



BÀI 24

Kiểu Xâu Ký Tự

GV: TRẦN ĐỨC HUY

Em đã biết kiểu dữ liệu chuỗi ký tự (gọi tắt là chuỗi) từ Bài 16 và chúng ta có thể tạo các biến kiểu chuỗi ký tự theo nhiều cách như sau:

```
>>> s = "Hoàng Thị Thanh Tâm"
```

```
>>> chuỗi = 'Hoa học trò'
```

```
>>> Cau_tho = ""Mình về mình có nhớ ta  
Mười lăm năm ấy thiết tha mặn nồng""
```

Liệu có lệnh nào trích ra từng ký tự của một chuỗi ký tự? Đếm số ký tự của một chuỗi?



1. Xâu là một dãy kí tự



Quan sát các ví dụ sau để biết cấu trúc xâu kí tự, so sánh với danh sách để biết sự khác nhau giữa xâu và danh sách?

```
>>> s = "Hoàng Thị Thanh Tâm"  
>>> len(s)  
  
19  
  
>>> s[0]  
  
'H'  
  
>>> s[10]  
  
'T'
```

- Một chuỗi ký tự được hiểu là một dãy các ký tự. Tương tự danh sách, ta có thể truy cập từng ký tự của chuỗi thông qua chỉ số, chỉ số bắt đầu từ 0.

Ví dụ 2. Quan sát các lệnh sau để thấy sự khác nhau giữa chuỗi và danh sách

```
>>> d=["t","â","m"]
>>> d[0]="T"
>>> s="tâm"
>>> s[0]="T"
```

```
Traceback (most recent call last):
```

```
  File "<pyshell>", line 1, in <module>
```

```
TypeError: 'str' object does not support item assignment
```

- Python không cho phép thay đổi từng kí tự của một xâu. Điều này khác với danh sách.
- Python không có kiểu dữ liệu kí tự. Kí tự chính là xâu có độ dài 1. Xâu rỗng được định nghĩa như sau: `empty = ""`

Ghi nhớ: Xâu kí tự trong Python là dãy các kí tự Unicode. Xâu có thể được coi là danh sách các kí tự nhưng không thay đổi từng kí tự của xâu. Truy cập từng kí tự của xâu qua chỉ số, chỉ số từ 0 đến độ dài `len()` – 1.

1. Các xâu kí tự sau có hợp lệ không?

a) “123&*()+-ABC”

b) “1010110&0101001”

c) “Tây Nguyên”

d) $11111111 = 256$

2. Mỗi xâu hợp lệ ở câu 1 có độ dài bằng bao nhiêu?





Quan sát các lệnh sau để biết cách duyệt từng kí tự của xâu kí tự bằng lệnh for. Có hai cách duyệt, theo chỉ số và theo phần tử của xâu kí tự.

```
>>> s="Hoàng Thị Thanh Tâm"  
>>> for i in range(len(s)):  
    print(s[i],end=" ")
```

H o à n g T h ị T h a n h T â m

```
>>> for ch in s:  
    print(ch,end=" ")
```

H o à n g T h ị T h a n h T â m

2. Lệnh duyệt kí tự của xâu

- Cách thứ nhất, biến i lần lượt chạy theo chỉ số của xâu kí tự s , từ 0 đến $\text{len}(s) - 1$. Kí tự tại chỉ số i là $s[i]$.
- Cách duyệt thứ hai duyệt theo từng kí tự của xâu s . Biến ch sẽ được gán lần lượt các kí tự của xâu s từ đầu đến cuối.

Chú ý: Từ khoá **in**, tùy trường hợp cụ thể, hoặc là toán tử logic dùng để kiểm tra một giá trị có mặt hay không trong một vùng giá trị/danh sách/xâu, hoặc để chọn lần lượt từng phần tử trong một vùng giá trị/danh sách/xâu.

```
>>> "a" in "abcd"
```

```
True
```

```
>>> "abc" in "abcd"
```

```
True
```

Ghi nhớ

- Có thể duyệt các kí tự của xâu bằng lệnh for tương tự với danh sách. s_1 in s_2 trả lại giá trị True nếu s_1 là xâu con của s_2



1. Sau khi thực hiện các lệnh sau, biến skq sẽ có giá trị bao nhiêu?

```
>>> s = "81723"
```

```
>>> skq = ""
```

```
>>> for ch in s:
```

```
    if int(ch) % 2 != 0:
```

```
        skq = skq + ch
```

2. Cho s1 = "abc", s2 = "ababcabca". Các biểu thức logic sau cho kết quả là đúng hay sai?

a) s1 in s2

b) s1 + s1 in s2

c) "abcabca" in s2

d) "abc123" in s2

THỰC HÀNH

Các lệnh cơ bản làm việc với xâu kí tự

Nhiệm vụ 1. Viết chương trình nhập số tự nhiên n là số học sinh, sau đó nhập họ và tên học sinh. Lưu họ và tên học sinh vào một danh sách. In danh sách ra màn hình, mỗi họ tên trên một dòng.

