

**ĐỀ THI CHÍNH THỨC**Môn: **TIN HỌC**Thời gian: **180 phút** (*không kể thời gian giao đề*)Ngày thi: **04/3/2022**

Đề thi gồm 04 trang, 03 bài

**TỔNG QUAN ĐỀ THI**

|              | Tên bài                 | File chương trình | File dữ liệu vào    | File kết quả        |
|--------------|-------------------------|-------------------|---------------------|---------------------|
| <b>Bài 1</b> | <b>Chọn cặp</b>         | <b>PAIR.*</b>     | <b>PAIR.INP</b>     | <b>PAIR.OUT</b>     |
| <b>Bài 2</b> | <b>Đặc trưng đồ thị</b> | <b>GRAPH.*</b>    | <b>GRAPH.INP</b>    | <b>GRAPH.OUT</b>    |
| <b>Bài 3</b> | <b>Kết nối internet</b> | <b>INTERNET.*</b> | <b>INTERNET.INP</b> | <b>INTERNET.OUT</b> |

Dấu \* được thay thế bởi PAS hoặc CPP của ngôn ngữ lập trình sử dụng tương ứng là Pascal hoặc C++.

*Hãy lập trình giải các bài toán sau:*

**Bài 1. Chọn cặp (7 điểm)**

Hưởng ứng phong trào thi đua học tập của trường, cô giáo chủ nhiệm lớp của Nam muốn chọn các cặp đôi bạn giúp đỡ nhau cùng tiến. Lớp của Nam có  $n$  học sinh, được đánh số từ 1 đến  $n$ , học sinh thứ  $i$  có chỉ số học lực  $a_i$ . Để có được sự cân bằng giữa các cặp bạn cùng tiến, cô giáo muốn chọn các cặp bạn có tổng chỉ số học lực đôi một giữa các cặp chênh nhau không quá một giá trị nhỏ  $d$  sao cho mỗi bạn xuất hiện trong không quá một cặp. Cô giáo mong muốn chọn được nhiều cặp như vậy nhất.

**Yêu cầu:** Cho biết số học sinh của lớp và chỉ số học lực của từng học sinh, hãy tính số cặp bạn cùng tiến nhiều nhất mà có tổng chỉ số học lực đôi một giữa các cặp chênh nhau không quá  $d$ .

**Dữ liệu:** Vào từ file văn bản **PAIR.INP**:

- Dòng đầu tiên chứa hai số nguyên không âm  $n, d$  ( $n \geq 2$ );
- Dòng thứ hai chứa  $n$  số nguyên dương, số thứ  $i$  là chỉ số học lực  $a_i$  của học sinh thứ  $i$  ( $1 \leq i \leq n$ ;  $a_i \leq 10^9$ ).

Các số trên cùng một dòng cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **PAIR.OUT** một số nguyên duy nhất là số cặp nhiều nhất tìm được.

**Ràng buộc:**

- Có 20% số test ứng với 20% số điểm của bài thỏa mãn:  $n \leq 10$ ;  $d = 0$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $n \leq 200$ ;  $d = 0$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $n \leq 2000$ ;  $d = 0$ ;
- 30% số test khác ứng với 30% số điểm của bài thỏa mãn:  $n \leq 200$ ;  $d = 1$ ;
- 10% số test còn lại ứng với 10% số điểm của bài thỏa mãn:  $n \leq 2000$ ;  $d = 1$ .

Ví dụ:

| PAIR . INP           | PAIR . OUT | Giải thích  |
|----------------------|------------|---|
| 7 1<br>9 1 2 4 5 6 8 | 3          | Một phương án chọn được nhiều nhất cặp các bạn học sinh với tổng học lực đôi một chênh nhau không quá 1 là:<br>$1 + 8 = 9; 2 + 6 = 8; 4 + 5 = 9.$<br>Một phương án khác cũng chọn được 3 cặp là:<br>$1 + 9 = 10; 2 + 8 = 10; 4 + 6 = 10.$ |

## Bài 2. Đặc trưng đồ thị (7 điểm)

Khi nghiên cứu về lý thuyết đồ thị, Nam đã tìm ra một đặc trưng cho đồ thị vô hướng. Cụ thể, với đồ thị vô hướng  $G$  gồm  $n$  đỉnh, xét lần lượt từng đỉnh từ 1 đến  $n$ , với đỉnh  $i$ , Nam tính được số  $a_i$  bằng số lượng đỉnh  $j$ , thỏa mãn  $1 \leq j < i$ , mà  $j$  có cạnh nối với  $i$ . Dãy số nguyên  $a_1, a_2, \dots, a_n$  được gọi là dãy đặc trưng của đồ thị. Dễ nhận thấy rằng một đồ thị chỉ có duy nhất một dãy đặc trưng, tuy nhiên, có thể có nhiều đồ thị khác nhau nhưng có cùng một dãy đặc trưng.

