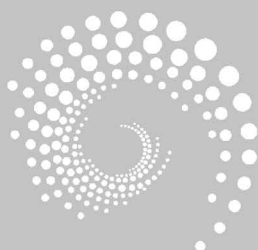


第3章

字符串操作与格式化

CHAPTER 3



3.1 实验目的与要求

- (1) 掌握字符串的编码、索引方式。
- (2) 掌握字符串的基本操作。
- (3) 掌握字符串格式化输出的方法。
- (4) 掌握基本数据类型的运算操作。

3.2 知识要点

- (1) 字符串是一个有序的字符集合,是不可变对象,以 Unicode 编码存储。
- (2) 字符串中的字符索引有两种方式:从左至右(正向递增序号,最左侧为 0)、从右至左(反向递减序号,最右侧为-1)。
- (3) 字符串的操作方法:使用内置运算符、内置函数、将字符串作为对象。
字符串处理函数: len()、str()、chr()、ord()。
字符串的分隔与合并方法: split()、join()。
字符串的大小写转换方法: lower()、upper()。
字符串的替换与删除方法: replace()、strip()。
字符串的子串出现次数: count()。
- (4) 字符串的输出可以使用 format() 格式化方法: <模板字符串>.format(<逗号分隔的参数>)。

3.3 实例验证

【实例 3-1】 索引和切片。已有一个字符串“坚持陆海统筹,加快建设海洋强国。”,根据用户输入的起始索引值和终止索引值,输出其对应的索引字符和切片结果。

解题指导: 索引从 0 开始,正数索引表示从左到右,负数索引表示从右到左,代码如下:

```
# 实例 3-1 索引和切片
sea = "坚持陆海统筹,加快建设海洋强国。"
print(sea)
i = eval(input("输入起始索引值:"))
print(sea[i])
j = eval(input("输入终止索引值:"))
print(sea[j])
print("切片为:", sea[i:j+1])
```

运行结果参考如下:

坚持陆海统筹,加快建设海洋强国。 输入起始索引值:2 陆 输入终止索引值:5 筹 切片为: 陆海统筹	坚持陆海统筹,加快建设海洋强国。 输入起始索引值:11 海 输入终止索引值:14 国 切片为: 海洋强国
---	---

【实例 3-2】 身份证号获取生日。通过用户输入的身份证号输出出生日期。

解题指导: 待切片的序列[起始索引(包含):结束索引(不包含):步长(默认为 1,可省略)]。代码如下:

```
# 实例 3-2 身份证号获取生日
No = input("请输入身份证号:")
y = No[6:10]
m = No[10:12]
d = No[12:14]
print("出生日期: {}年{}月{}日".format(y, m, d))
```

运行结果参考如下:

```
请输入身份证号:320705199605230529
出生日期:1996 年 05 月 23 日
```

【实例 3-3】 回文数判断。回文数是指一个数字从左边读和从右边读的结果一样,如 12321。

解题指导: 可以通过正向切片和反向切片进行比对。代码如下:

```
# 实例 3-3 回文数判断
n = input('请输入一个数:')
if n[0:-1] == n[-1:0:-1]:                                     # 等同于 if n[:] == n[::-1]
    print('是回文数!')
else:
    print('不是回文数!')
```

运行结果参考如下:

请输入一个数:12321 是回文数!	请输入一个数:123 不是回文数!	请输入一个数:131 是回文数!	请输入一个数:55 是回文数!
-----------------------	----------------------	---------------------	--------------------

【实例 3-4】 字符与 Unicode 码的互换。根据用户输入的字符串,将其中每个字符输出,并且将其对应的 Unicode 码以及 Unicode 码对应的字符后移两位的结果输出。

解题指导: 通过 for 循环依次取出一个字符进行转换。ord(s)函数是返回字符 s 所对应的 Unicode 编码,chr(x)函数是返回 Unicode 编码 x 所对应的字符。代码如下:

```
# 实例 3-4 字符与 Unicode 码的互换
s = input("请输入一个字符串:")
for i in range(len(s)):
    print(s[i], end = " ")                                     # 通过索引方式输出 i 所对应的字符
    print("对应的 Unicode 码是:", ord(s[i]), end = " ")
    print("Unicode 码对应的字符后移两位是:", chr((ord(s[i]) + 2)))
```

