

Strings em Python

Além de números, Python também pode manipular strings (sequências de caracteres), que podem ser expressas de diversas formas. Elas podem ser delimitadas por aspas simples ou duplas:

```
>>> 'spam eggs'
'spam eggs'
>>> 'doesn't'
"doesn't"
>>> "doesn't"
"doesn't"
>>> 'Yes, he said.'
```

'Yes, he said.'
Strings podem ser acessados como listas, mas eles são imutáveis, não podem ser alterados. Para alterar letras em strings, deve-se antes transformar o string em lista. Veja a ida e a volta

```
>>> L = list("Curitiba")
>>> L[0]='X'
>>> print(L)
['X', 'u', 'r', 'i', 't', 'i', 'b', 'a']
>>> s=''.join(L)
>>> s
'Xuritiba'
>>> print(s)
Xuritiba
```

Como se viu a função list transforma cada caractere da string em um elemento da lista retornada. O método join faz o inverso, transformando uma lista em um string.

O interpretador exibe o resultado de operações com strings da mesma forma como elas são digitadas na entrada: dentro de aspas, e com aspas, caracteres acentuados e outros caracteres especiais representados por sequências de escape com barras invertidas (como '\t', '\xc3\xa9' etc.), para mostrar o valor preciso. A string é delimitada entre aspas simples, exceto quando ela contém uma aspa simples e nenhuma aspa dupla. O comando print produz uma saída mais legível para tais strings com caracteres especiais.

Strings que contém mais de uma linha podem ser construídas de diversas maneiras. Linhas de continuação podem ser usadas, com uma barra invertida colocada na última posição para indicar que a próxima linha física é a continuação de uma linha lógica:

```
oi = "Eis uma string longa contendo\n\ndiversas linhas assim como se faria em C.\n"
```

Não existe um tipo específico para caracteres; um caractere é simplesmente uma string de tamanho 1. Substrings podem ser especificadas através da notação de slice (fatiamento ou intervalo): dois índices separados por dois pontos. É exatamente o mesmo conceito estudado lá nas listas.

```
>>> palavra[4]
'a'
>>> palavra[0:2]
'Aj'
>>> palavra[2:4]
'ud'
```

Diferentemente de C, strings em Python não podem ser alteradas. Tentar atribuir valor a uma posição (índice ou fatia) dentro de uma string resulta em erro:

```
>>> palavra[0] = 'x'
=== Erro ===
>>> palavra[:1] = 'Splat'
=== Erro ===
```

Entretanto, criar uma nova string combinando conteúdos é fácil e eficiente:

```
>>> 'x' + palavra[1:]
'xjudaZ'
>>> 'Splat' + palavra[5]
'SplatZ'
```

Eis uma invariante interessante das operações de fatiamento: s[:i] + s[i] é igual a s.

Intervalos fora de limites: um índice maior que o comprimento é trocado pelo comprimento, um limite superior menor que o limite inferior produz uma string vazia.

```
>>> palavra[1:100]
'judaZ'
>>> palavra[10:]
''
>>> palavra[2:1]
```

Indices podem ser números negativos, para iniciar a contagem pela direita. Observe que -0 é o mesmo que 0, logo neste caso não se conta a partir da direita!

```
>>> palavra[-0]
'A'
```

A função built-in (embutida) len() devolve o comprimento de uma string:

```
>>> s = 'anticonstitucionalissimamente'
>>> len(s)
29
```

Métodos de String

Tanto strings comuns quanto Unicode oferecem um grande número de métodos para busca e transformações básicas. Eis alguns

.capitalize() Transforma o primeiro caractere de uma string em maiúsculo. Ignora caracteres acentuados ou que não sejam letras.

```
>>> "curitiba".capitalize()
'Curitiba'
>>> "ágora".capitalize()
ágora
```

.center() Centraliza a string usando espaços à esquerda e à direita, até atingir a largura.

```
>>> "paraná".center(20)
'      paraná      '
>>> "paraná".center(2)
'paraná'
```

.count() Conta a quantidade de substrings dentro da string

```
>>> "abracadabra".count("ab")
2
>>> "abracadabra".count("a")
5
```

.find() Retorna a primeira ocorrência do substring dentro da string. Caso nada seja encontrado, retorna -1.

```
>>> "abracadabra".count("ab")
2
>>> "abracadabra".count("a")
5
```

Avaliadores São eles: .isalpha(), .isdecimal(), .isdigit(), .islower(), .isnumeric(), .isspace(), .istitle(), .isupper(). Retornam True caso a string seja composta exclusivamente de alfanuméricos, decimais, dígitos, minúsculos, numéricos, espaços, títulos (a primeira letra de cada palavra em maiúsculo) e maiúsculas. Todos ignoram as letras acentuadas.

