

ĐỀ THI CHÍNH THỨC

Đề thi có 04 trang

Môn: TIN HỌC

Thời gian: 300 phút (không kể thời gian giao đề)

Ngày thi thứ hai: 02/4/2021

TỔNG QUAN ĐỀ THI NGÀY THỨ HAI

Bài	Tên bài	Tên file chương trình	Hạn chế bộ nhớ
4	Mã ẩn	hcode.cpp	512 M
5	Đánh giá lỗi	ecount.cpp	512 M
6	Biến đổi gen	tgen.cpp	512 M

*Lập trình giải các bài toán sau đây:***Bài 4. (100 điểm) Mã ẩn**

Xét một dãy nhị phân A gồm n phần tử, phần tử A_i ($1 \leq i \leq n$) chỉ nhận giá trị 0 hoặc 1. Vì lí do bảo mật, ban đầu tất cả các phần tử của dãy A đều được ẩn giá trị và chỉ biết trong dãy A có chính xác k giá trị 0 ($0 \leq k \leq n$).

Thời điểm 0 là thời điểm bắt đầu xác định dãy A . Có m thông tin về mối quan hệ giữa các cặp phần tử trong dãy A sẽ lần lượt xuất hiện, thông tin thứ t ($1 \leq t \leq m$) xuất hiện tại thời điểm t . Cấu trúc của thông tin thứ t được mô tả thông qua ba giá trị u_t, c_t, v_t . Trong đó, u_t, v_t tương ứng với chỉ số của hai phần tử trong dãy A ($1 \leq u_t < v_t \leq n$) và c_t là một trong các kí tự $>$, $<$ hoặc $=$ để biểu diễn mối quan hệ giữa A_{u_t} và A_{v_t} . Nếu c_t là kí tự $>$ thì $A_{u_t} > A_{v_t}$, nếu c_t là kí tự $<$ thì $A_{u_t} < A_{v_t}$, còn c_t là kí tự $=$ thì $A_{u_t} = A_{v_t}$. Các thông tin đều bảo đảm tính chính xác, hợp lí trên dãy A .

Cần tìm thời điểm s ($0 \leq s \leq m$) nhỏ nhất để xác định duy nhất một cách gán giá trị cho tất cả n phần tử của dãy A sao cho có chính xác k giá trị 0 và thỏa mãn s thông tin đầu tiên.

Yêu cầu: Cho các thông tin về dãy A , hãy đưa ra thời điểm s nhỏ nhất sao cho dãy A được xác định duy nhất. Trong trường hợp sử dụng tất cả m thông tin mà vẫn có nhiều hơn một cách gán giá trị cho tất cả n phần tử của dãy A thì đưa ra -1 .

Dữ liệu: Vào từ thiết bị vào chuẩn:

- Dòng đầu tiên chứa một số nguyên dương q là số lượng bộ dữ liệu;
- Tiếp theo gồm q nhóm dòng, mỗi nhóm mô tả một bộ dữ liệu theo khuôn dạng sau:
 - Dòng thứ nhất chứa ba số nguyên n, m và k cách nhau bởi dấu cách ($1 \leq m \leq 4 \times 10^5$; $0 \leq k \leq n$);
 - Dòng thứ t trong số m dòng tiếp theo ($1 \leq t \leq m$) chứa ba giá trị u_t, c_t, v_t mô tả thông tin xuất hiện tại thời điểm t . Các giá trị cách nhau đúng một dấu cách.

Dữ liệu vào đảm bảo tồn tại ít nhất một dãy A có chính xác k giá trị 0 thỏa mãn tất cả m thông tin. Tổng các số m trong q bộ dữ liệu không quá 2×10^6 .

Kết quả: Ghi ra thiết bị ra chuẩn q dòng, mỗi dòng chứa một số nguyên s tương ứng với thời điểm nhỏ nhất để xác định duy nhất một cách gán giá trị cho tất cả n phần tử của dãy A trong dữ liệu vào tương ứng. Nếu không tìm được thời điểm s thỏa mãn thì ghi ra -1 .

Gọi N là tổng các số n trong q bộ dữ liệu.

Subtask 1 (30 điểm): $1 \leq n \leq 2 \times 10^3$ và $N \leq 10^4$;

Subtask 2 (32.5 điểm): $1 \leq n \leq 2 \times 10^4$ và $N \leq 10^5$;

Subtask 3 (37.5 điểm): $1 \leq n \leq 3 \times 10^5$ và $N \leq 1.5 \times 10^6$.

