

实用python编程 第4讲

字符串处理 string


2017-10-30

基本概念

- 字符串是 Python 中最常用的数据类型，用单引号‘或双引号“来创建字符串

```
var_str = "hello world"
```

- 字符串中的一个片段称为子字符串
 - 子字符串的长度不小于1，不超过原字符串
 - 通过方括号来截取
 - 不能被再次赋值

```
var_str[0] = "H" 
```

常用字符串操作(在notebook上练习)

```
a = "Hello", b = "Python"
```

操作符	描述	实例
+	字符串连接	<pre>>>>a + b 'HelloPython'</pre>
*	重复输出字符串	<pre>>>>a * 2 'HelloHello'</pre>
[]	通过索引获取字符串中字符	<pre>>>>a[1] 'e'</pre>
[:]	截取字符串中的一部分	<pre>>>>a[1:4] 'ell'</pre>
in	成员运算符 - 如果字符串中包含给定的字符/字符串返回 True	<pre>>>>"He" in a True</pre>
not in	成员运算符 - 如果字符串中不包含给定的字符返回 True	<pre>>>>"HE" not in a True</pre>
r/R	原始字符串 - 原始字符串：所有的字符串都是直接按照字面的意思来使用，没有转义特殊或不能打印的字符。原始字符串除在字符串的第一个引号前加上字母"r"或"R"以外，与普通字符串有着几乎完全相同的语法。	<pre>>>>print r'\n' \n >>> print R'\n' \n</pre>
%	格式字符串	

字符串操作举例

```
a = 'hello'
b = 'python'
c = a + b
print c
print len(c)
print c[1], c[1:2], c[1:3], c[1:-1], c[1:]
```

```
hellopython
11
e e el ellopytho ellopython
```

课堂练习1030-1

- 构造一个字符串，包含1个a，2个b，3个c，…，26个z

abbccdddeeeefffffgggggghhhhhhiiiiiiijjjjjjkkkkkkkkkkk
llllllllllllmmmmmmmmmmnnnnnnnnnnnooooooooooppppppppppp
ppppqqqqqqqqqqqrrrrrrrrrrsssssssstttttt
ttttttttttuuuuuuuuuuuvvvvvvvvvvwwwwwwwwww
wwwwwwwwwxxxxxxxxxxxxxxxxxyyyyyyyyyyyzzzz
zzzzzzzzzzzzzzzzzz

格式字符串 %

- 亲爱的xxx你好！你xx月的话费是xx，余额是xx
 - xxx的内容都是根据变量变化的
 - 需要一种简便的格式化字符串的方式
- 百分号 % 是占位符，用来格式化字符串
 - 在字符串内部，%s表示用字符串替换，%d表示用整数替换
 - 有几个%占位符，后面就跟几个变量或者值，一一对应。
 - 如果只有一个%括号可以省略

```
name = "Skywalker"
student_id = 2016000000
print "My name is %s" % (name)
print "My name is %s" % name
print "My name is %s and student id is %d" % (name, student_id)
print "My name is %s and student id is %d" % (student_id, name)
```

字符串格式化符号列表

符 号	描述
%c	格式化字符及其ASCII码
%s	格式化字符串
%d	格式化整数
%u	格式化无符号整型
%o	格式化无符号八进制数
%x	格式化无符号十六进制数
%X	格式化无符号十六进制数（大写）
%f	格式化浮点数字，可指定小数点后的精度
%e	用科学计数法格式化浮点数
%E	作用同%e，用科学计数法格式化浮点数
%g	%f和%e的简写
%G	%f 和 %E 的简写
%p	用十六进制数格式化变量的地址

```
a = 1.2
print '%f %g %d' % (a, a, a)

for a in range(20):
    print '%d %x %X' % (a, a, a)
```

字符串与其他数据类型的转换

转换	例子	结果
字符串 → 整数	<pre>a = '2' print int(a) + 1</pre>	3
字符串 → 浮点数	<pre>a = '2.14' print float(a) + 1</pre>	3.14
整数 → 字符串	<pre>a = 2 print str(a) print '%d' % a print '%g' % a</pre>	2 2 2
浮点数 → 字符串	<pre>a = 3.14 print '%f' % a print '%g' % a</pre>	3.140000 3.14

字符串自带函数

<http://www.runoob.com/python/python-strings.html>

方法	描述
string.capitalize()	把字符串的第一个字符大写
string.center(width)	返回一个原字符串居中,并使用空格填充至长度 width 的新字符串
string.count(str, beg=0, end=len(string))	返回 str 在 string 里面出现的次数, 如果 beg 或者 end 指定则返回指定范围内 str 出现的次数
string.decode(encoding='UTF-8', errors='strict')	以 encoding 指定的编码格式解码 string, 如果出错默认报一个 ValueError 的异常, 除非 errors 指定的是 'ignore' 或者 'replace'
string.encode(encoding='UTF-8', errors='strict')	以 encoding 指定的编码格式编码 string, 如果出错默认报一个 ValueError 的异常, 除非 errors 指定的是 'ignore' 或者 'replace'
string.endswith(obj, beg=0, end=len(string))	检查字符串是否以 obj 结束, 如果 beg 或者 end 指定则检查指定的范围内是否以 obj 结束, 如果是, 返回 True, 否则返回 False.
string.expandtabs(tabsize=8)	把字符串 string 中的 tab 符号转为空格, tab 符号默认的空格数是 8。
string.find(str, beg=0, end=len(string))	检测 str 是否包含在 string 中, 如果 beg 和 end 指定范围, 则检查是否包含在指定范围内, 如果是返回开始的索引值, 否则返回-1

在notebook中练习使用这些函数:

```
msg = "hello"  
msg.capitalize()
```

...