

Semana 09

Strings

CICOOO4 Algoritmos e Programação de Computadores

Prof. Pedro Garcia Freitas

https://pedrogarcia.gitlab.io/

pedro.garcia@unb.br

Brasilia

Este conjunto de slides não deve ser utilizado ou republicado sem a expressa permissão do autor.

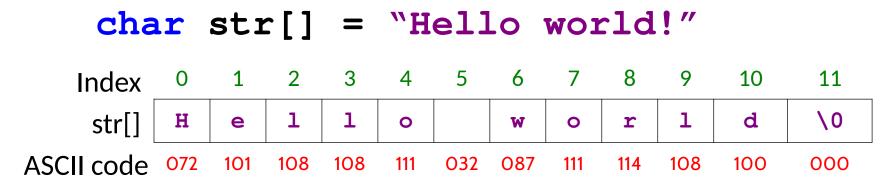
This set of slides should not be used or republished without the author's express permission.

1. Objetivos

Esta aula introduz o conceito de string (vetor de caracteres) e apresenta como tal conceito é implementado e explorado em linguagem Python.

• Em programação de computadores, uma string é tradicionalmente uma sequência de caracteres, seja como uma constante literal ou como algum tipo de variável.

 Em programação de computadores, uma string é tradicionalmente uma sequência de caracteres, seja como uma constante literal ou como algum tipo de variável.



 Variáveis de string podem permitir que seus elementos sejam modificados e o comprimento alterado, ou pode ser fixa (algumas linguagens não implementam strings com side effects).

- Uma string é usada para armazenar e manipular textos como palavras, nomes e sentenças.
- Quando uma string aparece literalmente no código-fonte, ela é conhecida como uma string literal ou uma string anônima.

```
# character literal in single quote
v = 'n'
# character literal in double quotes
w = "a"
variable_str = input()
print(v)
print(w)
print(w)
print(wariable_str)
print("My literal string")
```

```
character literal in single quote
    'n'
  character literal in double quotes
W
variable_str = input()
print(v)
                                   String literais
print(w)
print (variable_str)
print("My literal string")
```

- Um dos principais objetivos das strings é armazenar texto legível por humanos, como palavras e frases.
- As strings são usadas para comunicar informações de um programa de computador ao usuário do programa.

- Um programa também pode receber entradas de strings do usuário.
- Além disso, as strings podem armazenar dados expressos como caracteres, mas que não são necessariamente destinados à leitura humana.

- Exemplo de strings e seus propósitos:
 - A string "upload de arquivo concluído" pode ser usada no contexto de um software que exibe o status de um programa para os usuários finais.
 - No código-fonte do programa, essa mensagem provavelmente apareceria como uma string literal.

- Exemplo de strings e seus propósitos:
 - Um texto inserido pelo usuário, como "Consegui um novo emprego hoje", como uma atualização de status em um serviço de mídia social.
 - Em vez de uma *string* literal, o software provavelmente armazenaria essa *string* em um banco de dados.

- Exemplo de strings e seus propósitos:
 - Dados alfabéticos, como "AGATGCCGT", representando sequências de ácido nucleico do DNA.
 - Eis um exemplo de como uma string pode ser usada para conter uma informação sem necessariamente se destinar à leitura humana.



4. Histórico

• O uso da palavra "string" para significar "uma sequência de símbolos ou elementos linguísticos em uma ordem definida" surgiu da matemática, lógica simbólica e teoria linguística para falar sobre o comportamento formal de sistemas simbólicos, deixando de lado o significado dos símbolos.

- Um *string datatype* é um tipo de dado que modela uma string.
- As strings são um tipo de dado tão importante e útil que estão implementadas em praticamente todas as linguagens de programação.

- Em algumas linguagens, elas estão disponíveis como **tipos primitivos** e, em outras, como tipos compostos.
- A sintaxe da maioria das linguagens de programação de alto nível permite que uma string represente uma instância de um tipo de dado textual.

5.1 Comprimento de uma string:

 Embora as strings formais possam ter um comprimento finito arbitrário, o comprimento das strings em linguagens reais muitas vezes é limitado a um máximo artificial.