Ví dụ:

Dữ liệu	Kết quả	Giải thích
3	5	Trong dãy thứ nhất, tại thời điểm $s = 5$ xác định duy nhất dãy $\{1, 0, 1, 0, 1, 0\}$ thỏa mãn các thông tin tính đến thời điểm 5. Với thời điểm $t = 4$ có thêm ít nhất một dãy $\{1, 0, 0, 1, 0, 1\}$ thỏa mãn các thông tin tính đến thời điểm 4.
6 6 3	0	
1 > 2	-1	
3 = 5		Trong dãy thứ hai, không cần xét đến các thông tin bổ sung mà vẫn xác định dãy duy nhất $\{0, 0, 0, 0\}$. Do đó, đưa ra s bằng 0.
4 = 6		
1 > 2		
2 = 4		Trong dãy thứ ba, mặc dù sử dụng hết 2 thông tin nhưng vẫn có ít nhất hai dãy thỏa mãn tất cả 2 thông tin là: $\{0, 0, 1, 1\}$ và $\{1, 1, 0, 0\}$. Do đó, đưa ra -1.
1 = 3		
4 2 4		
1 = 2		
3 = 4		
4 2 2		
1 = 2		
3 = 4		

Bài 5. (100 điểm) Đánh giá lỗi

Một hệ thống tự động đóng hộp trà vừa được đưa vào vận hành. Hệ thống đã đóng được n hộp trà, các hộp được đánh số từ 1 đến n . Theo quy chuẩn, mỗi hộp trà nặng 1 kg và chứa đúng m gói trà nhỏ bên trong (tức là mỗi gói trà nặng $\frac{1}{m}$ kg). Do lỗi phát sinh trong quá trình đóng hộp, một số hộp trà chứa $m + 1$ gói, vì thế những hộp trà này sẽ nặng $(1 + \frac{1}{m})$ kg và được gọi là những hộp trà bị lỗi. Những hộp trà bị lỗi vẫn sẽ được bán ra, nhưng để đánh giá khoản lỗ gây ra bởi hệ thống, Ban quản lí muốn biết có bao nhiêu hộp trà bị lỗi trong số n hộp trà đã đóng. Việc đếm số lượng các hộp trà lỗi có thể thực hiện bằng cách cân thử ngay trên hệ thống. Tuy nhiên, vì chỉ được thiết kế để cân các hộp trà đúng quy chuẩn nên cân chỉ hiển thị được các số nguyên, số hiển thị trên cân là phần nguyên (làm tròn xuống) của tổng khối lượng thực tế các hộp được cân.

Tương tác

Bạn được cung cấp thư viện trong file `ecountlib.h` có các hàm và thủ tục sau:

- **int get_n()**
Hàm này trả về số lượng hộp trà đã được đóng hộp.
- **int get_m()**
Hàm này trả về số lượng gói trà trong một hộp chuẩn.
- **int get_w(vector<int> S)** với S là một tập con của $\{1, 2, 3, \dots, n\}$
Hàm này trả về phần nguyên của tổng khối lượng các hộp trà trong tập S . Lưu ý, S chỉ được phép chứa các số nguyên dương phân biệt trong đoạn từ 1 đến n .
- **void answer(int E)**
Thủ tục này trả lời số hộp trà lỗi là E hoặc giá trị -1 nếu không thể xác định số hộp trà bị lỗi.

Bạn phải viết một chương trình có khai báo `#include "ecountlib.h"` ở đầu chương trình để sử dụng các hàm tương tác trên. Bạn chỉ được cân không quá 400 lần. Chú ý, trong quá trình tương tác, nếu bạn gọi đến hàm `get_w()` quá 400 lần, hoặc gọi đến hàm `get_w()` với tham số không phù hợp, hoặc gọi đến thủ tục `answer()` thì chương trình sẽ tự động kết thúc. Xem thêm chương trình mẫu trong mục file đính kèm để hiểu hơn về cách tương tác với hệ thống.

Subtask 1 (13 điểm): $n \leq 8$ và $m \leq 10^9$;

Subtask 2 (17 điểm): $n \leq 300$ và $m \leq 10^9$;

Subtask 3 (17 điểm): $n \leq 10^4$ và $m \leq 20$;

Subtask 4 (19 điểm): $n \leq 10^4$ và $m \leq 40$;

Subtask 5 (23 điểm): $n \leq 2000$ và $m \leq 10^9$;

Subtask 6 (11 điểm): $n \leq 10^4$ và $m \leq 10^9$.

Cách chấm điểm

Có tất cả 6 subtask như mô tả ở trên. Với mỗi subtask, bạn sẽ nhận được đầy đủ điểm nếu tất cả các test trong subtask đó đều được đưa ra câu trả lời đúng, ngược lại bạn sẽ được 0 điểm của subtask đó.

Ví dụ:

Lời gọi hàm/thủ tục	Kết quả trả về
<code>get_n()</code>	3
<code>get_m()</code>	2
<code>get_w({1, 2})</code>	3
<code>get_w({2, 3})</code>	2
<code>answer(2)</code>	chương trình tự động kết thúc

Bài 6. (100 điểm) Biến đổi gen

Khi nghiên cứu quá trình biến đổi gen của một loại virus mới, Alice mô hình hóa bộ gen của virus là một dãy số nguyên dương và sự biến đổi gen là một dãy các thao tác thêm và xóa phần tử ở một số vị trí. Sau một thời gian nghiên cứu, giải mã và thu thập dữ liệu, Alice đã có một cặp dãy số, trong đó, dãy thứ nhất là dãy số mô tả bộ gen của virus trước khi biến đổi, dãy thứ hai là dãy số mô tả bộ gen sau quá trình biến đổi. Alice muốn dựa vào cặp dãy số để phân tích, tìm quy luật biến đổi gen của loại virus mới này. Đây là một vấn đề rất phức tạp, trước mắt Alice muốn tìm ra một dãy ít nhất các thao tác, mỗi thao tác là một hành động thêm hoặc xóa phần tử trên dãy thứ nhất để nhận được dãy thứ hai. Dựa trên các kết quả nghiên cứu trước, Alice sẽ chỉ quan tâm đến câu trả lời nếu số thao tác ít nhất cần sử dụng là không lớn hơn k .

