



BÀI 7

String in Python Kiểu Chuỗi trong Python

Tóm Tắt Nội Dung

Trong bài học này chúng ta sẽ đi tìm hiểu lần lượt các nội dung

- 1 Tổng quan về Chuỗi (String)
- 2 Các phương thức xử lý với Chuỗi
- 3 Định dạng chuỗi với khuôn
- 4 Dãy thoát Escape trong Python

7.1 Tổng quan về Chuỗi (String)



Giới thiệu về Kiểu dữ liệu String

- ◆ Trong Python, kiểu dữ liệu chuỗi (string) là một dạng dữ liệu dùng để lưu trữ văn bản, chữ số, ký tự, hoặc bất kỳ dãy các ký tự nào
- ◆ Chuỗi trong Python được biểu diễn bằng các ký tự nằm trong cặp dấu nháy đơn ('...') hoặc nháy kép ("..."). Python hỗ trợ các thao tác và phương thức mạnh mẽ để làm việc với kiểu dữ liệu chuỗi

```
# Khai báo chuỗi bằng cặp dấu nháy đơn
string1 = 'Hello, World!'
# Khai báo chuỗi bằng cặp dấu nháy kép
string2 = "Python is awesome"
# Chuỗi có thể chứa các ký tự đặc biệt
string3 = 'Các ký tự đặc biệt: !@#$%^&*()'
# Chuỗi có thể chứa các số và chữ số
string4 = '12345'
```

7.1 Tổng quan về Chuỗi (String)

- ◆ Chuỗi nhiều dòng với 3 kí tự nháy kép `"""`

```
a = """Lorem ipsum dolor sit amet,  
consectetur adipiscing elit,  
sed do eiusmod tempor incididunt  
ut labore et dolore magna aliqua."""  
print(a)
```

- ◆ Hoặc Chuỗi nhiều dòng với 3 kí tự nháy đơn `'''`

```
a = '''Lorem ipsum dolor sit amet,  
consectetur adipiscing elit,  
sed do eiusmod tempor incididunt  
ut labore et dolore magna aliqua.'''  
print(a)
```

7.1 Tổng quan về Chuỗi (String)

- ◆ Trong Python String cũng được coi là một Mảng (Array) tập hợp của nhiều kí tự, vị trí sắp xếp bắt đầu là số 0

```
a = "Hello, World!"  
print(a[1]) #output: e
```

- ◆ Vì là mảng nên bạn có thể lặp với vòng lặp **for**

```
for x in "banana":  
    print(x)
```

- ◆ Dùng phương thức len() để kiểm tra độ dài của chuỗi có bao nhiêu kí tự

```
a = "Hello, World!"  
print(len(a)) #output: 13
```

7.1 Tổng quan về Chuỗi (String)

- ◆ Kiểm tra sự tồn tại của một kí tự trong chuỗi với từ khóa **in**

```
txt = "The best things in life are free!"  
print("free" in txt)  
  
if "free" in txt:  
    print("Có tồn tại")
```

- ◆ Hoặc kết hợp với **not** để kiểm tra không tồn tại

```
if "fresh" not in txt:  
    print("Đúng, fresh không tồn tại trong chuỗi")
```

7.2 Các phương thức xử lí Chuỗi



Cắt chuỗi

Bạn có thể tách chuỗi dựa vào vị trí index của nó trong chuỗi

Cú pháp: **string[start:end]**

- start mặc định là 0, có thể là số nguyên âm
- end mặc định là đến cuối chuỗi, có thể là số nguyên âm

```
b = "Hello, World!"  
print(b[2:5]) #output: llo
```

Ví dụ này sẽ lấy kí tự từ vị trí thứ 2 đến 4 (5 – 1, không bao gồm vị trí số 5)

```
print(b[-2:-5]) #output: orl
```

Dùng số nguyên âm để đảo chiều vị trí lấy

7.2 Các phương thức xử lí Chuỗi



Cắt chuỗi

Ví dụ không điền **start**, thì start mặc định là số 0

```
b = "Hello, World!"  
print(b[0:5]) #output: Hello
```