Bằng đi một thời gian, một hôm, Nam tìm lại được một file văn bản ghi một dãy số là một dãy đặc trưng của một đồ thị vô hướng. Tuy nhiên, dãy bị khuyết mất một số số và biết rằng bậc mỗi đỉnh của đồ thị này đều nhỏ hơn hoặc bằng  $b$ . Nam muốn tính số lượng cách điền các số vào các vị trí bị khuyết để nhận được dãy  $f$  là dãy đặc trưng biết rằng:

- 1) Tồn tại ít nhất một đồ thị có dãy đặc trưng là  $f$ ;
- 2) Tất cả các đồ thị có dãy đặc trưng là  $f$  đều có bậc của mỗi đỉnh trong đồ thị nhỏ hơn hoặc bằng  $b$ .

**Yêu cầu:** Cho số nguyên  $b$  và dãy đặc trưng bị khuyết, gọi  $S$  là số lượng cách điền các số vào các vị trí bị khuyết để nhận được dãy  $f$  là dãy đặc trưng thỏa mãn, hãy tính  $S \% (10^9 + 7)$ , trong đó  $\%$  là phép toán chia lấy dư. Hai cách điền gọi là khác nhau nếu như tồn tại một vị trí khuyết mà giá trị điền vào trong cách này khác với giá trị điền vào trong cách kia.

**Dữ liệu:** Vào từ file văn bản **GRAPH . INP**:

- Dòng đầu tiên chứa hai số nguyên dương  $n, b$  ( $b \leq n$ );
- Dòng thứ hai chứa  $n$  số nguyên, số thứ  $i$  có giá trị  $a_i$  ( $1 \leq i \leq n$ ;  $0 \leq a_i \leq i - 1$  hoặc  $a_i = -1$ ) mô tả dãy đặc trưng. Giá trị  $a_i = -1$  có nghĩa là vị trí thứ  $i$  bị khuyết.

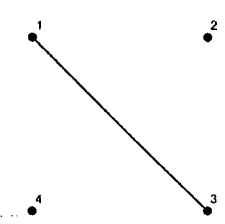
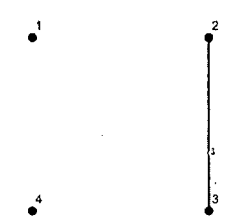
Các số trên cùng một dòng cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **GRAPH . OUT** một số nguyên duy nhất là giá trị  $S \% (10^9 + 7)$ .

**Ràng buộc:**

- Có 25% số test ứng với 25% số điểm của bài thỏa mãn:  $n \leq 6$ ;
- 30% số test khác ứng với 30% số điểm của bài thỏa mãn:  $n \leq 200$  và chỉ có duy nhất một ô bị khuyết;
- 25% số test khác ứng với 25% số điểm của bài thỏa mãn:  $n \leq 200$ ;
- 20% số test còn lại ứng với 20% số điểm của bài thỏa mãn:  $n \leq 2000$ .

Ví dụ:

| GRAPH . INP                 | GRAPH . OUT | Giải thích   |
|-----------------------------|-------------|--|
| <pre> 4 1 -1 -1 1 -1 </pre> | 1           | <p>Có duy nhất một dãy đặc trưng thỏa mãn là <math>(0, 0, 1, 0)</math> và có tất cả 2 đồ thị cùng có dãy đặc trưng này như trong hình sau:</p> <div style="display: flex; justify-content: space-around; align-items: flex-end;"> <div style="text-align: center;">  <p>Đồ thị 1</p> </div> <div style="text-align: center;">  <p>Đồ thị 2</p> </div> </div> |

### Bài 3. Kết nối internet (6 điểm)

Thành phố nơi Nam sinh sống đã xây dựng hệ thống Wi-Fi công cộng để phục vụ khách du lịch và người dân trong thành phố. Hệ thống bao gồm  $n$  trạm phát sóng được đánh số từ 1 đến  $n$ , trạm thứ  $i$  được đặt ở vị trí tương ứng với tọa độ  $(x_i, y_i)$  trên bản đồ mặt phẳng tọa độ của thành phố. Các trạm phát sóng này đều có mức độ phủ sóng là  $s$ . Hai trạm phát sóng  $i$  và  $j$  có thể truyền thông tin cho nhau bằng sóng nếu  $|x_i - x_j| + |y_i - y_j| \leq s$ , hoặc trạm phát sóng  $i$  có thể truyền thông tin qua một số trạm phát sóng trung gian để tới được trạm phát sóng  $j$ . Một nhóm các trạm phát sóng gọi là liên thông nếu như hai trạm bất kỳ trong nhóm có thể truyền thông tin cho nhau. Mỗi nhóm liên thông chỉ cần một đường truyền cung cấp internet để tất cả các trạm phát sóng này đều có thể phát được mạng internet cho người dân sử dụng.