.join() Concatena todos os elementos dentro de iterável usando a string especificada como concatenador. A string pode ser vazia. Os elementos do iterável devem ser string.

```
>>> a=['agua', 'fogo', 'xabu']
>>> ''.join(a)
'aguafogoxabu'
>>> '\n'.join(a)
'agua\nfogo\nxabu'
>>> print('\n'.join(a))
agua
fogo
xabu
```

.partition() Divide a string em uma tupla com 3 elementos: a parte anterior à primeira ocorrência do separador, o separador e a parte posterior ao separador

```
>>> "curitibaparanabrasil".partition('ra')
('curitibapa', 'ra', 'nabrasil')
>>> "curitibaparanabrasil".partition('z')
('curitibaparanabrasil', '', '')
```

.replace() Substitui n ocorrências da substring antiga pela nova. Se n não é fornecido, ele substitui todas.

```
>>> "curitibaparanabrasil".replace("ra", "agora")
'curitibapaganabagorasil'
>>> "curitibaparanabrasil".replace("ra", "zz", 1)
'curitibapazznabrasil'
```

.split() Divide a string transformando-a em uma lista usando um separador n vezes. Se o separador não é informado ele usa um espaço em branco.

```
>>> "curitibaparanabrasil".split('a')
['curitib', 'p', 'r', 'n', 'br', 'sil']
>>> "curitiba parana brasil".split()
['curitiba', 'parana', 'brasil']
```

Muitos outros Aqui estão só alguns, talvez os mais úteis e/ou principais. Consulte a documentação da versão de Python que você está usando. Estão todos lá.

Formatação

Além da formatação padrão C (com %d, %s, %f, etc), Python introduziu a formatação representando os parâmetros por números em chaves. Veja o exemplo

```
>>> "{0} traga {1} aqui".format("Ze", 4)
'Ze traga 4 aqui'
```

Graças a esta facilidade, pode-se repetir o mesmo valor na frase e reordenar os parâmetros. Eis algumas coisas que se podem fazer

```
estabelecer o tamanho          {0:10}
alinhar à esquerda              {0:<10}
centralizar no espaço          {0:^10}
preencher o espaço com .      {0:.<10}
número: tamanho 5 com zeros   {0:05}
idem: 3 casas após a vírgula  {0:10.3}
com + ou -                     {0:-10}{1:+3}
em binário                     {0:b}
em hexadecimal*{0:x} {1:X}
```

Strings Unicode

Um novo tipo para armazenar textos foi introduzido: o tipo unicode. Ele pode ser usado para armazenar e manipular dados no padrão Unicode (veja http://www.unicode.org/) e se integra bem aos objetos string pré-existent, realizando conversões automáticas quando necessário.

Unicode tem a vantagem de associar um único número ordinal a cada caractere, para todas as formas de escrita usadas em textos modernos ou antigos. Anteriormente, havia somente 256 números ordinais possíveis para identificar caracteres. Cada texto era tipicamente limitado a uma "code page" (uma tabela de códigos) que associava ordinais aos caracteres. Isso levou a muita confusão especialmente no âmbito da internacionalização de software. Unicode resolve esses problemas ao definir uma única tabela de códigos para todos os conjuntos de caracteres.

Criar strings Unicode em Python é tão simples quanto criar strings normais:

```
>>> u'Hello World !'
u'Hello World !'
```

O u antes das aspas indica a criação de uma string Unicode. Se você desejar incluir caracteres especiais na string, você pode fazê-lo através da codificação Unicode-Escape de Python. O exemplo a seguir mostra como:

```
>>> u'Hello\u0020World !'
u'Hello World !'
```

O código de escape \u0020 insere um caractere Unicode com valor ordinal 0x0020 (o espaço em branco) naquela posição.

Para você fazer

- 1. Considere ED="OQBRMY"
Efetue todos os fatiamentos.
a) ED[3:4]
b) ED[3:5:2]
c) ED[1:4]
d) ED[:6]
e) ED[2:6:2]

- 2. Considere ZC="ZKTFNVPM"
Efetue todos os fatiamentos.
a) ZC[1:7]
b) ZC[-7:6:2]
c) ZC[-8:8:2]
d) ZC[-7:7:2]
e) ZC[2:-3:2]

Responda aqui:

Table with 5 columns (1a-1e, 2a-2e) and 3 rows.

