

# 第三章：字符串格式化和编码格式代码笔记

## 转义字符

### 什么是转义？

在Python中\ 加一些英文字符的特定组合具有特殊意义。

#### \n 换行

```
# \n 换行符号
print("床前明月光，\n疑是地上霜。\n举头望明月，\n低头思故乡。")

# 拓展内容：
# print() 默认换行，为什么具有换行功能？
# 是因为在 Python的print源码中，设置了print的结束参数end 为\n
print("1")
print("2")

# 鼠标移动到 print 上，按下ctrl 键，待鼠标箭头变为小手点击 print 即可进入源码。

# 如果想不实现换行功能怎么办？
# 手动设置 结束参数，实现不换行
# end="。" 手动的将 print 的结尾换行的 \n 换成了 其他东西
print("这是一句话,我想不换行",end="。")
print("有没有换行呢? ")
```

#### \t 空格

```
# \t 水平制表符，自动补齐空格。
print("床前明月光，疑是地上霜。举头望明月，低头思故乡。")
print("床前明月光，\t疑是地上霜。\t举头望明月，\t低头思故乡。")

# 补齐空格。如果前面的字符不满四个字节的的话，就按照4字节补齐，如果前面的字节满4字节的话，按照8字节补齐。
# 英文字母,数字占一个字节,汉字占两个字节
print("abc\t12345\t哈哈")
```

#### \取消转义

```
# 在Python中 \ 加一些英文字符的特定组合具有特殊意义。
# 如果有些情况下，我希望能够取消这些特殊的意义
print("abc\\t12345\\t哈哈")
```

## 字符串编码

## 编码规范发展历史：

- 1、ASCII 编码，表示 127 个英文字符。刚开始只支持英语，其它语言不能够在计算机上存储和显示。
- 2、GBK 编码，中国编码。随着计算机普及，各国家语言不同，文字多少更是不同，127个根本不够用。各国都要为自己的语言编码
- 3、Unicode 规范，为了使国际间信息交流更加方便，国际组织制定了 UNICODE 字符集，为各种语言中的每一个字符设定了统一并且唯一的数字编号，以满足跨语言、跨平台进行文本转换、处理的要求。目前90%的网站使用的是 UTF-8 编码。UTF-8 把 unicode 进行优化。减小了硬盘存储空间浪费与传输效率低下。不过在内存中都要转换成 unicode。

```
# 字符串编码（将指定字符串，转换成特定编码格式的数据）
name = "不渝"
# encoding 必填参数，填写你要转变成什么编码格式
x = name.encode(encoding="GBK") # 返回值
print("编码后的数据：",end=" ")
print(x)

new_name = b'\xb2\xbb\xd3\xe5' #b 前缀用于表示字节字符串（bytes string）。
# 将 UTF-8格式的数据解码
y = new_name.decode(encoding="GBK") # 有返回值
print("解码后的数据：",end=" ")
print(y)

"注意：进行编码解码过程中，编码格式要统一。"
```

## 字符串拼接（格式化）

### +号拼接

```
# 字符串增加： + 拼接符号，用于拼接两个字符串
str1 = "不渝"
str2 = "上课暗号：1024"
str3 = str1+"老师，"+str2
print(str3)
```

### %s 万能占位符

```
# %s 万能占位符，先占一个位置，这个位置的内容待会再告诉电脑
print("我叫%s,上课暗号：%s"%( "不渝",1024))
a = "不渝"
b = 1024
print("我叫",a,"上课暗号：",b) #正常写法，稍微复杂，标点符号多要注意，容易出错
print("我叫%s,上课暗号：%s"%(a,b)) #占位符写法,更简单
```

### join函数拼接

```
# 拼接字符.join(元素)
# 列表中存了两个元素，分别是： 当前日期 ， 当前具体时间
dateTime = ('2023/7/31','21:36')
# 用 - 作为拼接符号，拼接 dateTime 中的元素
# join() 被拼接的必须是字符串
newTime = '-'.join(dateTime)
print(newTime)      # 2023/7/31-21:36
```

## format拼接

```
# format 格式化输出普通版
a = "不渝"
b = 1024
print("我叫%s,上课暗号: %s"%(a,b))
print("我叫{},上课暗号: {}".format(a,b))

# # format 格式化输出升级版
a = "不渝"
b = 1024
print(f"我叫{a},上课暗号: {b}")

# 举例
qiang = ["M762猛男枪","7.62mm子弹","平底锅"]
zuoBiao = "东149,西北11,东北132,西南168"
print(f"枪械列表: {qiang}\n敌人坐标: {zuoBiao}")

print("枪械列表: ",qiang,"\n敌人坐标: ",zuoBiao)
```