

# Dữ liệu kiểu tập hợp trong Python

Tập hợp và các thao tác với tập hợp

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

Bộ môn Khoa học Máy tính và Công nghệ  
Khoa Công nghệ Thông tin

Tháng 09 năm 2021



# Định nghĩa tập hợp

Tập hợp trong Python hay còn được gọi là kiểu dữ liệu Set, cho phép biểu diễn và thực hiện các phép toán trên tập hợp giống như trong Toán học. Đây là một điểm mạnh của Python, các kiểu Set trong Python có tính chất sau:

- Tập hợp là các phần tử không có thứ tự và không được đánh chỉ số. Các phần tử trong tập hợp phải khác nhau. Nếu các phần tử trong tập hợp là giống nhau thì sẽ bị loại bỏ. Để khởi tạo một tập hợp chúng ta sử dụng dấu ngoặc nhọn {}
- Có 2 cách để khởi tạo tập hợp, sử dụng khai báo trực tiếp và sử dụng khai báo qua từ khóa `set(object-name)`, có thể hiểu như ép dữ liệu kiểu tập hợp.
- Để duyệt qua các phần tử của tập hợp, chúng ta sử dụng vòng lặp `for`.
- **Chú ý:** `set1 = {}` là một câu lệnh sai, đây không phải là tập hợp rỗng. Để khởi tạo một tập hợp rỗng, sử dụng cú pháp `set1 = set()`.



# Các phép toán trên tập hợp

Chúng ta có một số phép toán trên tập hợp như sau:

- **Phép giao** - Giao của hai tập hợp A và B là một tập hợp, bao gồm tất cả các phần tử thuộc cùng cả hai tập hợp. Ví dụ:  $C = A \cap B$ . Ngoài ra chúng ta có thể sử dụng phương thức **intersection** để tìm giao của hai tập hợp. Ví dụ:  $C = B.intersection(A)$ .
- **Phép hợp** Hợp của hai tập hợp A và B là một tập hợp bao gồm tất cả các phần tử thuộc tập A hoặc thuộc tập B. Ký hiệu  $A \cup B$ . Ngoài ra chúng ta có thể sử dụng phương thức **union** để tìm hợp của hai tập hợp. Ví dụ:  $C = A.union(B)$ .
- Để thêm một phần tử x vào tập hợp A, ta hợp tập hợp A với tập x, hoặc sử dụng phương thức **add(element)**.
- **Phép hiệu**: Hiệu của tập hợp A và tập hợp B là một tập hợp bao gồm tất cả các phần tử thuộc A nhưng không thuộc B. Ký hiệu là  $A - B$  hoặc sử dụng phương thức **A.difference(B)**



# Các phép toán trên tập hợp-tiếp

- **Phương thức discard** dùng để loại bỏ một phần tử  $x$  khỏi tập hợp  $A$ , ta sử dụng phép hiệu giữa tập hợp  $A$  và  $x$  hoặc sử dụng phương thức **discard**. Ví dụ: `A.discard(4)`
- **Phép hiệu đối xứng** - Hiệu đối xứng của hai tập hợp  $A$  và  $B$  là một tập hợp bao gồm tất cả các phần tử thuộc  $A$  hoặc thuộc  $B$  nhưng không thuộc cả  $A$  và  $B$ . Ngoài ra chúng ta có thể sử dụng phương thức **symmetric-difference** để tìm hiệu đối xứng của hai tập hợp.
- Để kiểm tra một phần tử có thuộc một tập hợp hay không, có thể sử dụng toán tử **in** và toán tử **not in**



# Các phép toán đối với tập hợp con

- ➊ **Phép toán so sánh bằng giữa hai tập hợp.** Ký hiệu  $A == B$ ; biểu thức trả về giá trị là True nếu hai tập hợp đó giống nhau, tức là mọi phần tử thuộc tập hợp A thì cũng thuộc tập hợp B và ngược lại, trường hợp khác nhau thì trả về False
- ➋ **Phép toán kiểm tra tập con** - Trả về giá trị True nếu A là tập hợp con của B, tức là mọi phần tử thuộc tập hợp A đều thuộc tập hợp B; ngược lại trả về giá trị False. Ký hiệu  $A \leq B$
- ➌ **Phép toán kiểm tra tập con thực sự** - Trả về giá trị True nếu A là tập con thực sự của B, tức là mọi phần tử thuộc tập hợp A đều thuộc tập B nhưng có phần tử thuộc B lại không thuộc A.
- ➍ **Phép toán all và any** - Trong toán học, chúng ta sử dụng ký hiệu  $\forall$  để nói về mọi phần tử trong một tập hợp và ký hiệu  $\exists$  để nói sự tồn tại của một phần tử trong một tập hợp.



