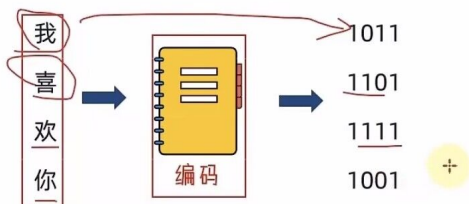
### 1.1掌握文件编码的概念和常见编码

思考：文本文件如何被计算机识别，并被计算机存储呢？

答案：使用编码技术（密码本）将内容翻译成0和1存储

#### 文件编码

编码技术即：翻译的规则，记录了如何将内容翻译成二进制，以及如何将二进制翻译回可识别内容。



计算机中有许多可用编码：

- UTF-8
- GBK
- Big5
- 等

#### 在我们电脑中呢

不同的编码，将内容翻译成二进制也是不同的。

编码有许多，所以要使用正确的编码，才能对文件进行正确的读写操作呢。



如上，如果你给喜欢的女孩发送文件，使用编码A进行编码（内容转二进制）。

女孩使用编码B打开文件进行解码（二进制反转回内容）

自求多福吧。

### 1.2文件编码的总结

#### 1. 什么是编码？

编码就是一种**规则集合**，记录了**内容和二进制**间进行相互转换的逻辑。编码有许多种，我们最常用的时UTF-8编码

#### 2. 为什么需要使用编码？

计算机只认识0和1，所以需要将内容翻译成0和1才能保存在计算机中，同时也需要编码，将计算机保存的0和1，反向翻译回可以识别的内容

## 2.文件的读取

### 2.1了解文件操作的作用

文件操作可以分为三个步骤：、

- 1. 打开文件
- 2. 读写文件
- 3. 关闭文件

### 2.2掌握文件的打开，读取，关闭操作

#### 1.创建文件对象

##### 1.open()打开函数

在Python中，使用open函数，可以打开一个**已经存在的文件**，或者创建一个新文件，语法如下：

`open(name,mode,encoding)`

name:是要打开文件名的字符串(可以包含文件所在的具体路径)

mode: 设置打开文件的模式(访问模式)：只读，写入，追加等。

encoding:编码格式(推荐使用UTF-8)

示例代码：

```
1 f = open('python.txt','r',encoding='utf-8')
2 注意：此时的'f'是'open'函数的文件对象。
```

##### 2.with open() as 对象名

```
1 with open('./demo/sss.txt','r',encoding='UTF-8') as f:
2     cont = f.read()
3     print(cont)
```

##### mode常用的三种基础访问模式

模式	描述
r	以只读方式打开文件。文件的指针将放在文件的开头，这是默认模式。
w	打开一个文件只用于写入。如果该文件已存在则打开文件，并从头开始编辑， <b>原有内容会被删除。如果该文件不存在，创建新文件。</b>
a	打开一个文件用于追加。如果该文件已存在， <b>新的内容将会写入到已有内容之后。</b> 如果该文件 <b>不存在，创建新文件</b> 进行写入

## 2.3文件对象的读操作方法

### 1.read()方法：有指针

文件对象.read(num)

num表示要从文件中读取的数据的长度(单位是字节), 如果没有传入num, 那么就表示读取文件中的所有数据。

**注意：**如果多次调用read()方法，下一次读取会在上一次read()读取结尾处接着读取。

### 2.readlines()方法：

readlines可以按照行的方式把整个文件中的内容进行一次读取，并且返回一个**列表**，其中每一行的数据为一个元素。

```
1 f = open('python.txt')
2 content = f.readlines()
3
4 #['哈哈\n', '嘿嘿\n', 'QQQQQQQQQQQQQ']
5 print(content)
6
7 f.close()
```

### 3.readline()方法：

readline一次读取一行内容

```
1 f = open('python.txt')
2 content = f.readline()
3
4 #['哈哈\n']
5 print(content)
6
7 f.close()
```

### 4.for循环读取文件行

```
1 for line in open('python.txt', "r"):
2     print(line)
3 #每一个line临时变量，就记录了文件的一行数据
```

### 5.close()文件的关闭

f.close()

## 2.4文件读取操作总结

操作	功能
文件对象 = open(file, mode, encoding)	打开文件获得文件对象
文件对象.read(num)	读取指定长度字节 不指定num读取文件全部
文件对象.readline()	读取一行
文件对象.readlines()	读取全部行，得到列表
for line in 文件对象	for循环文件行，一次循环得到一行数据
文件对象.close()	关闭文件对象
with open() as f	通过with open语法打开文件，可以自动关闭

## 2.5案例:

通过windows的文本编辑器软件，复制并保存到：word.txt，文件可以存储在任意位置，通过文件读取操作，读取此文件，统计itheima单词出现的次数。

```
1 count = 0
2 with open('./demo/word.txt.', 'r', encoding='UTF-8') as f:
3     for line in f:
4         line = line.strip() #去除开头与结尾的空格与换行符
5         words = line.split(' ')
6         for word in words:
7             if word == 'itheima':
8                 count += 1
9     print(f"itheima出现的次数是: {count}")
```

