

SAU BÀI NÀY EM SẼ:

- Hiểu được chuỗi ký tự là kiểu dữ liệu cơ bản của Python.
- Biết và thực hiện được lệnh for để xử lý chuỗi ký tự.



Em đã biết 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 = "Thời khóa biểu"
>>> xau = '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 CÁC KÍ TỰ

Hoạt động 1 Tìm hiểu cấu trúc của chuỗi ký tự

Quan sát các ví dụ sau để biết cấu trúc chuỗi ký tự, so sánh với danh sách để biết sự khác nhau giữa chuỗi (string) và danh sách (list).



Ví dụ 1. Chuỗi ký tự và cách truy cập đến từng ký tự của chuỗi.

```
>>> s = "Thời khóa biểu"
>>> len(s)
14
>>> s[0]
'T'
>>> s[10]
'b'
```

Lệnh `len()` sẽ tính độ dài của chuỗi hay số lượng các ký tự có trong chuỗi.

Có thể truy cập từng ký tự của chuỗi thông qua chỉ số. Chỉ số bắt đầu từ 0.

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 = ["a", "b", "c"]
>>> d[0] = "A"
>>> s = "abc"
>>> s[0] = "A"
Traceback (most recent call last):
  File "<pyshell#15>", line 1, in <module>
    s[0] = "A"
TypeError: 'str' object does not support item assignment
```

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 = ""
```

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?

2. LỆNH DUYỆT KÍ TỰ CỦA XÂU

Hoạt động 2 Tìm hiểu lệnh duyệt từng kí tự của xâu

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 = "Thời khóa biểu"
>>> for i in range(len(s)):
    print(s[i], end = " ")
```

← Duyệt theo chỉ số với lệnh `range()`.

```
T h ờ i   k h o á   b i ế u
>>> for ch in s:
    print(ch, end = " ")
```

← Duyệt theo kí tự của xâu kí tự.

T h ờ i k h o á b i ế u



- Cách thứ nhất, biến `i` lần lượt chạy theo chỉ số của xâu kí tự `s`, từ 0 đến `len(s) - 1`. Kí tự tại chỉ số `i` là `s[i]`.
- Cách 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
```

Có thể duyệt các kí tự của chuỗi bằng lệnh `for` tương tự như với danh sách.
`s1 in s2` trả lại giá trị `True` nếu `s1` là chuỗi con của `s2`.



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 chuỗi 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:

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

Nhiệm vụ 2. Nhập một chuỗi kí tự `S` từ bàn phím rồi kiểm tra xem chuỗi `S` có chứa chuỗi con "10" không.

Hướng dẫn. Cách 1. Nếu chuỗi `S` chứa chuỗi con "10" thì sẽ có chỉ số `k` mà `S[k] = "1"` và `S[k+1] = "0"`. Cách 2. Dùng toán tử `in` để kiểm tra chuỗi "10" có là chuỗi con của `S`.

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

```
S = input("Nhập chuỗi kí tự bất kì: ")
kq = False
```

```

for i in range(len(S)-1):
    if S[i] == "1" and S[i+1] == "0":
        kq = True
        break
if kq:
    print("Xâu gốc có chứa xâu '10'")
else:
    print("Xâu gốc không chứa xâu '10'")

```

Cách 2: Sử dụng toán tử `in`.

```

S = input("Nhập xâu kí tự bất kì: ")
s10 = "10"
if s10 in S:
    print("Xâu gốc có chứa xâu '10'")
else:
    print("Xâu gốc không chứa xâu '10'")

```



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".



VẬN DỤNG

1. 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.
2. 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.