运行结果参考如下：

```
请输入一个字符串:aD3
a 对应的 Unicode 码是: 97  Unicode 码对应的字符后移两位是: c
D 对应的 Unicode 码是: 68  Unicode 码对应的字符后移两位是: F
3 对应的 Unicode 码是: 51  Unicode 码对应的字符后移两位是: 5
```

3.4 实验任务

1. 程序填空

【填空 3-1】 转换成大写字母。将用户输入的一个字符串中的字母全部转换成大写字母,请在如下代码中填空。

```
# tk3-1.py
m = input("请输入:")
print(_____)
```

运行结果参考如下：

```
请输入:ab1CD2eF
AB1CD2EF
```

【填空 3-2】 替换子串。用户输入一个带有“江海大”的字符串,请将其中“江海大”替换为“江苏海洋大学”并输出字符串,请在如下代码中填空。

```
# tk3-2.py
a = "江海大"
b = "江苏海洋大学"
s = input("请输入:")
print(_____)
```

运行结果参考如下：

```
请输入:江海大欢迎你!
江苏海洋大学欢迎你!
```

【填空 3-3】 4 位回文字符串判断。回文字符串是一个正读和反读都一样的字符串,如 noon 或“蜜蜂蜂蜜”等。现对用户输入的 4 个字符进行判断,如果是回文字符串,则显示“是”,否则“不是”,请在如下代码中填空。

```
# tk3-3.py
s = input("请输入 4 个字符:")
if s == s[3:4] + s[2:3] + s[1:2] + _____:
    print("是")
else:
    _____
```

运行结果参考如下：

请输入 4 个字符:noon
是

请输入 4 个字符:moon
不是

【填空 3-4】 汉字统计。已知变量 `s="岂曰无衣？与子同袍。王于兴师，修我戈矛。与子同仇！岂曰无衣？与子同泽。王于兴师，修我矛戟。与子偕作！岂曰无衣？与子同裳。王于兴师，修我甲兵。与子偕行！"`，统计并输出字符串 `s` 中汉字和标点符号的总个数，请在如下代码中填空。

解题指导：用 `str.count(sub)` 函数计算字符串 `str` 中 `sub` 子串出现的次数，用 `len(str)` 函数计算字符串 `str` 的长度。

```
# tk3-4.py
s = "岂曰无衣？与子同袍。王于兴师，修我戈矛。与子同仇！\
岂曰无衣？与子同泽。王于兴师，修我矛戟。与子偕作！\
岂曰无衣？与子同裳。王于兴师，修我甲兵。与子偕行！"
b = s.count('?') + s.count('.') + _____ + s.count('！')
a = _____ - b
print("汉字总数:{} 标点符号总数:{} ". _____))
```

运行结果参考如下：

汉字总数:60 标点符号总数:15

2. 编程

【编程 3-1】 字符统计。已有一个字符串内容为“中国人民解放军海军航空兵现已装备了轰炸机、巡逻机、电子干扰机、水上飞机、运输机等勤务飞机。”，请统计并输出其中“机”出现的次数。

解题指导：统计某字符串中的子串出现的次数用 `count` 方法，形式是：`str.count(sub[,start[,end]])`，即返回 `str[start:end]` 中 `sub` 子串出现的次数。

运行结果参考如下：

"机"的个数为: 6

【编程 3-2】 计算苹果重量。水果篮里装满了橘子和苹果，现由用户输入：水果篮总质量、橘子总质量以及苹果数量。请计算单个苹果的质量，保留两位小数位。

解题指导：第一行用户输入水果篮总质量(变量名：`weight`)，第二行用户输入橘子总质量(变量名：`orange`)，第三行用户输入苹果数量(变量名：`appleQty`)，第四行计算单个苹果的质量(变量名：`average`)，第五行输出结果。