5.1 Comprimento de uma string:

• Em geral, existem dois tipos de tipos de dado string: strings de comprimento fixo, e strings de comprimento variável, cujo comprimento não é arbitrariamente fixo e que podem utilizar quantidades variáveis de memória dependendo dos requisitos reais em tempo de execução.

5.1 Comprimento de uma string:

- A maioria das strings nas linguagens de programação modernas são strings de comprimento variável.
- Claro, mesmo as strings de comprimento variável têm um limite de comprimento: o tamanho da memória do computador disponível.

5.1 Comprimento de uma string:

• O comprimento da *string* pode ser armazenado como um **inteiro separado** (o que pode impor um limite artificial ao comprimento) ou **implicitamente por meio de um caractere de terminação**, geralmente um valor de caractere com todos os bits iguais a zero.

5.1 Comprimento de uma string:

```
class String:
```

text: list = []

size: int = 0

Inteiro separado indicando a quantidade de caracteres

```
char str[] = "Hello world!\0"
```

Caractere de terminação

```
typedef struct {
    PyObject_VAR_HEAD
    long ob_shash;
    int ob_sstate;
    char ob_sval[1];
```

PyStringObject;

Como o interpretador oficial do Python (CPython) implementa o tipo string.

Fonte: https://svn.python.org/projects/python/trunk/Include/stringobject.h

5.2 Codificação de caracteres (character encoding)

- Os tipos de dados de string historicamente alocavam um byte por caractere.
- Esses conjuntos de caracteres eram tipicamente baseados em ASCII ou EBCDIC.

5.2 Codificação de caracteres (character encoding)

 Se um texto gerado num sistema de codificação for exibido noutro sistema que usa uma codificação diferente, o texto geralmente é apresentado distorcido.

Fonte: https://www.rapidtables.com/code/text/ascii-table.html



Universidade de Brasília

Departamento de Ciências da Computação

5. String datatypes

0	00	NUL
1	01	SOH
2	02	STX
3	03	ETX
4	04	EOT
5	05	ENQ
6	06	ACK
7	07	BEL
8	80	BS
9	09	HT
10	0A	LF
11	0B	VT
12	0C	FF
13	0D	CR
14	0E	SO
15	0F	SI
16	10	DLE
17	11	DC1
18	12	DC2
19	13	DC3
20	14	DC4
21	15	NAK
22	16	SYN
23	17	ETB
24	18	CAN

```
25
    19
        EM
    1A SUB
26
27
    1B ESC
28
    1C FS
29
    1D GS
30
    1E RS
31
    1F US
32
    20 space
33
    21
34
    22
35
    23 #
36
    24
37
    25
38
    26
39
    27
40
    28
41
    29
42
    2A
43
    2B +
44
    2C
45
    2D -
    2E
46
47
    2F
48
    30 0
49
    31
```

-	54	22	2
	51	33	3
	52	34	4
	53	35	5
	54	36	6
,	55	37	7
	56	38	8
	57	39	9
	58	3A	:
	59	3B	
(60	3C	<
(61	3D	; < = >
(52	3E	>
(63	3F	?
(54	40	@
(65	41	Α
(66	42	В
(67	43	С
(68	44	D
(69	45	E
7	70	46	F
	71	47	G
7	72	48	Н
7	73	49	I
7	74	4A	J
	75	4B	K
7	76	4C	L

77	4D	М
78	4E	N
79	4F	0
80	50	Р
81	51	Q
82	52	R
83	53	S
84	54	T
85	55	U
86	56	V
87	57	W
88	58	X
89	59	Υ
90	5A	Z
91	5B	[
92	5C	\
93	5D]
94	5E	٨
95	5F	_
96	60	•
97	61	a
98	62	b
99	63	С
100	64	d
101	65	е
102	66	f

103	67	g
104	68	h
105	69	i
106	6A	j
107	6B	k
108	6C	1
109	6D	m
110	6E	n
111	6F	0
112	70	p
113	71	q
114	72	r
115	73	S
116	74	t
117	75	u
118	76	V
119	77	W
120	78	X
121	79	y
122	7A	Z
123	7B	{
124	7C	1
125	7D	}
126	7E	~
127	7F	DEL

5.3 Mutabilidade de *strings*

 Algumas linguagens, como C++, Perl e Ruby, normalmente permitem que o conteúdo de uma string seja alterado após sua criação; essas são chamadas de strings mutáveis.

