

**Bài 1.** Xây dựng chương trình/thủ tục sắp xếp tiếng Việt.

**Bài 2.** Xây dựng cách mã hóa tiếng Việt sao cho có thể tích hợp vào câu lệnh SQL của hệ quản trị cơ sở dữ liệu để sắp xếp các trường dữ liệu tiếng Việt.

**Bài 3.** Cho một mảng các chuỗi ký tự. Ví dụ:  $s[] = \{\text{"flower"}, \text{"flow"}, \text{"flight"}\}$ .

- Vẽ hình minh họa quá trình chèn các ký tự của các chuỗi trong mảng  $s[]$  vào cây tiền tố (trie).
- Viết chương trình chèn các chuỗi ký tự vào cây tiền tố và tìm tiền tố chung dài nhất của các chuỗi ký tự đó. Ví dụ: tiền tố chung dài nhất của mảng  $s[]$  là “fl”.

**Bài 4.** Cho một mảng các ký tự  $ch[]$  và tần suất của mỗi ký tự là  $freq[]$  như sau:

Đầu vào:  $ch[] = \{\text{'a'}, \text{'b'}, \text{'c'}, \text{'d'}, \text{'e'}, \text{'f'}\}$ ,  $freq[] = \{5, 9, 12, 13, 16, 45\}$ .

Đầu ra:

f 0  
c 100  
d 101  
a 1100  
b 1101  
e 111

- Vẽ cây Huffman với dữ liệu đầu ra như trên.
- Viết chương trình xây dựng cây Huffman và gán từ mã (code word) cho mọi ký tự trong  $ch[]$  bằng cách sử dụng Hàng đợi ưu tiên (có thể khai thác thư viện của ngôn ngữ lập trình).
- Vẽ minh họa các bước xây dựng cây Huffman với dữ liệu đầu vào như trên.

## **Bài 5. GOOD OR BAD?**

Mật rất thích học về chuỗi ký tự nên anh của cậu đã dành 1 tối để dạy cậu. Với tư chất thông minh Mật đã tiếp thu được hết những gì được dạy. Để kiểm tra Mật, anh cậu đã nghĩ ra một bài về chuỗi ký tự và hứa rằng nếu cậu làm được sẽ đưa cậu đi xem phim vào cuối tuần.

Đề bài là cho  $N$  chuỗi kí tự chỉ gồm các kí tự từ ‘a’ đến ‘j’. Kiểm tra xem  $N$  chuỗi này là GOOD hoặc BAD.  $N$  chuỗi được gọi là GOOD khi mà không có chuỗi nào là tiền tố của 1 chuỗi khác, còn lại thì là BAD.

### Input

Dòng đầu tiên là  $N$  ( $N \leq 10^5$ )

$N$  dòng sau mỗi dòng là 1 chuỗi kí tự dài không quá 60 kí tự

### Output

Ghi ra GOOD SET khi  $N$  chuỗi là GOOD.

Nếu  $N$  chuỗi là BAD thì ghi ra BAD SET và chuỗi đầu tiên không thỏa mãn điều kiện.

### Example

#### Test 1:

##### Input:

3  
ab  
bc  
cd

##### Output:

GOOD SET

#### Test 2:

##### Input:

4  
aabd  
jjjjj  
cccc  
aabdj

##### Output:

BAD SET  
aabdj

### Bài 6. Chuỗi từ

Chuỗi từ có độ dài  $n$  là một dãy các từ  $w_1, w_2, \dots, w_n$  sao cho với mọi  $1 \leq i < n$ , từ  $w_i$  là tiền tố của từ  $w_{i+1}$ . Từ  $u$  có độ dài  $k$  là tiền tố của từ  $v$  có độ dài  $l$  nếu  $l > k$  và các ký tự đầu tiên của  $v$  trùng với từ  $u$ . Cho tập hợp các từ  $S = s_1, s_2, \dots, s_m$ . Tìm chuỗi từ dài nhất có thể xây dựng được bằng cách dùng các từ trong tập hợp  $S$  (có thể không sử dụng hết các từ).

## Bài 7. Similar String Groups

Hai chuỗi X và Y là tương tự nhau nếu chúng ta có thể hoán đổi hai chữ cái (ở các vị trí khác nhau) của X, sao cho nó bằng Y. Ngoài ra hai chuỗi X và Y cũng là tương tự nhau nếu chúng bằng nhau.

Ví dụ: "tars" và "rats" là tương tự nhau (hoán đổi vị trí 0 và 2), và "rats" và "arts" là tương tự nhau, nhưng "star" không tương tự với "tars", "rats" hoặc "arts".

Những nhóm này cùng tạo thành hai nhóm được kết nối với nhau bởi tính tương tự: {"tars", "rats", "arts"} và {"star"}. Chú ý rằng "tars" và "arts" ở cùng một nhóm mặc dù chúng không tương tự nhau. Về mặt hình thức, mỗi nhóm sao cho một từ nằm trong nhóm nếu và chỉ khi nó tương tự với ít nhất một từ khác trong nhóm.

Cho một danh sách các chuỗi A. Mọi chuỗi trong A là đảo chữ của mọi chuỗi khác trong A. Có bao nhiêu nhóm?

**Ví dụ 1:**

**Đầu vào:** A = ["tars", "rats", "arts", "star"]

**Đầu ra:** 2

**Các ràng buộc:**

- $1 \leq A.length \leq 2000$
- $1 \leq A[i].length \leq 1000$
- $A.length * A[i].length \leq 20000$
- Tất cả các từ trong A chỉ bao gồm các chữ thường.
- Tất cả các từ trong A có độ dài bằng nhau và là các đảo chữ của từ khác.