

So sánh văn bản

Phân loại văn bản là một vấn đề có ý nghĩa ứng dụng quan trọng. Xét bài toán cơ bản trong phân loại văn bản sau đây: Cho V là tập gồm m văn bản v_1, v_2, \dots, v_m và P là tập gồm n thuộc tính p_1, p_2, \dots, p_n . Đối với mỗi văn bản v_i người ta đã khảo sát xem nó có những thuộc tính nào trong số n thuộc tính trong P . Kết quả khảo sát đối với mỗi văn bản được ghi nhận dưới dạng một xâu gồm n ký tự '+' hoặc '-', trong đó ký tự thứ k là '+' nếu như văn bản có thuộc tính p_k và là '-' nếu như văn bản không có thuộc tính đó ($k = 1, 2, \dots, n$). Ta nói văn bản u là bao hàm văn bản v nếu như u có tất cả các thuộc tính như v và hơn nữa u phải có ít nhất một thuộc tính mà v không có.

Ví dụ:

- Nếu kết quả khảo sát 2 văn bản u và v tương ứng là "++----+" và "++-----" thì u bao hàm v .
- Nếu kết quả khảo sát 2 văn bản u và v tương ứng là "-+-----" và "++-----" thì v bao hàm u .
- Nếu kết quả khảo sát 2 văn bản u và v tương ứng là "+----++" và "+----++" thì u không bao hàm v và v không bao hàm u .
- Nếu kết quả khảo sát 2 văn bản u và v tương ứng là "+---+++" và "++----+" thì u không bao hàm v và v không bao hàm u .

Vấn đề đặt ra là: Dựa vào kết quả khảo sát cần xác định xem trong tập văn bản đã cho có bao nhiêu cặp văn bản (u, v) mà u bao hàm v .

Yêu cầu: Cho biết m, n và m xâu ghi nhận kết quả khảo sát, hãy tính số lượng cặp số (i, j) , $1 \leq i, j \leq m$ mà văn bản v_i bao hàm văn bản v_j .

Dữ liệu: Vào từ file văn bản TEXTCOM.INP:

- Dòng đầu tiên chứa hai số nguyên dương m và n ;
- Dòng thứ i trong số m dòng tiếp theo chứa xâu gồm n ký hiệu, mỗi ký hiệu là '+' hoặc '-' là kết quả khảo sát văn bản v_i , $i = 1, 2, \dots, m$.

Kết quả: Ghi ra file văn bản TEXTCOM.OUT một số nguyên duy nhất là số lượng cặp số tìm được.

TEXTCOM . INP	TEXTCOM . OUT
5 5	3
++++-	
++----	
+---++	
-----+	
----+-	

TEXTCOM . INP	TEXTCOM . OUT
3 4	0
-+--	
--++	
+--+	

Subtask 1: $2 \leq m \leq 200$; $1 \leq n \leq 10$;

Subtask 2: $2 \leq m \leq 200000$; $n = 3$;

Subtask 3: $2 \leq m \leq 200000$; $1 \leq n \leq 20$;