

SAU BÀI NÀY EM SẼ:

- Biết và thực hiện được một số lệnh thường dùng với chuỗi ký tự.



Bài toán tìm kiếm chuỗi con trong một chuỗi là một trong những bài toán tin học được ứng dụng nhiều trong thực tế. Công cụ tìm kiếm thông tin trên Internet hay lệnh tìm kiếm trong soạn thảo văn bản được xây dựng trên cơ sở bài toán tìm chuỗi con.

Cho chuỗi $c = \text{"Trường Sơn"}$ và chuỗi $m = \text{"Bước chân trên dải Trường Sơn"}$. Em hãy cho biết chuỗi c có là chuỗi con của chuỗi m không? Nếu có thì tìm vị trí của chuỗi c trong chuỗi m .

1. XÂU CON VÀ LỆNH TÌM VỊ TRÍ XÂU CON

Hoạt động 1 Một số lệnh tìm kiếm chuỗi con trong chuỗi ký tự

Quan sát các ví dụ sau để tìm hiểu cách kiểm tra chuỗi con và tìm kiếm vị trí chuỗi con trong chuỗi ký tự.



Ví dụ 1. Dùng toán tử `in` để kiểm tra một chuỗi có là chuỗi con của chuỗi khác không.

```
>>> "abc" in "123abc"
```

```
True
```

```
>>> "010" in "1101"
```

```
False
```

Biểu thức kiểm tra `<chuỗi 1> nằm trong <chuỗi 2>` là:

```
<chuỗi 1> in <chuỗi 2>
```

Nếu đúng thì trả lại giá trị `True`, nếu sai trả lại giá trị `False`.

Ví dụ 2. Lệnh `find()` tìm vị trí xuất hiện của một chuỗi trong chuỗi khác.

```
>>> s = "ab bc cd 123 456 00"
```

```
>>> s.find("b")
```

```
1 ← Vị trí xuất hiện đầu tiên của "b" trong chuỗi s là chỉ số 1.
```

```
>>> s.find("12")
```

```
9 ← Vị trí tìm thấy đầu tiên của "12" trong chuỗi s là chỉ số 9.
```

```
>>> s.find("AB")
```

```
-1 ← Không tìm thấy chuỗi "AB" trong chuỗi s nên trả về -1.
```

Tương tự danh sách, Python cũng có một số lệnh đặc biệt dành riêng cho chuỗi ký tự (phương thức). Cách thực hiện phương thức là:

```
<chuỗi>.<phương thức>
```

Cú pháp đơn của lệnh `find()`:

```
<chuỗi mẹ>.find(<chuỗi con>)
```

Lệnh sẽ tìm vị trí đầu tiên của chuỗi con trong chuỗi mẹ và trả về vị trí đó. Nếu không tìm thấy thì trả về -1.

Cú pháp đầy đủ của lệnh `find()`:

```
<chuỗi mẹ>.find(<chuỗi con>, start)
```

Lệnh sẽ tìm vị trí chuỗi con bắt đầu từ vị trí `start`.

Ví dụ 3

```
>>> sub = "Đà Nẵng"
>>> s = "Hà Nội - Đà Nẵng - Hồ Chí Minh"
>>> s.find(sub)
9
>>> s.find(sub, 10)
-1
```

Để tìm một chuỗi trong một chuỗi khác có thể dùng toán tử `in` hoặc lệnh `find()`. Lệnh `find()` trả về vị trí của chuỗi con trong chuỗi mẹ.



1. Biểu thức logic sau là đúng hay sai?

```
>>> "010" in "001100"
```

2. Lệnh sau trả lại giá trị gì?

```
>>> "ababababab".find("ab", 4)
```

2. MỘT SỐ LỆNH THƯỜNG DÙNG VỚI CHUỖI KÝ TỰ

Hoạt động 2 Một số lệnh thường dùng với chuỗi ký tự

Quan sát các ví dụ sau để biết cách sử dụng một số lệnh thường dùng với chuỗi ký tự như: `split()`, `join()`.



Ví dụ 1. Lệnh `split()` tách một chuỗi thành danh sách các từ.

```
>>> s = "Tiên học lễ hậu học văn"
>>> s.split()  ← Tách chuỗi dùng dấu cách để phân biệt tách.
['Tiên', 'học', 'lễ', 'hậu', 'học', 'văn']
>>> st = "0,1,2,3,4,5,6,10"
>>> st.split(",")  ← Tách chuỗi dùng dấu "," để phân biệt tách.
['0', '1', '2', '3', '4', '5', '6', '10']
```

Lệnh `split()` tách một chuỗi thành các từ và đưa vào một danh sách. Kí tự tách dùng để phân tách các từ mặc định là dấu cách, tuy nhiên có thể thay thế kí tự tách bằng kí tự khác.

