

BÀI THỰC HÀNH SỐ 7: 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 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 xâu ký tự s chỉ chứa các chữ cái, chữ số và dấu cách.

Output:

- Cho biết xâu ký tự p = “1000” có xuất hiện trong xâu 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 xâu p trong xâu s.
- Khởi tạo một xâu ký tự s.
- Sử dụng thuật toán Boyer Moore Horspool cho biết xâu p = “1000” có xuất hiện trong xâu s hay không? Thông báo kết quả.
- Thay thế xâu p = “1000” tìm được trong xâu s bằng xâu q = “một nghìn dong”, hiển thị xâu 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.