5.3 Mutabilidade de *strings*

• Em outras linguagens, como Java, JavaScript, Lua, Python e Go, o valor é fixo e uma nova string é criada se alguma alteração for feita; essas são chamadas de strings imutáveis.

5.3 Mutabilidade de strings

 Algumas dessas linguagens com strings imutáveis também fornecem outro tipo que é mutável, como o StringBuilder do Java e .NET, o Java StringBuffer,.

5.3 Mutabilidade de strings

 Existem vantagens e desvantagens na imutabilidade: embora as strings imutáveis demandam mais memória (cada operação gera uma nova cópia), elas são mais simples e completamente seguras em termos de concorrência.

<u>6.1 Declaração de strings</u>

```
# create a string using double quotes
string1 = "Python programming"

# create a string using single quotes
string2 = 'Python programming'

# create a string using triple quotes
string3 = '''Both ' and " available!'''
```

<u>6.1 Declaração de strings</u>

```
exemplo01.py
    # create string type variables
  3 # create a string using double quotes
    s1 = "Double quotes string"
    # create a string using single quotes
    s2 = 'Single quotes string'
    # create a string using triple quotes
    s3 = '''Both ' and " available!'''
 11
 12 for s in (s1,s2,s3):
 13
        print(s)
```

6.1 Declaração de *strings*

```
>>> %Run exemplo01.py

Double quotes string
Single quotes string
Both ' and " available!
>>>
```

6.2 Acesso de caracteres pertencentes à string

Podemos acessar os caracteres em uma string de três maneiras:

- Indexação
- Indexação negativa
- Slicing

6.2 Acesso de caracteres pertencentes à string

Indexação (indexing): Uma maneira é tratar as strings como uma lista e usar valores de índice.

```
greet = 'hello'
# access 2nd index element
print(greet[1]) # "e"
```

6.2 Acesso de caracteres pertencentes à string

Indexação negativa (negative indexing): Semelhante a uma lista, o Python permite a indexação negativa para suas strings. Por exemplo,

```
greet = 'hello'
# access 4th element reversely
print(greet[-4]) # "e"
```

6.2 Acesso de caracteres pertencentes à string

Indexação negativa (negative indexing): Semelhante a uma lista, o Python permite a indexação negativa para suas strings. Por exemplo,

```
greet = 'hello'
# access 1th element reversely
print(greet[-1]) # "o"
```

6.2 Acesso de caracteres pertencentes à string

Slicing: Python permite o acesso a um intervalo de caracteres em uma string usando o "operador de fatiamento" ":"

```
greet = 'hello'
# access 1th element reversely
print(greet[1:4]) # "ell"
```

6.2 Acesso de caracteres pertencentes à string



6.3 Mutabilidade/Imutabilidade de strings

 Em Python, as strings são imutáveis.
 Isso significa que os caracteres de uma string não podem ser alterados. Por exemplo,

```
message = 'Hola Amigos'
message[0] = 'X'
```

6.3 Mutabilidade/Imutabilidade de strings

```
TypeError: 'str' object does not support item assignment res de uma dos. Por ex o,

essage = 'Hola Amigos'
message[0] = 'X'
```

6.3 Mutabilidade/Imutabilidade de strings

 Portanto, toda operação feita com strings como operandos produz uma nova string.

```
message = 'Hola Amigos'
message = 'X' + message[1:]
print(message)
```

6.3 Mutabilidade/Imutabilidade de strings

```
exemplo02.py x

1  message = 'Hola Amigos'
2  message = 'X' + message[1:]
3  print(message)
```

```
>>> %Run exemplo02.py

Xola Amigos
>>>
```

6.4 Multiline strings

Também é possível criar uma *string* multilinha em Python. Para isso, utilizamos três aspas duplas """ ou três aspas simples "".

