

KHẢO SÁT CHẤT LƯỢNG HỌC SINH GIỎI
NĂM HỌC 2024-2025

MÔN: TIN HỌC

Thời gian: 180 phút (không kể thời gian giao đề)

TỔNG QUAN BÀI THI

Đề thi có 04 trang. Học sinh dùng ngôn ngữ lập trình C/C++ hoặc Python viết chương trình đọc file dữ liệu vào, xử lý thông tin rồi ghi kết quả vào file dữ liệu ra.

Bài	Tên file mã nguồn	Tên file dữ liệu vào	Tên file dữ liệu ra	Bộ nhớ	Điểm
Bài 1	HSG.*	HSG.INP	HSG.OUT	1024 MB	6.0
Bài 2	PACK.*	PACK.INP	PACK.OUT	1024 MB	7.0
Bài 3	MATRIX.*	MATRIX.INP	MATRIX.OUT	1024 MB	8.0
Bài 4	TOUR.*	TOUR.INP	TOUR.OUT	1024 MB	9.0

Chú ý:

- Bài thi được làm trên máy tính. Học sinh nộp bài thi là các file mã nguồn.
- Dấu * trong tên file mã nguồn là **cpp** hoặc **py** tương ứng với ngôn ngữ **C++** hoặc **Python**;
- Đặt tên file mã nguồn theo đúng quy định của từng bài, không ghi thông tin cá nhân (họ tên, số báo danh, ngày sinh, trường,...) và các ký hiệu khác thường vào file bài làm;
- File dữ liệu vào và file dữ liệu ra ở cùng thư mục với file mã nguồn, học sinh không khai báo đường dẫn đến file dữ liệu vào và file dữ liệu ra;

ĐỀ BÀI

Bài 1. Chữ HSG

Trong kì thi chọn học sinh giỏi năm nay, Việt và Nam dự thi môn toán. Hai người thích chữ viết tắt "**HSG**" nên Việt đã đồ Nam bài toán như sau:

Cho dãy kí tự X có n kí tự chữ cái Latin viết in hoa hoặc dấu cách trông $x_1x_2 \dots x_n$, hỏi có bao nhiêu cách chọn ra 3 kí tự "**H**", "**S**" và "**G**" theo thứ tự xuất hiện trước/sau ban đầu trong dãy kí tự X để ghép thành chữ "**HSG**", hay nói cách khác là đếm số bộ chỉ số (i, j, k) thỏa mãn: $1 \leq i < j < k \leq n$ và $\overline{x_i x_j x_k} = HSG$

Yêu cầu: viết chương trình giúp Nam giải quyết bài toán trên.

Dữ liệu: vào từ file văn bản **HSG.INP** có cấu trúc:

- ✓ Dòng đầu tiên là số nguyên dương n ($n \leq 10^6$);
- ✓ Dòng thứ hai có n kí tự chữ cái Latin viết in hoa hoặc dấu cách trông $x_1x_2 \dots x_n$.

Kết quả: ghi ra file **HSG.OUT** một số nguyên là số cách chọn ra bộ kí tự "**HSG**".

Ví dụ:

HSG.INP	HSG.OUT	Giải thích
20 HUNG THI HSG MON TIN	3	Các bộ chỉ số thỏa mãn là: (1, 11, 12), (7, 11, 12), (10, 11, 12)

Ràng buộc:

- ✓ 40% số điểm tương ứng với 40% số test có: $n \leq 10^2$ và X không có dấu cách;
- ✓ 40% số điểm tương ứng với 40% số test có: $n \leq 10^4$;
- ✓ 20% số test còn lại không có ràng buộc gì thêm.

Bài 2. Đóng gói

Kho hàng của một công ty vận chuyển có n kiện hàng đang đóng gói. Khối lượng hiện tại của n kiện hàng lần lượt là các số nguyên dương $a_1, a_2, \dots, a_n$. Công ty đang có m gói hàng chờ để bổ sung vào n kiện hàng. Khối lượng của m gói hàng lần lượt là các số nguyên dương $b_1, b_2, \dots, b_m$. Bờm có nhiệm vụ xếp lần lượt m gói hàng vào n kiện hàng theo quy ước: ở lượt xếp gói hàng thứ i ($i = 1 \div m$) chỉ chọn một trong hai cách sau:

- ✓ Bỏ qua (không xếp) gói hàng thứ i (có khối lượng b_i);
- ✓ Xếp gói hàng thứ i vào kiện hàng thứ j ($j = 1 \div n$) nếu từ kiện hàng thứ j đến kiện hàng thứ n chưa từng được xếp gói hàng nào trước đó.

