

BÀI 11. THỰC HÀNH CÁC THUẬT TOÁN XỬ LÝ XÂU KÝ TỰ

- Chiến lược vét cạn đối với bài toán tìm xâu con
- Thuật toán Boyer Moore Horspool và ứng dụng
- Thuật toán z và ứng dụng

BÀI TẬP THỰC HÀNH TRÊN LỚP

Bài tập 1: Xâu đối xứng

Input:

- Cho xâu ký tự s .
- Xâu s là đối xứng nếu đọc từ trái qua phải hoặc từ phải qua trái đều giống nhau (ví dụ: memoomem).

Output:

- Xác định xâu s có phải là xâu đối xứng hay không?
- Tìm xâu con đối xứng dài nhất trong xâu s có tâm là $s[k]$, với $k \geq 1$.

Yêu cầu thực hiện

- Thiết kế thuật toán A1 để kiểm tra xâu ký tự s có phải là xâu đối xứng hay không?
- Thiết kế thuật toán A2 để tìm xâu con đối xứng dài nhất trong xâu s có tâm là $s[k]$.
- Viết chương trình thực hiện:
 - o Cài đặt hàm F1 và F2 biểu diễn các thuật toán A1 và A2.
 - o Khởi tạo xâu ký tự s .
 - o Sử dụng thuật toán A1 để xác định xâu s có phải là một xâu đối xứng hay không? Thông báo kết quả.
 - o Nếu s không phải là xâu đối xứng, hãy nhập vào số nguyên k , tìm xâu con đối xứng dài nhất có tâm là $s[k]$, hiển thị kết quả.

Bài 2: Chuẩn hóa xâu

Input:

- Cho một đoạn văn bản tiếng anh p , không quá 50 từ.

Output:

- Đếm số lượng các từ “child” trong p và thay thế tất cả các từ “child” trong p bằng từ “children”.

Yêu cầu thực hiện

- Thiết kế thuật toán A3 để đếm và trả về số lượng các từ “child” trong p .

- Thiết kế thuật toán A4 để thay thế tất cả các từ “child” trong p bằng từ “children”.
- Cài đặt chương trình ứng dụng:
- Cài đặt các hàm biểu diễn các thuật toán A3 và A4.
- Khởi tạo một đoạn văn bản s không quá 50 từ.
- Sử dụng thuật toán A3 để đếm số từ “child” trong s. Thông báo kết quả.
- Sử dụng thuật toán A4 để thay thế tất cả các từ child (nếu có) trong p bằng từ “children”. Thông báo kết quả.

Bài 3: Thuật toán boyer moore horspool

Input:

- Cho chuỗi ký tự s chỉ chứa các chữ cái, chữ số và dấu cách.

Output:

- Cho biết chuỗi ký tự p = “1000” có xuất hiện trong chuỗi s hay không?

Cài đặt chương trình thực hiện:

- Cài đặt hàm biểu diễn thuật toán Boyer Moore Horspool tìm và trả về vị trí xuất hiện (đầu tiên) của chuỗi p trong chuỗi s.
- Khởi tạo một chuỗi ký tự s.
- Sử dụng thuật toán Boyer Moore Horspool cho biết chuỗi p = “1000” có xuất hiện trong chuỗi s hay không? Thông báo kết quả.
- Thay thế chuỗi p = “1000” tìm được trong chuỗi s bằng chuỗi q = “một nghìn dong”, hiển thị chuỗi mới.

Bài 4: Thuật toán z

Input:

- Cho một đoạn văn bản s chỉ chứa các chữ cái, chữ số và dấu cách.

Output:

- Tìm và trả về w là từ đầu tiên của đoạn văn bản s (một từ là một dãy ký tự liên tiếp dài nhất không có chứa dấu cách).
- Đếm và trả về số lần xuất hiện của từ w trong đoạn văn bản s.

Cài đặt chương trình thực hiện:

- Khởi tạo đoạn văn bản s.
- Cài đặt hàm F1 tìm và trả về từ đầu tiên của của đoạn văn bản s.
- Cài đặt hàm Z biểu diễn thuật toán z.
- Sử dụng hàm F1 để tìm và trả về w là từ đầu tiên của s. Hiển thị kết quả.
- Sử dụng hàm Z để đếm và trả về số lần xuất hiện của w trong s. Hiển thị kết quả.

- Xóa tất cả các từ w trong đoạn văn bản s. Hiển thị kết quả.

BÀI TẬP THỰC HÀNH SAU KHI LÊN LỚP

Bài tập 1: Xâu tiền tố, hậu tố

- Xâu a được gọi là tiền tố của xâu b nếu xâu a trùng với phần đầu của xâu b. Ví dụ “pre” là tiền tố của “prefix”.
- Xâu a được gọi là hậu tố của xâu b nếu xâu a trùng với phần cuối của xâu b. Ví dụ “fix” là hậu tố của “suffix”.
- Cho 2 xâu a và b chỉ gồm các chữ cái latin. Hãy tìm một xâu c thỏa mãn:
 - o Xâu a là tiền tố của xâu c
 - o Xâu b là hậu tố của xâu c
 - o Độ dài xâu c là ngắn nhất.

Bài tập 2: Xâu tiền tố, hậu tố

- Cho một đoạn văn bản không quá 100 từ.
- Hãy chuẩn hóa đoạn văn bản sao cho:
 - o Hai đầu của đoạn văn bản không còn chứa dấu cách.
 - o Các từ trong đoạn văn bản chỉ cách nhau đúng một dấu cách.
 - o Ký tự đầu tiên của câu viết hoa, các ký tự khác viết thường.
 - o Trước dấu chấm, dấu phẩy không có dấu cách, sau dấu chấm, dấu phẩy có đúng một dấu cách.
- Hiển thị kết quả.

Bài tập 3: Thuật toán z

Cho một danh sách d chứa thông tin về các máy tính xách tay khác nhau được bày bán trong một cửa hàng, thông tin về mỗi máy tính xách tay gồm hãng sản xuất, thông tin mô tả (ví dụ: “CPU I7 12650H – RAM 16GB – 15.7 inches – new”), giá bán (USD). Cho một số tiền m (USD) và một xâu t là thông tin về kích thước màn hình của một máy tính (ví dụ “14 inches”).

Hãy viết chương trình thực hiện các yêu cầu:

- Cài đặt các thuật toán:
 - o Thuật toán Z xử lý xâu ký tự để biết một xâu s có phải là một xâu con của một xâu p hay không?
 - o Khởi tạo một danh sách gồm ít nhất 6 máy tính xách tay sao cho danh sách được sắp xếp theo thứ tự giảm dần của giá bán máy tính, khởi tạo số tiền m và khởi tạo xâu t. Áp dụng các thuật toán đã thiết kế ở trên để

tính sau đó hiển thị ra màn hình các giá trị x , q và giá trị y là số máy tính xách tay có kích thước màn hình t , hoặc đưa ra thông báo nếu không có phương án.