# Phép toán all và any trong tập hợp

Xét ví dụ như sau: Cho  $A = \{1, 2, 3, 4, 5\}$ ,  $B = \{\forall x \in A | x : 2\}$  khi đó  $B = \{2, 4\}$ .

- ➊ **Phép toán all()** câu lệnh của phép toán này là `all([x > 0 for x in SetA])`. Có nghĩa là đây là khẳng định mọi giá trị trong tập hợp A đều  $> 0$ , kết quả của biểu thức là True hoặc False tùy thuộc vào điều kiện.
- ➋ **Phép toán any()** câu lệnh của phép toán này tương đương với phép tồn tại trong tập hợp. Cú pháp của câu lệnh này là: `any([x < 0 for x in SetA])`. Tức là nếu tồn tại một giá trị nhỏ hơn không trong A thì biểu thức trả kết quả về là True, ngược lại là False.



- ➊ **Tổng chữ số chung.** Viết chương trình trong Python nhập vào hai số tự nhiên  $a$  và  $b$ . Ví dụ  $a=1123499$ ,  $b = 1112229$ ; hai số này đều có chữ số chung là 1,2,9; nên tổng các chữ số chung bằng  $1+2+9=12$ . Lưu ý: Nếu có một chữ số cùng xuất hiện nhiều lần trong cả  $a$  và  $b$  thì chữ số đó chỉ tính một lần.
- ➋ **Phần thưởng vở.** Một lớp mẫu giáo gồm có  $n$  em bé và các em này được đánh số thứ tự từ 1 đến  $n$ , nhà trường tổ chức phát tặng vở cho các cháu trong lớp học theo phiếu bé ngoan, 1 phiếu nhận 1 quyển vở, 2 phiếu nhận 2 quyển vở, 0 phiếu không nhận vở. Nhận vở chia làm hai đợt.
  - Đợt 1 gồm có  $k$  em theo thứ tự là  $l_1, l_2, \dots, l_k$
  - Đợt 2 gồm  $t$  các em có số thứ tự là  $ll_1, ll_2, \dots, ll_t$ .



Trong mỗi đợt phát thưởng, mỗi em trong từng đợt sẽ nhận được 1 quyển vở, tuy nhiên sẽ có em nhận được 2 quyển vở. Yêu cầu: Nhập vào số học sinh  $n$  của lớp, số lượng học sinh nhận phần thưởng đợt 1 ( $k$  em) và số lượng học sinh nhận phần thưởng đợt 2 ( $t$  em). In ra danh sách các em nhận 2 quyển vở, nhận 1 quyển vở và không nhận quyển nào.

- 3 **Danh sách đại diện.** Nhập danh sách  $a$  gồm  $n$  số nguyên  $a[0], a[1], \dots, a[n-1]$ . Tạo danh sách  $b$  chỉ gồm các ký tự đại diện trong  $a$ , tức là nếu phần tử xuất hiện nhiều lần thì chỉ lấy một phần tử. Các phần tử trong  $b$  được sắp xếp theo chiều tăng dần.
- 4 **Học sinh giỏi Toán, Lý, Hóa** Lớp học có  $n$  học sinh tham gia kỳ thi học sinh giỏi cấp tỉnh về ba môn Toán, Lý, Hóa. Các học sinh được đánh số thứ tự từ 1 đến  $n$ .
  - Môn Toán có  $a$  học sinh thi
  - Môn Lý có  $b$  học sinh thi





- Môn Hóa có c học sinh thi.

Yêu cầu: Hãy đưa ra danh sách các em chỉ thi một môn, các em thi hai môn và các em thi cả ba môn Toán, Lý, Hóa. Ví dụ như sau: Giả sử  $n=10$ ; thi môn Toán gồm các em trong tập hợp  $\{1, 2, 3, 7, 8\}$ ; thi môn Lý gồm các em trong tập hợp  $\{1, 2, 6, 9, 10\}$ ; thi môn Hóa gồm các em trong tập hợp  $\{4, 5, 6, 7, 9\}$ . Thi một môn Toán gồm các em số 3,8; Lý:10; Hóa 4,5 Thi hai môn Toán, Lý:1,2. Thi ba môn không có em nào.