Công việc kết thúc sau khi Bờm xếp xong gói hàng thứ m hoặc kiện hàng thứ n . Ban quản lý lấy khối lượng nhỏ nhất của n kiện hàng để tính tiền thưởng cho Bờm. Vì có nhiều cách xếp các gói hàng nên để có tiền thưởng cao hơn thì Bờm chọn cách xếp để kiện hàng có khối lượng nhỏ nhất đạt giá trị lớn nhất.

Yêu cầu: tính giá trị lớn nhất có thể của kiện hàng có khối lượng nhỏ nhất.

Dữ liệu: vào từ file văn bản **PACK.INP** có cấu trúc:

- ✓ Dòng đầu tiên có 2 số nguyên dương n và m ($n, m \leq 10^5$);
- ✓ Dòng thứ hai có n số nguyên dương $a_1, a_2, \dots, a_n$ ($a_i \leq 10^{12} \forall i = 1 \div n$);
- ✓ Dòng thứ ba có m số nguyên dương $b_1, b_2, \dots, b_m$ ($b_i \leq 10^{12} \forall i = 1 \div m$).

Các số nguyên trong cùng dòng dữ liệu cách nhau bởi dấu cách trống.

Kết quả: ghi ra file văn bản **PACK.OUT** một số là kết quả tìm được.

Ví dụ:

PACK.INP	PACK.OUT	Giải thích
6 4 2 5 1 6 4 7 1 3 4 2	5	✓ Bỏ qua gói hàng 1 ✓ Xếp gói hàng 2 vào kiện hàng 1 ✓ Xếp gói hàng 3 vào kiện hàng 3 ✓ Xếp gói hàng 4 vào kiện hàng 5
7 4 4 7 2 8 9 3 5 5 9 7 6	7	✓ Xếp gói hàng 1 vào kiện hàng 1 ✓ Xếp gói hàng 2 vào kiện hàng 3 ✓ Xếp gói hàng 3 vào kiện hàng 6 ✓ Xếp gói hàng 4 vào kiện hàng 7

Ràng buộc:

- ✓ 20% số điểm tương ứng với 20% số test có: $m = 1; n \leq 10^3$;
- ✓ 20% số điểm tương ứng với 20% số test có: $m = 2; n \leq 10^3$;
- ✓ 30% số điểm tương ứng với 30% số test có: $n, m \leq 10^3$, các kiện hàng và gói hàng có khối lượng tối đa là 10^6 ;
- ✓ 30% số test còn lại không có ràng buộc gì thêm.

Bài 3. Ma trận

Cho ma trận số nguyên dương có m hàng và n cột. Ô giao giữa hàng i ($i = 1 \div m$) và cột j ($j = 1 \div n$) gọi là ô (i, j) có giá trị là số nguyên a_{ij} ($1 \leq a_{ij} \leq n$) thỏa mãn:

- ✓ Hàng thứ nhất: các số trong hàng đôi một khác nhau;
- ✓ Các hàng khác: các số trong hàng có thể giống nhau.

Thực hiện xóa một số cột của ma trận sao cho các hàng của ma trận còn lại sau khi được sắp xếp các số theo thứ tự tăng dần thì giống nhau.

Yêu cầu: tính số cột tối thiểu cần xóa theo quy tắc trên.

Dữ liệu: vào từ file văn bản **MATRIX.INP** có cấu trúc:

- ✓ Dòng đầu tiên có 2 số nguyên dương m ($m \leq 10$) và n ($n \leq 10^5$);
- ✓ m dòng tiếp theo mỗi dòng có n số nguyên có giá trị thuộc đoạn $[1, n]$; ở dòng thứ nhất các số đôi một khác nhau.

Các số nguyên trong cùng dòng dữ liệu cách nhau bởi dấu cách trống.

Kết quả: ghi ra file văn bản **MATRIX.OUT** một số là số cột tối thiểu cần xóa.

