

第16课 编码与解码的战争

一、计算机的编码

1、编码（0或1）

2、存储单位：位/比特（bit），字节（byte）；

3、1 byte=8bit；1KB=1024B；1MB=1024KB；1GB=1024MB

4、编码表：ASCII码，GBK码，Unicode码，UTF-8码；Unicode是内存编码的规范，而UTF-8是如何保存和传输Unicode的手段

5、encode()和decode()

用法：'你想编码的内容'.encode('你使用的编码表')

'你想解码的内容'.decode('你使用的编码表')

6、b'\xce\xe2\xb7\xe3'，字母b代表的是字节，\x是分隔符

7、所谓的编码，其实本质就是把str（字符串）类型的数据，利用不同的编码表，转换成bytes（字节）类型的数据

二、读写文件

8、读取文件：打开文件-读文件-关闭文件

用法：

```
myfile = open('test.txt','r')
```

```
myfilecontent = myfile.read()
```

```
print(myfilecontent)
```

```
myfile.close()
```

9、写文件：打开文件-写文件-关闭文件

用法：

```
myfile = open('test1.txt','w')
```

```
myfile.write('从你的全世界路过')
```

```
myfile.close()
```

```
open('test1.txt')
```

10、open()函数

模式	说明
r	以只读方式打开文件。文件的指针将会放在文件的开头。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将其覆盖。如果该文件不存在，创建新文件。
a	打开一个文件用于追加。如果该文件已存在，文件指针将会放在文件的结尾。也就是说，新的内容将会被写入到已有内容之后。如果该文件不存在，创建新文件进行写入。
rb	以二进制格式打开一个文件用于只读。文件指针将会放在文件的开头。这是默认模式。
wb	以二进制格式打开一个文件只用于写入。如果该文件已存在则将其覆盖。如果该文件不存在，创建新文件。
ab	以二进制格式打开一个文件用于追加。如果该文件已存在，文件指针将会放在文件的结尾。也就是说，新的内容将会被写入到已有内容之后。如果该文件不存在，创建新文件进行写入。
r+	打开一个文件用于读写。文件指针将会放在文件的开头。
w+	打开一个文件用于读写。如果该文件已存在则将其覆盖。如果该文件不存在，创建新文件。
a+	打开一个文件用于读写。如果该文件已存在，文件指针将会放在文件的结尾。文件打开时会是追加模式。如果该文件不存在，创建新文件用于读写。
rb+	以二进制格式打开一个文件用于读写。文件指针将会放在文件的开头。
wb+	以二进制格式打开一个文件用于读写。如果该文件已存在则将其覆盖。如果该文件不存在，创建新文件。
ab+	以二进制格式打开一个文件用于追加。如果该文件已存在，文件指针将会放在文件的结尾。如果该文件不存在，创建新文件用于读写。

11、关键字with，无需用close()关闭

用法：

```
with open(r'test.txt','a') as myfile:
    myfile.write('你好')
```