Ví dụ không điền **end**, thì start mặc định là đến vị trí cuối

```
print(b[2:]) #output: llo, World!
```


7.2 Các phương thức xử lí Chuỗi



Biến đổi chuỗi

- ◆ Biến thành chuỗi viết HOA

```
a = "Hello, World!"  
print(a.upper()) #output: HELLO, WORLD!
```

- ◆ Biến thành chuỗi viết thường

```
print(a.lower()) #output: hello, world!
```

- ◆ Kí tự đầu tiên mỗi chữ viết HOA

```
txt = "Welcome to my world"  
print(txt.title()) #output: Welcome To My World
```

7.2 Các phương thức xử lí Chuỗi



Biến đổi chuỗi

- ◆ Tách chuỗi với **split()** dựa vào một kí tự nằm giữa chuỗi

```
a = "Hello, World!"  
print(a.split(",")) #returns list: ['Hello', ' World!']
```

- ◆ Xóa khoảng trắng 2 đầu chuỗi với **strip()**

```
a = " Hello, World! "  
print(a.strip()) # returns "Hello, World!"
```

- ◆ Thay thế kí tự trong chuỗi với **replace()**

```
txt = "Welcome to my world"  
print(txt.replace("world", "home")) #output: Welcome To My home
```

7.3 Định dạng Chuỗi



Định dạng (đóng khuôn) chuỗi

- ◆ Trong python, không thể nối chuỗi với số với toán tử + cộng, nhưng chúng ta có thể làm điều đó với hàm **format()**

```
age = 36
txt = "Tôi là John, Tôi {} tuổi"
print(txt.format(age))
#Kết quả in ra: Tôi là John, Tôi 36 tuổi
```

Trong ví dụ trên kí tự **{}** đóng vai trò giữ chỗ cho biến age, và age sẽ hiển thị ra thay cho **{}**. Các vị trí **{}** này được gọi là **trường** (hay là field)

7.3 Định dạng Chuỗi



Định dạng (đóng khuôn) chuỗi

- ◆ Thay thế đơn giản không chỉ số

```
age = 36
name = 'Jonh'
txt = "Tôi là {}, Tôi {} tuổi"
print(txt.format(name, age))
#Kết quả in ra: Tôi là John, Tôi 36 tuổi
```

- Khi đó biến **name** sẽ thay thế vào kí tự {} đầu tiên trong chuỗi txt
- Và biến **age** sẽ thay thế vào vị trí {} thứ 2 trong chuỗi txt
- Nếu có một biến khác thì biến và {} được tham chiếu một cách tuần tự như vậy khi bạn thêm vào hàm **format()**

7.3 Định dạng Chuỗi



Định dạng (đóng khuôn) chuỗi

◆ Thay thế đơn giản không chỉ số

```
age = 36
name = 'Jonh'
txt = "Tôi là {}, Tôi {} tuổi"
print(txt.format(name, age))
#Kết quả in ra: Tôi là John, Tôi 36 tuổi
```

- Khi đó biến **name** sẽ thay thế vào kí tự {} đầu tiên trong chuỗi txt
- Và biến **age** sẽ thay thế vào vị trí {} thứ 2 trong chuỗi txt
- Nếu có một biến khác thì biến và {} được tham chiếu một cách tuần tự như vậy khi bạn thêm vào hàm **format()**

Lưu ý: bạn có thể đưa nhiều tham số biến hơn số lượng {} cần thay thế

7.3 Định dạng Chuỗi



Định dạng (đóng khuôn) chuỗi

◆ Thay thế có chỉ số

```
quantity = 2
price = 47.95
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"
print(txt.format(quantity, price))
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

Với cách dùng chỉ số này thì nhìn vào dễ hiểu hơn:

- **quantity**: sẽ tương ứng với field số {0}

- **price**: sẽ tương ứng với field số {1}

Và mình có thể đặt các field bất kỳ ở đâu chỉ cần ghi số vào là được.

Lưu ý: nếu đã đánh số thì tất cả các field đều đánh số, chứ không được cái thì {}, cái thì {0}. Python sẽ không xử lý đúng.

7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh thiết lập chiều rộng Width