Ví dụ:

MATRIX.INP	MATRIX.OUT	Giải thích
2 8 1 2 3 4 5 6 7 8 2 5 2 3 4 4 3 1	4	Xóa cột: 1, 6, 7, 8
3 8 5 4 3 2 1 6 7 8 5 5 1 1 3 4 7 3 3 7 1 4 5 6 2 3	5	Xóa cột: 2, 4, 6, 7, 8

Ràng buộc:

- ✓ 30% số điểm tương ứng với 30% test có: $m = 2$; $n \leq 10^3$;
- ✓ 40% số điểm tương ứng với 40% test có: $m = 3$; $n \leq 10^3$;
- ✓ 30% số test còn lại không có ràng buộc gì thêm.

Bài 4. Du lịch

Sau kì thi học sinh giỏi, Việt đi du lịch ở thành phố HP - nơi có n địa điểm du lịch được đánh số từ 1 đến n , giá vé của mỗi địa điểm là c_i ($i = 1 \div n$). Có m con đường 2 chiều, mỗi con đường kết nối 2 địa điểm du lịch u và v ($1 \leq u, v \leq n$; $u \neq v$). Ban quản lý lập danh sách các tour du lịch, mỗi tour gồm tất cả các địa điểm có đường đi trực tiếp hoặc gián tiếp đến nhau. Chi phí một tour bằng tổng tiền vé của các địa điểm trong tour.

Là người thích trải nghiệm đến nhiều vùng khác nhau nên Việt chọn k tour du lịch và quan tâm đến độ chênh lệch chi phí nhỏ nhất của 2 tour bất kì trong k tour. Việt muốn tính giá trị tối đa của độ chênh lệch chi phí nhỏ nhất trong các cách chọn k tour.

Yêu cầu: hãy giúp Việt tính giá trị tối đa trên.

Dữ liệu: vào từ file văn bản **TOUR.INP** có cấu trúc:

- ✓ Dòng đầu tiên có 3 số nguyên dương n, m, k ($n, m, k \leq 10^5$);

- ✓ Dòng thứ hai có n số nguyên dương $c_1, c_2, \dots, c_n$ ($c_i \leq 10^9, \forall i = 1 \div n$) là giá vé du lịch của các địa điểm.
- ✓ m dòng tiếp theo, mỗi dòng có 2 số nguyên dương u và v mô tả có một con đường 2 chiều nối 2 địa điểm du lịch u và v ($u, v \leq n; u \neq v$).

Các số nguyên trong cùng dòng dữ liệu cách nhau bởi dấu cách trống.

Kết quả: ghi ra file văn bản **TOUR.OUT** một số là giá trị tối đa của độ chênh lệch chi phí nhỏ nhất tìm được.

Ví dụ:

TOUR.INP	TOUR.OUT	Giải thích																					
9 5 4 1 2 3 5 7 4 4 6 8 4 3 5 6 8 9 6 5 3 4	3	Có 6 tour: <table border="1"> <thead> <tr> <th>Tour</th><th>Địa điểm</th><th>Chi phí</th></tr> </thead> <tbody> <tr> <td>1</td><td>1</td><td>1</td></tr> <tr> <td>2</td><td>2</td><td>2</td></tr> <tr> <td>3</td><td>3, 4</td><td>8</td></tr> <tr> <td>4</td><td>5, 6</td><td>11</td></tr> <tr> <td>5</td><td>7</td><td>4</td></tr> <tr> <td>6</td><td>8, 9</td><td>14</td></tr> </tbody> </table> Có nhiều cách chọn 4 tour; Chọn: {1,3,4,5} hoặc {2,3,4,6}	Tour	Địa điểm	Chi phí	1	1	1	2	2	2	3	3, 4	8	4	5, 6	11	5	7	4	6	8, 9	14
Tour	Địa điểm	Chi phí																					
1	1	1																					
2	2	2																					
3	3, 4	8																					
4	5, 6	11																					
5	7	4																					
6	8, 9	14																					

Ràng buộc:

- ✓ 20% số điểm tương ứng với 20% test có:

$$k = 3; n, m \leq 10^3; |u_i - v_i| = 1 (i = 1 \div m);$$

- ✓ 30% số điểm tương ứng với 30% test có: $k = 3; n, m \leq 10^3$;

- ✓ 30% số điểm tương ứng với 30% test có: $n, m, k \leq 10^3; c_i \leq 10^4 (i = 1 \div n)$;

- ✓ 20% số test còn lại không có ràng buộc gì thêm.

----- HẾT -----

Thí sinh không sử dụng tài liệu. Cán bộ coi thi không giải thích gì thêm.

Họ và tên: **Số báo danh:**

CBCT số 1:..... CBCT số 2: