

**BÀI TẬP PHẦN XÂU KÝ TỰ**

- **Cấu trúc:** có thể coi một chuỗi (hay chuỗi) ký tự s là một mảng một chiều, mỗi phần tử là 1 ký tự (kiểu char). Ký tự kết thúc là ký tự '\0'. Còn có cách khác để sử dụng kiểu chuỗi ký tự là khai báo dưới dạng con trỏ, tuy nhiên cách này khá phức tạp.

- **Tài liệu:** đọc thêm tài liệu phần 8.2.2 Mảng chuỗi, trang 82 và 16. Một số hàm chuẩn thường dùng, trang 182.

- **Chương trình mẫu:**

```
#include <stdio.h>
#include <string.h> //Khai báo thư viện chứa các hàm xử lý chuỗi ký tự (string)
#include <ctype.h> //Khai báo thư viện chứa các hàm xử lý ký tự (char)
int main()
{
    FILE *f1, *f2;
    char s[200], x[200]; //Khai báo chuỗi ký tự s và x dưới dạng mảng ký tự
    //1. Nhập dữ liệu
    f1 = fopen("input.txt", "r");
    fgets(s, sizeof(s), f1); //so sánh với fscanf(f1, "%s", s); fscanf(f1, "%d", &n);
    fclose(f1);
    //2. Xử lý + Xuất kết quả
    f2 = fopen("output.txt", "w");
    //B1: Loại bỏ các ký tự khác chữ và khoảng trắng
    int n = strlen(s); //Hàm lấy độ dài chuỗi s
    int j = 0;
    for(int i = 0; i < n; i++)
        if (isalpha(s[i]) || isspace(s[i])) //Hàm kiểm tra chữ cái, khoảng trắng
        {
            x[j] = s[i];
            j++;
        }
    x[j] = '\0'; //Kết thúc chuỗi ký tự x

    //B2: Thao tác xử lý khác

    fprintf(f2, "%s", x); //Xuất kết quả là chuỗi ký tự x
    fclose(f2);
    return 0;
}
```

- 1) Tập tin nguồn: Cau1.cpp
  - Nhập: từ tập tin văn bản Cau1.inp chuỗi ký tự S có không quá 200 ký tự.
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.
    - + Sau đó, chuyển thành chuỗi ký tự X là dạng **UPPER CASE** của S.
  - Xuất: ra tập tin văn bản Cau1.out chuỗi ký tự X.

Ví dụ: Nếu S = ‘ Khoa1 Cong nghe thong@ tin’ thì X = ‘ KHOA CONG NGHE THONG TIN’.

2) Tập tin nguồn: Cau2.cpp

- Nhập: từ tập tin văn bản Cau2.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, chuyển thành xâu ký tự X là dạng **lower case** của S.

- Xuất: ra tập tin văn bản Cau2.out xâu ký tự X.

Ví dụ: Nếu S = ‘ Khoa1 Cong nghe thong@ tin’ thì X = ‘ khoa cong nghe thong tin’.

3) Tập tin nguồn: Cau3.cpp

- Nhập: từ tập tin văn bản Cau3.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, chuyển thành xâu ký tự X là dạng **Title Case** của S.

- Xuất: ra tập tin văn bản Cau3.out xâu ký tự X.

Ví dụ: Nếu S = ‘ Khoa1 Cong nghe thong@ tin’ thì X = ‘ Khoa Cong Nghe Thong Tin’.

4) Tập tin nguồn: Cau4.cpp

- Nhập: từ tập tin văn bản Cau4.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, chuyển thành xâu ký tự X là dạng **tOGGLE cASE** của S.

- Xuất: ra tập tin văn bản Cau4.out xâu ký tự X.

Ví dụ: Nếu S = ‘ Khoa1 Cong nghe thong@ tin’ thì X = ‘ kHOA cONG NGHE tHONG tIN’.

