

Dữ liệu kiểu chuỗi trong Python

Chuỗi và các thao tác trên chuỗi

TS. Trần Quang Quý
e-mail: tqquy@ictu.edu.vn
Contact: 0818981166(zalo)

Trường Đại học Công nghệ Thông tin và
Truyền thông Thái Nguyên

Tháng 09 năm 2021



Định nghĩa chuỗi

- Chuỗi là một dãy các ký tự. Mọi ký tự nằm trong dấu " " hoặc ' ' thì đều được xem là chuỗi trong Python. Chuỗi rỗng là chuỗi không có ký tự nào.
- Các ký tự trong chuỗi thuộc bảng mã Unicode hoặc ASCII, do đó chuỗi trong Python có hỗ trợ tiếng Việt. Chỉ số trong chuỗi được đánh số từ 0, 1, ..., và đánh -1, -2, ... theo chiều từ trái sang phải và từ phải sang trái tương ứng.
- Số ký tự nằm trong chuỗi chính là độ dài của chuỗi đó.
- Chúng ta lưu ý là máy tính không thể nhận ký tự vào để đọc được mà nó cần mã hóa. Quá trình chuyển ký tự thành số nhị phân gọi là mã hóa và quá trình ngược lại chúng ta gọi là giải mã. Unicode và ASCII là một trong những dạng mã hóa phổ biến.



Một số thao tác với chuỗi

- Để lấy được độ dài của chuỗi hay chính là số ký tự của chuỗi chúng ta sử dụng hàm `len(namestring)`;
- Để truy cập vào từng phần tử của chuỗi, chúng ta sử dụng `namestring[index]`
- Để in lần lượt từng phần tử của chuỗi, sử dụng vòng lặp `for`.
- Để ghép (cộng các chuỗi) chúng ta sử dụng cú pháp (chuỗi 1) + (chuỗi 2)+...
- Để thực hiện việc lặp chuỗi sử dụng cú pháp: (số lần lặp)*(chuỗi cần lặp).
- So sánh chuỗi: Việc so sánh chuỗi sẽ thực hiện bằng cách so sánh từng phần tử trong các chuỗi từ trái sang phải, nếu chuỗi nào có phần tử lớn hơn thì chuỗi đó lớn hơn.



Một số thao tác với chuỗi - tiếp

- Để thực hiện việc lấy chuỗi con từ chuỗi ban đầu, sử dụng toán tử slicing (`:`)- `namestring[a:b]`, thực hiện việc lấy chuỗi con từ chỉ số là a đến chỉ số $b-1$ với $a \leq b$
- Dạng tham số của slicing là: `namestring[a,b,c]` với chỉ số lần lượt là a , $a+c$, $a+2c, \dots, a+jc$ với $a + jc \leq b$
- Để đảo ngược một chuỗi, chúng ta sử dụng phương thức: `chuỗi[::-1]`
- **Nhận xét:** Trong Python không có kiểu ký tự như nhiều ngôn ngữ lập trình khác. Mỗi ký tự trong Python được xem như là một chuỗi có độ dài bằng 1.
- Khi khai báo một chuỗi, chúng ta không thể thay đổi các ký tự trong biến chuỗi đó.
- Để ghép chuỗi với số hoặc kiểu dữ liệu khác, ta có thể thực hiện ép kiểu bằng phương thức **`str(biến kiểu số)`**



Bài tập thực hành với chuỗi

- ❶ **Bài 1:** Sử dụng vòng lặp for, hãy kiểm tra xem chuỗi nhập vào có bao nhiêu ký tự thỏa mãn yêu cầu đề bài.
- ❷ **Bài 2:** Tách ký tự số và ký tự chữ cái: Hãy nhập vào một chuỗi gồm các ký tự số và ký tự chữ cái La-tinh thường 'a','b',...,'z'. Hãy tách chuỗi được nhập thành hai chuỗi, một chuỗi gồm các ký tự số và chuỗi còn lại gồm các ký tự chữ cái.



Một số thao tác khác với chuỗi

- Chúng ta sử dụng phương thức `namestring.split()` để tách chuỗi thành danh sách các chuỗi con mà ký tự phân tách là các ký tự trắng.
- Phương thức `split()` nêu lên ở phía trên có thể ở dạng tham số nếu sử dụng dạng như sau: `namestring.split(",")`.