Hướng dẫn. Chương trình có thể như sau:

HoangThiThanhTam_TL.py ×

```
1 n = int(input("Nhập số học sinh trong lớp: "))
2 ds_lop = []
3 for i in range(n):
4     hoten = input("Nhập họ tên học sinh thứ " + str(i+1)+" :")
5     ds_lop.append(hoten)
6 print("Danh sách lớp học:")
7 for i in range(n):
8     print(ds_lop[i])
9
```

Nhiệm vụ 2. Nhập một xâu kí tự S từ bàn phím rồi kiểm tra xem xâu S có chứa xâu con “10” không.

Hướng dẫn. Cách 1. Nếu xâu S chứa xâu con “10” thì sẽ có chỉ số k mà $S[k] = \text{“1”}$ và $S[k+1] = \text{“0”}$. Cách 2. Dùng toán tử in để kiểm tra xâu “10” có là xâu con của S.

Cách 1: Duyệt kí tự của xâu theo chỉ số.

HoangThiThanhTam_TL.py ×

```
1 S = input("Nhập chuỗi kí tự bất kì: ")
2 kq = False
3 for i in range(len(S)-1):
4     if S[i] == "1" and S[i+1] == "0" :
5         kq = True
6         break
7 if kq:
8     print("Chuỗi gốc có chứa chuỗi '10'")
9 else:
10    print("Chuỗi gốc không chứa chuỗi '10'")
11
```

Cách 2: Sử dụng toán tử in.

HoangThiThanhTam_TL.py ×

```
1 S = input("Nhập xâu kí tự bất kì: ")
2 s10 = "10"
3 if s10 in S:
4     print("Xâu gốc có chứa xâu '10'")
5 else:
6     print("Xâu gốc không chứa xâu '10'")
7
```

LUYỆN TẬP

1. Cho xâu S, viết đoạn lệnh trích ra xâu con của S bao gồm ba kí tự đầu tiên của S.
2. Viết chương trình kiểm tra xâu S có chứa chữ số không. Thông báo “S có chứa chữ số” hoặc “S không chứa chữ số nào”.
3. Cho hai xâu s_1 , s_2 . Viết đoạn chương trình chèn xâu s_1 vào giữa s_2 , tại vị trí $\text{len}(s_2)//2$. In kết quả ra màn hình.
4. Viết chương trình nhập số học sinh và họ tên học sinh. Sau đó đếm xem trong danh sách có bao nhiêu bạn tên là “Hương”.

Gợi ý: Sử dụng toán tử **in** để kiểm tra một xâu có là xâu con của một xâu khác.

BÀI TẬP

Bài 1: Hãy dự đoán kết quả đưa ra màn hình sau mỗi câu lệnh xuất dữ liệu `print()` trong chương trình ở hình bên và sau đó dùng cửa sổ Shell để đối chiếu, kiểm tra từng kết quả dự đoán

```
*thay_the_xau.py - C:/Users/TAM/AppData/Local/Programs/Python/Pyt...
File Edit Format Run Options Window Help
xau1 = "Hà Nội là thủ đô của nước Việt Nam, "
xau2 = "Nam Khánh sinh ra ở Hà Nội."
xau = xau1 + xau2
print(xau) -----> ?
print(xau.count("N", 6)) -----> ?
print(xau.find("Khánh")) -----> ?
print(xau[4:9]) -----> ?
print(xau.replace("Khánh", "An")) ----> ?
|
```

Bài 2: Em hãy viết chương trình nhập từ bàn phím chuỗi ghi ngày tháng dạng dd/mm/yyyy, trong đó dd là hai ký tự chỉ ngày, mm là hai ký tự chỉ tháng, yyyy là bốn ký tự chỉ năm. Sau đó đưa ra màn hình ngày, tháng, năm dưới dạng chuỗi “Ngày dd tháng mm năm yyyy”.

Ví dụ:

Input	Output
15/12/2022	Ngày 15 tháng 12 năm 2022

Bài 3: Nhập vào từ bàn phím hai xâu s1 và s2, mỗi xâu không chứa kí tự dấu cách ở đầu và cuối xâu cũng như không chứa hai hay nhiều dấu cách liên tiếp nhau. Nếu xâu không chứa dấu cách thì nó là một từ, trong trường hợp ngược lại, dấu cách là dấu phân tách các từ trong xâu. Ví dụ, xâu “Bước tới Đèo Ngang, bóng xế tà” chứa bảy từ. Em hãy viết chương trình xác định và đưa ra màn hình tổng số từ trong hai xâu s1 và s2 đã cho

Ví dụ:

Input	Output
Dưới trăng quỳên đã gọi hè Đầu tường lửa lựu lập lòe đâm bông	14

Bài 4: Trong các câu sau đây, những câu nào đúng?

- 1) Có thể ghép các xâu để được xâu mới
- 2) Có thể tìm vị trí một xâu con trong một xâu
- 3) Không thể xóa một xâu con trong một xâu
- 4) Không thể thay đổi một xâu con trong một xâu