5) Tập tin nguồn: Cau5.cpp

- Nhập: từ tập tin văn bản Cau5.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, chuyển thành xâu ký tự X là dạng **xâu chuẩn** của S. Biết: xâu chuẩn là 1 xâu ký tự mà không có 2 khoảng trắng liên tiếp, bắt đầu và kết thúc không phải là khoảng trắng.

- Xuất: ra tập tin văn bản Cau5.out xâu ký tự X.

Ví dụ: Nếu S = ‘ Co@ng1 nghe2 34 th\$ong tin5 ’ thì X = ‘Cong nghe thong tin’.

6) Tập tin nguồn: Cau6.cpp

- Nhập: từ tập tin văn bản Cau6.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, chuyển thành xâu ký tự X là dạng **câu chuẩn** của S. Biết: câu chuẩn là 1 xâu ký tự mà không có 2 khoảng trắng liên tiếp, bắt đầu và kết thúc không phải là khoảng trắng, ký tự bắt đầu là ký tự chữ HOA.

- Xuất: ra tập tin văn bản Cau6.out xâu ký tự X.

- Ví dụ: Nếu  $S = \text{' co@ng1 nghe2 34 th$ong tin5 '}$  thì  $X = \text{'Cong nghe thong tin'}$ .
- 7) Tập tin nguồn: Cau7.cpp
- Nhập: từ tập tin văn bản Cau7.inp xâu ký tự  $S$  có không quá 200 ký tự.
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi  $S$ .
    - + Sau đó, chuyển thành xâu ký tự  $X$  là dạng **câu đảo** của  $S$ . Biết: câu đảo là 1 xâu ký tự mà không có 2 khoảng trắng liên tiếp, bắt đầu và kết thúc không phải là khoảng trắng, ký tự bắt đầu là ký tự chữ thường, các ký tự còn lại là chữ HOA.
  - Xuất: ra tập tin văn bản Cau7.out xâu ký tự  $X$ .
- Ví dụ: Nếu  $S = \text{' co@ng1 nghe2 34 th$ong tin5 '}$  thì  $X = \text{'cONG NGHE THONG TIN'}$ .
- 8) Tập tin nguồn: Cau8.cpp
- Nhập: từ tập tin văn bản Cau8.inp xâu ký tự  $S$  có không quá 200 ký tự.
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi  $S$ .
    - + Sau đó, chuyển thành xâu ký tự  $X$  là dạng **xâu đảo** của  $S$ . Biết: xâu đảo là xâu nhận được bằng cách viết xâu ký tự  $S$  theo thứ tự ngược lại.
  - Xuất: ra tập tin văn bản Cau7.out xâu ký tự  $X$ .
- Ví dụ: Nếu  $S = \text{' co@ng1 nghe2 34 th$ong tin5 '}$  thì  $X = \text{'nit gnoht ehgn gnoc'}$ .
- 9) Tập tin nguồn: Cau9.cpp
- Viết chương trình thực hiện các yêu cầu sau:
- Nhập: từ tập tin văn bản Cau9.inp xâu ký tự  $S$  có không quá 50 ký tự là họ tên Việt Nam của 1 người, trong đó có một số ký tự gõ nhầm (không phải ký tự chữ và gõ thừa khoảng trắng).
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi  $S$ .
    - + Sau đó, tìm xâu ký tự  $X$  là **tên** của người đó.
  - Xuất: ra tập tin văn bản Cau9.out xâu ký tự  $X$ .
- Ví dụ: Nếu  $S = \text{' Nguyen1 Van234 A5n67h '}$  thì  $X = \text{'Anh'}$ .
- 10) Tập tin nguồn: Cau10.cpp
- Nhập: từ tập tin văn bản Cau10.inp xâu ký tự  $S$  có không quá 50 ký tự là họ tên Việt Nam của 1 người, trong đó có một số ký tự gõ nhầm (không phải ký tự chữ và gõ thừa khoảng trắng).
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi  $S$ .
    - + Sau đó, tìm xâu ký tự  $X$  là **họ** của người đó.
  - Xuất: ra tập tin văn bản Cau10.out xâu ký tự  $X$ .
- Ví dụ: Nếu  $S = \text{' Nguyen1 Van234 A5n67h '}$  thì  $X = \text{'Nguyen'}$ .
- 11) Tập tin nguồn: Cau11.cpp
- Viết chương trình thực hiện các yêu cầu sau:
- Nhập: từ tập tin văn bản Cau11.inp xâu ký tự  $S$  có không quá 200 ký tự.
  - Xử lý:
    - + Loại bỏ các ký tự không phải ký tự chữ khỏi  $S$  để được xâu  $X$ .