Cú pháp của lệnh `split()`.

```
<chuỗi>.split(<kí tự tách>)
```

Ví dụ 2. Lệnh `join()` nối danh sách gồm các từ thành một chuỗi.

```
>>> A = ['Tiên', 'học', 'lễ', 'hậu', 'học', 'văn']
>>> " ".join(A)
'Tiên học lễ hậu học văn'
>>> B = ['0', '1', '2', '3', '4', '5', '6', '10']
>>> ",".join(B)
'0,1,2,3,4,5,6,10'
```

Lệnh `join()` này sẽ nối các phần tử của danh sách A bởi dấu cách.

Lệnh `join()` này sẽ nối các phần tử của danh sách B bởi dấu ",".

Lệnh `join()` có tác dụng ngược với lệnh `split()`, có chức năng nối các phần tử (là chuỗi) của một danh sách thành một chuỗi. Cú pháp của lệnh `join()` là:

```
"kí tự nối".join(<danh sách>)
```

Python có các lệnh đặc biệt để xử lí chuỗi là `split()` dùng để tách chuỗi thành danh sách và lệnh `join()` dùng để nối danh sách các chuỗi thành một chuỗi.



Cho chuỗi kí tự: "gà,vịt,chó,lợn,ngựa,cá". Em hãy trình bày cách làm để xoá các dấu "," và thay thế bằng dấu " " trong chuỗi này.



THỰC HÀNH

Một số bài toán liên quan đến chuỗi kí tự.

Nhiệm vụ 1. Viết chương trình nhập nhiều số nguyên từ bàn phím, các số cách nhau bởi dấu cách. Khi nhập xong thông báo số lượng các số đã nhập và in các số này thành hàng ngang.

Hướng dẫn. Dữ liệu nhập vào là một chuỗi. Dùng lệnh `split()` để tách thành danh sách. Chuyển các phần tử của danh sách này thành số và in ra màn hình.

```
s = input("Nhập các số nguyên cách nhau bởi dấu cách: ")
sline = s.split()
n = len(sline)
nline = []
for x in sline:
    nline.append(int(x))
print("Bạn đã nhập",n,"số.")
for k in nline:
    print(k, end = " ")
```

Nhiệm vụ 2. Viết chương trình nhập một chuỗi ký tự có thể có nhiều dấu cách giữa các từ. Sau đó chỉnh sửa chuỗi ký tự đó sao cho giữa các từ chỉ có một dấu cách. In chuỗi kết quả ra màn hình.

Hướng dẫn. Chuyển chuỗi ký tự ban đầu thành danh sách các từ đơn bằng lệnh `split()`, sau đó nối các từ đơn này bằng lệnh `join()`.

```
s = input("Nhập đoạn văn bản: \n")
sline = s.split()
skq = " ".join(sline)
print(skq)
```

Nhiệm vụ 3. Viết chương trình nhập số tự nhiên n , rồi nhập họ tên của n học sinh. Sau đó in ra danh sách tên học sinh theo hai cột, cột 1 là tên, cột 2 là họ đệm.

Hướng dẫn. Họ tên ban đầu tách ra thành tên và họ đệm bằng lệnh `split()`. Các tên được đưa vào danh sách `ten`, các họ đệm được đưa vào danh sách `hodem`. Sau đó in ra danh sách theo yêu cầu.

```
n = int(input("Nhập số học sinh trong lớp: "))
ten = []
hodem = []
for i in range(n):
    s = input("Nhập họ tên học sinh thứ "+str(i+1)+" : ")
    sline = s.split()
    m = len(sline)-1
    ten.append(sline[m])
    del sline[m]
    hodem.append(" ".join(sline))
print("Danh sách học sinh:")
for i in range(n):
    print(ten[i],hodem[i])
```



LUYỆN TẬP

1. Viết chương trình nhập nhiều số (số nguyên hoặc số thực) từ bàn phím, các số cách nhau bởi dấu cách. Sau đó in ra màn hình tổng các số đã nhập.
2. Viết chương trình nhập họ tên đầy đủ của người dùng, sau đó in thông báo tên và họ đệm của người đó.



VẬN DỤNG

1. Viết chương trình nhập hai số tự nhiên từ bàn phím, cách nhau bởi dấu cách và đưa ra kết quả là ƯCLN của hai số này.
2. Viết chương trình nhập số tự nhiên n rồi nhập n họ tên học sinh. Sau đó yêu cầu nhập một tên và thông báo số bạn có cùng tên đó trong lớp.