```
quantity = 2
price = 47.95
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"
print(txt.format(quantity, price))
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh thiết lập canh thẳng hàng

```
quantity = 2
price = 47.95
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"
print(txt.format(quantity, price))
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```


7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh chèn kí tự vào khoảng trống

```
quantity = 2
price = 47.95
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"
print(txt.format(quantity, price))
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh thể hiện số gần đúng

```
quantity = 2  
price = 47.95  
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"  
print(txt.format(quantity, price))  
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh # (hash)

```
quantity = 2  
price = 47.95  
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"  
print(txt.format(quantity, price))  
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

7.3 Định dạng Chuỗi



Chỉ lệnh trong field

◆ Chỉ lệnh kiểu (type)

```
quantity = 2
price = 47.95
txt = "Mua {0} cái bánh hết {1} đô la. Em tôi ăn hết {0} cái bánh"
print(txt.format(quantity, price))
#Kết quả in ra: Mua 2 cái bánh hết 47.95 đô la. Em tôi ăn hết 2 cái bánh
```

7.4 Escape Character



Giới thiệu về khái niệm Escape

- ◆ Trong Python, ký tự escape (escape character) là một ký tự đặc biệt được sử dụng để thể hiện các ký tự không thể hiện hoặc có ý nghĩa đặc biệt trong chuỗi. Khi bạn sử dụng ký tự escape, nó cho phép bạn đưa vào chuỗi những ký tự đặc biệt như dấu nháy kép, dấu nháy đơn, dấu gạch chéo ngược, và các ký tự khác mà không bị xem là kết thúc của chuỗi hoặc gây ra sự hiểu lầm về cú pháp.

Ký tự escape được ký hiệu bằng một dấu gạch chéo ngược (") theo sau là ký tự đại diện cho ý nghĩa cụ thể. Dưới đây là một số ví dụ về ký tự escape phổ biến trong Python:

\n: Ký tự xuống dòng (newline)
\t: Ký tự tab ngang (tab)
": Dấu nháy kép
' : Dấu nháy đơn

7.4 Escape Character



Ví dụ về cách sử dụng Escape

Sử dụng ký tự escape để in ra các ký tự đặc biệt trong chuỗi

```
print("Dòng 1\nDòng 2")  
# In ra hai dòng cách nhau bởi dấu xuống dòng (\n)  
print("Tab:\t\tKý tự tiếp theo sau tab")  
# In ra ký tự tab (\t) để tạo khoảng cách  
print("Dấu nháy kép: \"Ví dụ\" ")  
# In ra dấu nháy kép (") trong chuỗi  
print('Dấu nháy đơn: \'m a student \'')  
# In ra dấu nháy đơn (') trong chuỗi  
print("Dấu gạch chéo ngược: \\")  
# In ra dấu gạch chéo ngược (\) trong chuỗi
```

Tham khảo link sau để tìm hiểu tất cả phương thức xử lý chuỗi:
https://www.w3schools.com/python/python_strings_methods.asp

7.5 Nguồn tham chiếu



Nguồn tham chiếu

- ◆ Chuỗi trong Python

https://www.w3schools.com/python/python_strings.asp

- ◆ Các phương thức Xử lý chuỗi trong Python

https://www.w3schools.com/python/python_strings_methods.asp

Tổng kết lại bài

- 1 Hiểu được thế nào gọi là dữ liệu kiểu Chuỗi (String)
- 2 Nắm được các phương thức xử lý với Chuỗi
- 3 Biết cách định dạng chuỗi với khuôn format()
- 4 Hiểu thế nào là dãy thoát Escape trong Python và cách sử dụng