+ Sau đó, kiểm tra xem X có phải là một **xâu ghép**. Biết: xâu ghép là 1 xâu ký tự được tạo thành bằng cách viết liên tiếp K lần ( $K > 1$ ) 1 xâu có độ dài ngắn hơn.

- Xuất: ra tập tin văn bản Cau11.out xâu ký tự X và giá trị lớn nhất của K nếu X là xâu ghép.

Ví dụ: Nếu  $S = 'AB@1 A2 BA B\$ A5B '$  thì  $X = 'ABABABAB'$  và  $K = 4$ .

12) Tập tin nguồn: Cau12.cpp

- Nhập: từ tập tin văn bản Cau12.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ khỏi S để được xâu X.

+ Sau đó, kiểm tra xem X có phải là một **xâu ghép**. Biết: xâu ghép là 1 xâu ký tự được tạo thành bằng cách viết liên tiếp K lần ( $K > 1$ ) 1 xâu có độ dài ngắn hơn.

- Xuất: ra tập tin văn bản Cau12.out xâu ký tự X và kết luận.

Ví dụ: Nếu  $S = 'AB@1 A2 BA B\$ A5B '$  thì  $X = 'ABABABAB'$  là xâu ghép.

13) Tập tin nguồn: Cau13.cpp

- Nhập: từ tập tin văn bản Cau13.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, tìm MAX là số ký tự của **từ dài nhất** trong xâu S. Biết: từ là một nhóm các ký tự liên tiếp khác khoảng trắng.

- Xuất: ra tập tin văn bản Cau13.out giá trị của MAX.

Ví dụ: nếu  $S = ' Cong1 nghe2 34 tho@ng tin5 '$  thì từ dài nhất là 'thong' có số ký tự là  $MAX = 5$ .

14) Tập tin nguồn: Cau14.cpp

- Nhập: từ tập tin văn bản Cau14.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, **đếm số từ** C của S. Biết: từ là một nhóm các ký tự liên tiếp khác khoảng trắng.

- Xuất: ra tập tin văn bản Cau14.out giá trị của C.

Ví dụ: nếu  $S = ' Cong1 nghe2 34 tho@ng tin5 '$  thì S có  $C = 4$  từ là 'Cong', 'nghe', 'thong', 'tin'.

15) Tập tin nguồn: Cau15.cpp

- Nhập: từ tập tin văn bản Cau15.inp xâu ký tự S có không quá 200 ký tự.

- Xử lý:

+ Loại bỏ các ký tự không phải ký tự chữ và khoảng trắng khỏi S.

+ Sau đó, kiểm tra xem xâu X nhận được có phải là xâu đối xứng hay không? Biết: xâu đối xứng là xâu ký tự mà khi viết ngược lại cũng được chính xâu đó. Nếu đúng, kết quả là 1, ngược lại kết quả là 0.

- Xuất: ra tập tin văn bản Cau15.out xâu X và kết quả kiểm tra Kt.

Ví dụ: nếu  $S = ' Cong1 nghe2 34 tho@ng tin5 '$  thì  $X = ' Cong nghe thong tin '$  và  $Kt = 0$ .