## 3.文件的写入

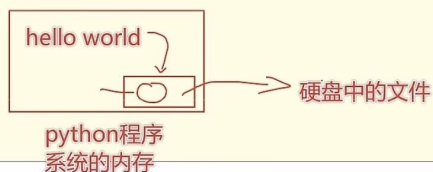
### 写操作快速入门

案例演示:

```
# 1. 打开文件
f = open('python.txt', 'w')

# 2. 文件写入
f.write('hello world')

# 3. 内容刷新
f.flush()
```



注意:

- 直接调用write，内容并未真正写入文件，而是会积攒在程序的内存中，称之为缓冲区
- 当调用flush的时候，内容会真正写入文件
- 这样做是避免频繁的操作硬盘，导致效率下降（攒一堆，一次性写磁盘）

```
1 r打开一个文件只用于写入。如果该文件已存在则打开文件，并从头开始编辑，原有内容会被删除。如果该文件不存在，创建新文件。
2 import time
3
4 f = open('./demo/word.txt', 'w', encoding='UTF-8')
5 f.write("hello world!") #将内容写入到内存中
6 f.flush() #将内存中积攒的内容写入到硬盘中
7 f.close() #close方法，内置了flush的功能
```

### 3.1文件写入的总结

1. 写入文件使用open函数中的“w”模式进行写入
2. 写入的方法有：
  - write(), 写入内容
  - flush(),刷新内容到硬盘中
3. 注意事项
  - w模式，文件不存在，会创建新文件
  - w模式，文件存在，会清空原有内容
  - close()方法中，带有flush()方法的功能

## 4.文件的追加

案例演示：

```
# 1. 打开文件，通过a模式打开即可
f = open('python.txt', 'a')

# 2. 文件写入
f.write('hello world')

# 3. 内容刷新
f.flush()
```

注意：

- a模式，文件不存在会创建文件
- a模式，文件存在会在最后，追加写入文件

### 4.1文件的追加总结

1. 写入文件使用open函数中的“a”模式进行写入
2. 写入的方法有：(和w模式一致):
  - write(), 写入内容
  - flush(),刷新内容到硬盘中
3. 注意事项
  - a模式，文件不存在，会创建新文件
  - a模式，文件存在，会在原有内容后面继续写入

- 可以使用“\n”来写出换行符
- close()方法中，带有flush()方法的功能

## 5.文件操作综合案例

### 5.1 完成文件备份的案例

需求：有一份账单文件，记录了消费收入的具体记录，内容如下：

name	date	money	type	remarks
周杰轮	2022-01-01	100000	消费	正式
周杰轮	2022-01-02	300000	收入	正式
周杰轮	2022-01-03	100000	消费	测试
林俊杰	2022-01-01	300000	收入	正式
林俊杰	2022-01-02	100000	消费	测试
林俊杰	2022-01-03	100000	消费	正式
林俊杰	2022-01-04	100000	消费	测试
林俊杰	2022-01-05	500000	收入	正式
张学油	2022-01-01	100000	消费	正式
张学油	2022-01-02	500000	收入	正式
张学油	2022-01-03	900000	收入	测试
王力鸿	2022-01-01	500000	消费	正式
王力鸿	2022-01-02	300000	消费	测试
王力鸿	2022-01-03	950000	收入	正式
刘德清	2022-01-01	300000	消费	测试
刘德清	2022-01-02	100000	消费	正式
刘德清	2022-01-03	300000	消费	正式

同学们可以将内容复制并保存为 bill.txt文件

我们现在要做的就是：

- 读取文件
- 将文件写出到bill.txt.bak文件作为备份
- 同时，将文件内标记为测试的数据行丢弃

实现思路：

- open和r模式打开一个文件对象，并读取文件
- open和w模式打开另一个文件对象，用于文件写出
- for循环内容，判断是否是测试不是测试就write写出，是测试就continue跳过
- 将2个文件对象均close()

代码

```
1 #打开文件对象，准备读取
2 fr = open("D:/bill.txt",'r',encoding="UTF-8")
3 #打开文件对象准备写入
4 fw = open('D:/bill.txt.bak','w',encoding='UTF-8')
5 #for循环读取文件
6 for line in fr:
7     line = line.strip() #去除每行开头结尾的空格和换行符
8     #判断内容，将满足的内容写出
9     if line.split(",")[4] == "测试":
10         continue #continue进入下一次循环，这一次后面的内容就跳过来
11     #将内容写出去
```

```
12     fw.write(line)
13     #手动添加换行符
14     fw.write("\n")
15 fr.close()
16 fw.close()
```