Yêu cầu: Cho cặp dãy số mô tả bộ gen trước và sau khi biến đổi, hãy tìm dãy các thao tác thêm hoặc xóa phần tử thỏa mãn yêu cầu của Alice.

Dữ liệu: Vào từ thiết bị vào chuẩn:

- Dòng đầu tiên chứa một số nguyên dương q ($q \leq 5$) là số bộ dữ liệu.
- Tiếp theo là q nhóm dòng, mỗi nhóm dòng mô tả một bộ dữ liệu theo khuôn dạng:

- Dòng thứ nhất chứa ba số nguyên dương m, n, k ($m, n \leq 5 \times 10^5; k \leq 10^4$), trong đó m là độ dài dãy số mô tả bộ gen trước khi biến đổi, n là độ dài dãy số mô tả bộ gen sau quá trình biến đổi, k là giới hạn trên của số thao tác ít nhất mà Alice quan tâm;
- Dòng thứ hai chứa m số nguyên dương là dãy số mô tả bộ gen trước khi biến đổi, mỗi phần tử có giá trị không vượt quá 10^4 ;
- Dòng thứ ba chứa n số nguyên dương là dãy số mô tả bộ gen sau quá trình biến đổi, mỗi phần tử có giá trị không vượt quá 10^4 .

Hai số liên tiếp trên cùng một dòng được ghi cách nhau bởi dấu cách.

Kết quả: Ghi ra thiết bị ra chuẩn q nhóm dòng, mỗi nhóm là kết quả tương ứng với một bộ dữ liệu vào. Nếu số thao tác lớn hơn k thì chỉ ghi ra số -1 trên một dòng duy nhất, ngược lại ghi ra theo khuôn dạng:

- Dòng thứ nhất chứa số nguyên s , trong đó s là số thao tác thêm hoặc xóa ít nhất để biến đổi từ dãy số mô tả bộ gen ban đầu thành dãy số mô tả bộ gen sau quá trình biến đổi;
- Dòng thứ i trong s dòng tiếp theo là thao tác thứ i trên dãy số mô tả bộ gen trước khi biến đổi. Một thao tác có thể là thêm một số nguyên dương x vào vị trí p trên dãy, hoặc là xóa phần tử ở vị trí p của dãy. Gọi n' là độ dài của dãy trước khi thực hiện thao tác thứ i :
 - Nếu thao tác là thêm, ghi ra một dòng theo định dạng " $add\ p\ x$ ". Thao tác được coi là hợp lệ nếu $1 \leq p \leq n' + 1$ và $1 \leq x \leq 10^4$;
 - Nếu thao tác là xóa, ghi ra một dòng theo định dạng " $del\ p$ ". Thao tác được coi là hợp lệ nếu $1 \leq p \leq n'$.

Subtask 1 (10 điểm): Tồn tại đáp án tối ưu chỉ sử dụng nhiều nhất một thao tác xóa;

Subtask 2 (20 điểm): $m, n \leq 50000$ và $k \leq 1000$;

Subtask 3 (20 điểm): $m, n \leq 10^5$ và $k \leq 5000$;

Subtask 4 (30 điểm): $m, n \leq 4 \times 10^5$;

Subtask 5 (20 điểm): Không có ràng buộc nào thêm.

Ví dụ:

Dữ liệu	Kết quả	Giải thích
2 7 6 6 3 2 1 2 1 3 2 1 2 3 1 2 3 3 3 5 1 1 1 2 2 2	5 del 1 del 1 del 3 add 4 1 add 6 3 -1	Trong bộ dữ liệu thứ nhất, cần ít nhất 5 phép biến đổi, quá trình biến đổi như sau: $(3, 2, 1, 2, 1, 3, 2) \rightarrow (2, 1, 2, 1, 3, 2) \rightarrow (1, 2, 1, 3, 2)$ $\rightarrow (1, 2, 3, 2) \rightarrow (1, 2, 3, 1, 2) \rightarrow (1, 2, 3, 1, 2, 3)$ Trong bộ dữ liệu thứ hai cần ít nhất 6 phép biến đổi, số lượng phép biến đổi ít nhất vượt quá giới hạn $k = 5$ nên ghi ra -1 .

----- HẾT -----

- *Thí sinh không được sử dụng tài liệu.*
- *Giám thị không giải thích gì thêm.*