6.4 Multiline strings

```
# multiline string
message = """
Never gonna give you up
Never gonna let you down
"""
print (message)
```

6.4 Multiline strings

```
exemplo03.py ×

1  # multiline string
2  message = """
3  Never gonna give you up
4  Never gonna let you down
5  """
6
7  print(message)
```

```
>>> %Run exemplo03.py

Never gonna give you up
Never gonna let you down
>>>
```

6.5 Comparação de strings

Utilizamos o operador == para comparar duas strings. Se as duas strings forem iguais, o operador retorna True. Caso contrário, retorna False. Por exemplo,

6.5 Comparação de strings

```
Shell x

>>> %Run exemplo04.py
  'Hello, world!' = 'I love Python.'? False
  'Hello, world!' = 'Hello, world!'? True
>>>
```

6.6 Junção de strings (concatenação)

Em Python, podemos juntar (concatenar) duas ou mais *strings* usando o operador +.

```
greet = "Hello, "
name = "Jack"
# using + operator
result = greet + name
print(result)
```

6.6 Junção de strings (concatenação)

```
greet = "Hello, "
name = "Jack"
    # using + operator
result = greet + name
print(result)
```

```
>>> %Run exemplo05.py
Hello, Jack
>>>
```

6.7 Iterar através de uma string em Python

Podemos iterar através de uma string usando um loop **for** ou **while** em Python. Por exemplo,

```
greet = 'Hello'
# iterating through greet string
for letter in greet:
    print(letter)
```

6.7 Iterar através de uma string em Python

```
greet = 'Hello'

for letter in greet:
    print(letter)
```

```
Shell ×

>>> %Run exemplo06.py

H
e
1
1
0
>>>
```

6.8 Comprimento/tamanho de uma string

Em Python, usamos o método len () para encontrar o comprimento de uma string. Por exemplo,

```
greet = 'Hello'
# count length of greet string
print(len(greet))
```

6.8 Comprimento/tamanho de uma string

```
exemplo07.py ×

1  greet = 'Hello'
2
3  # count length of greet string
4  print(len(greet))
```

```
Shell ×
>>> %Run exemplo07.py
5
>>>
```

6.9 Teste da existência de uma substring

Podemos testar se uma substring existe ou não em uma string usando a palavra-chave in.

```
print('a' in 'program') # True
print('at' not in 'battle') # False
```

6.9 Teste da existência de uma substring

```
exemplo08.py ×

1 print('a' in 'program')
2 print('at' not in 'battle')
```

```
>>> %Run exemplo08.py
True
False
>>>
```

6.10 Busca de *substring*

- O método **find** () retorna o índice da primeira ocorrência da substring (se encontrada). Se não for encontrada, **retorna -1**.
- O método index () retorna o índice de uma substring dentro da string (se encontrada). Se a substring não for encontrada, ele gera uma exceção.

6.10 Busca de *substring*

```
exemplo09.py x

1  string = "This string has a substring!"
2  3  print(string.find("substring"))
4  print(string.index("substring"))
```

6.10 Busca de *substring*

```
exemplo10.py x

1   string = "This string has a substring!"
2   substr = "not_exist"
3   print("find:", string.find(substr))
4   print("index:|", string.index(substr))
```

```
>>> %Run exemplo10.py
find: -1
Traceback (most recent call last):
   File "/home/pedro/Dropbox/UnB/APC/2023.1/T05/meus
   slides/Aula09/exemplo10.py", line 4, in <module>
        print("index:", string.index(substr))
ValueError: substring not found
>>>
```

6.11 Substituição de *substring*

O método replace () substitui cada ocorrência correspondente de uma substring por outra string. Sintaxe:

str.replace(old, new [, count])



Universidade de Brasília

Departamento de Ciências da Computação

- old a antiga substring que queremos substituir
- new a nova substring que substituirá a antiga substring
- count (opcional) o número de vezes que você deseja substituir a antiga substring pela nova string

Nota: Se o parâmetro count não for especificado, o método **replace**() substituirá todas as ocorrências da antiga substring pela nova string.

uma son outra string. Sintaxe:

str.replace(old, new [, count])

6.11 Substituição de *substring*

```
exemplo12.py x

1   string = "Multi str str str!"
2   old = "str"
3   new = "NEW"
4   print("old:", string)
5   print("new:", string.replace(old, new))
```

```
>>> %Run exemplo12.py
old: Multi str str str!
new: Multi NEW NEW NEW!
>>>
```

6.12 Conversão capitular

- O método upper () converte todos os caracteres minúsculos em uma string em caracteres maiúsculos e retorna a nova string.
- O método lower () converte todos os caracteres maiúsculos em uma string em caracteres minúsculos e retorna a nova string.