运行结果参考如下：

请输入水果篮总质量(kg):6
请输入橘子总质量(kg):3
请输入苹果数量(个):12
单个苹果的质量:0.25kg

3.5 难点分析

1. 字符串的索引和切片

(1) 索引 $s[i]$: 取索引值为 i 的字符。

① 正向索引: 从 0 开始, 向右依次递增, 如: $s[3]$ 。

② 反向索引: 从 -1 开始, 向左依次递减, 如: $s[-2]$ 。

(2) 切片 $s[i:j:k]$: k 为步长, 取整数, 默认值为 1。

① $k=1$: 从索引值 i 开始取到 $j-1$ 位置的字符串, k 为 1, 可省略不写。如: $s[1:4]$ 等同于 $s[1:4:1]$ 。

② $k>1$ 或 $k<0$: 从索引值 i 开始取到 $j-1$ 位置的字符串, k 为正数时表示从左向右取值, 如: $s[1:6:2]$; k 为负数时, 表示从右向左取值, 如: $s[6:0:-2]$ 。

(3) 举例: 现有一个字符串 $s="victory"$, 输出索引和切片的结果。索引序号可以正向和反向进行标识, 如图 3-1 所示。



图 3-1 字符串"victory"的索引序号

针对"victory"字符串的索引和切片程序以及运行结果如表 3-1 所示。

表 3-1 字符串"victory"索引和切片的程序以及运行结果示例

程 序	运 行 结 果
<code>s = "victory"</code>	
<code>print('索引 s[0]:', s[0])</code>	索引 s[0]: v
<code>print('索引 s[-1]:', s[-1])</code>	索引 s[-1]: y
<code>print('索引 s[2]:', s[2])</code>	索引 s[2]: c
<code>print('索引 s[-2]:', s[-2])</code>	索引 s[-2]: r
<code>print('切片 s[:]:', s[:])</code>	切片 s[:]: victory
<code>print('切片 s[0:]:', s[0:])</code>	切片 s[0:]: victory
<code>print('切片 s[0:-1]:', s[0:-1])</code>	切片 s[0:-1]: victor
<code>print('切片 s[:-1]:', s[:-1])</code>	切片 s[:-1]: victor
<code>print('切片 s[0:7]:', s[0:7])</code>	切片 s[0:7]: victory
<code>print('切片 s[0:7:1]:', s[0:7:1])</code>	切片 s[0:7:1]: victory
<code>print('切片 s[0:-2]:', s[0:-2])</code>	切片 s[0:-2]: victo
<code>print('切片 s[0:7:2]:', s[0:7:2])</code>	切片 s[0:7:2]: vcoy
<code>print('切片 s[6:0:-1]:', s[6:0:-1])</code>	切片 s[6:0:-1]: yrotci
<code>print('切片 s[6:0:-2]:', s[6:0:-2])</code>	切片 s[6:0:-2]: yoc
<code>print('切片 s[6::-2]:', s[6::-2])</code>	切片 s[6::-2]: yocv
<code>print('切片 s[::-1]:', s[::-1]) # 反转</code>	切片 s[::-1]: yrotciv

2. 转义字符

(1) 含义：需要在字符中使用特殊字符时,用反斜杠(\)作为转义字符来输出特殊效果。常用转义字符如表 3-2 所示。

表 3-2 转义字符及其描述

转 义 字 符	描 述
\(在行尾时)	续行符
\n	换行
\t	横向制表符
\v	纵向制表符
\\	反斜杠符号
\'	单引号
\"	双引号

说明：转义字符作为一个整体,长度是 1。

(2) 举例：以常用的“\n”和“\t”为例,程序及运行结果如表 3-3 所示。

表 3-3 转义字符程序及运行结果示例

程 序	运 行 结 果
print("江苏省\n 连云港市\t 海州区")	江苏省 连云港市 海州区
print(len("江苏省\n 连云港市\t 海州区"))	12