6.12 Conversão capitular

```
exemplo13.py x

1  title = "This is a Title Case String!"
2  low = title.lower()
3  up = title.upper()
4  print("title=", title|)
5  print("low=", low)
6  print("up=", up)
```

```
>>> %Run exemplo13.py

title -> This is a Title Case String!
low -> this is a title case string!
up -> THIS IS A TITLE CASE STRING!
>>>
```

6.13 Particionamento de *strings*

- O método partition () divide a string na primeira ocorrência da string de argumento e retorna uma tupla contendo a parte antes do separador, a string de argumento e a parte após o separador.
- O método split () divide uma string no separador especificado e retorna uma lista de substrings.

6.13 Particionamento de strings

```
cars = 'BMW-Tesla-Range Rover'

print("cars -> ", cars)
print("split -> ", cars.split('-'))
print("partition -> ", cars.partition('-'))
```

```
>>> %Run exemplo14.py

cars -> BMW-Tesla-Range Rover
split -> ['BMW', 'Tesla', 'Range Rover']
partition -> ('BMW', '-', 'Tesla-Range Rover')
>>>
```

6.14 Format strings

Existem pelo menos 3 formas de formatar strings em Python:

- Old-fashion style
- Método String.format ()
- f-strings

6.14 Format strings

```
Old '%s %s' % ('one', 'two')

New '{} '.format('one', 'two')

Output one two
```

6.14 Format strings

```
Old '%d %d' % (1, 2)

New '{} '.format(1, 2)

Output 1 2
```

6.14 Format strings

This operation is not available with old-style formatting.

```
New '{1} {0}'.format('one', 'two')
```

Output

t w o o n e

6.14 Format strings

6.14 Format strings

```
Old
    '%-10s' % ('test',)

New
    '{:10}'.format('test')

Output
    t e s t
```

6.14 Format strings

This operation is not available with old-style formatting.

New

Output

```
t e s t _ _ _ _ _ _
```

6.14 Format strings

This operation is not available with old-style formatting.

```
New '{:^10}'.format('test')
```

Output t e s t

6.14 Format strings

Output

'{:d}'.format(42)

6.14 Format strings

```
Old '%f' % (3.141592653589793,)
```

New '{:f}'.format(3.141592653589793)

Output 3 . 1 4 1 5 9 3

6.14 Format strings

6.14 Format strings

```
Old '%06.2f' % (3.141592653589793,)
```

```
New '{:06.2f}'.format(3.141592653589793)
```

```
Output 0 0 3 . 1 4
```

6.14 Format strings

```
data = {'first': 'Hodor', 'last': 'Hodor!'}
Old
          '%(first)s %(last)s' % data
New
          '{first} {last}'.format(**data)
Output
          Hodor Hodor!
```

6.14 Format strings

```
New
    '{first} {last}'.format(first='Hodor', last='Hodor!')
Output
    H o d o r H o d o r!
```



7. Exercício

Escreva uma função que recebe como entrada um <nome> e imprime na tela do computador "Hola, <nome>" emoldurada conforme ilustrado abaixo. A largura da moldura deve adaptar-se de acordo com o tamanho da string de entrada. Por exemplo, se o nome for "Ivo", a largura do quadro impresso na tela será diferente da largura do quadro impresso para o nome "Maersalalhasbaz".

+----+ | | | | Hola, Ivo | | | | 

Universidade de Brasília

Departamento de Ciências da Computação



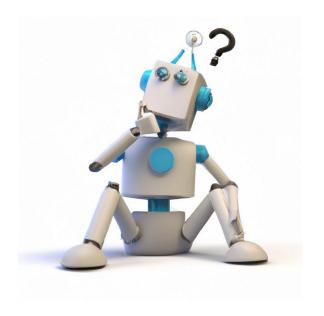
Prof. Pedro Garcia Freitas

https://pedrogarcia.gitlab.io/

pedro.garcia@unb.br



Dúvidas?



Prof. Pedro Garcia Freitas

https://pedrogarcia.gitlab.io/

pedro.garcia@unb.br