Tuy nhiên, sau một thời gian sử dụng, chi phí duy trì hoạt động của hệ thống này là rất lớn, trong đó chi phí sử dụng mạng internet hàng tháng là chiếm nhiều nhất. Chi phí sử dụng mạng internet tỉ lệ thuận với số đường truyền cung cấp internet. Do đó, chính quyền thành phố muốn giảm thiểu số lượng đường truyền cung cấp internet mà vẫn bảo đảm tất cả các trạm đều có mạng internet bằng cách thiết kế các đường dây cáp kết nối trực tiếp một số trạm phát sóng với nhau. Với hai trạm phát sóng  $u$  và  $v$  chưa truyền được thông tin cho nhau, khi kết nối trực tiếp chúng bằng đường dây cáp, hai trạm phát sóng này có thể truyền tin cho nhau, vì vậy, nếu một trong hai trạm có mạng internet thì trạm còn lại cũng có mạng internet và toàn bộ nhóm liên thông chứa trạm này đều phát được mạng internet, do đó, có thể giảm đi một đường truyền cung cấp internet. Chi phí để kết nối hai trạm phát sóng  $u$  và  $v$  bằng đường dây cáp là  $|x_u - x_v| + |y_u - y_v|$ . Chính quyền thành phố muốn xây dựng phương án giảm đi  $k$  đường truyền cung cấp internet mà vẫn bảo đảm trạm phát sóng nào cũng có mạng internet bằng cách sử dụng thêm  $k$  đường dây cáp kết nối trực tiếp sao cho tổng chi phí kết nối bằng dây cáp là nhỏ nhất.

**Yêu cầu:** Cho biết số lượng và vị trí của  $n$  trạm phát sóng, mức độ phủ sóng  $s$  của các trạm phát sóng và số lượng  $k$  đường truyền cung cấp internet cần giảm đi, hãy tính chi phí nhỏ nhất để kết nối các trạm phát sóng bằng dây cáp sao cho trạm phát sóng nào cũng có mạng internet.

**Dữ liệu:** Vào từ file văn bản **INTERNET . INP**:

- Dòng đầu tiên chứa ba số nguyên dương  $n, s, k$  ( $n \leq 10^5; s \leq 10^9; k \leq 20$ );
- Dòng thứ  $i$  trong số  $n$  dòng tiếp theo chứa hai số nguyên không âm  $x_i, y_i$  mô tả vị trí của của trạm phát sóng thứ  $i$  ( $1 \leq i \leq n; x_i, y_i \leq 10^9$ ). Dữ liệu bảo đảm không có hai trạm phát sóng nào có tọa độ trùng nhau và số lượng đường truyền cung cấp internet cần sử dụng cho hệ thống ban đầu lớn hơn  $k$ .

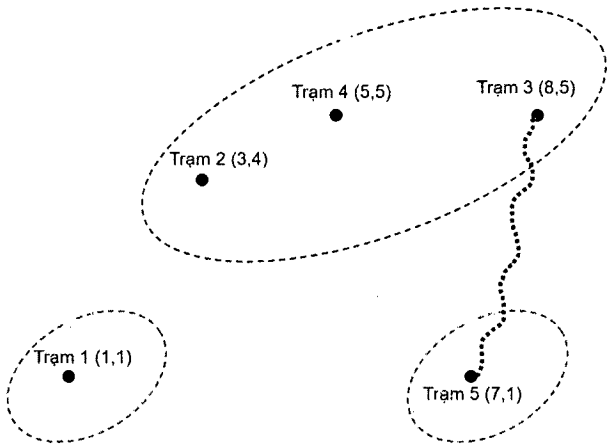
Các số trên cùng một dòng cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **INTERNET . OUT** một số nguyên duy nhất là chi phí nhỏ nhất để kết nối các trạm phát sóng thỏa mãn yêu cầu.

**Ràng buộc:**

- Có 20% số test ứng với 20% số điểm của bài thỏa mãn:  $n \leq 1000$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn tọa độ tất cả các trạm phát sóng đều nằm trên một đường thẳng;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn giá trị tọa độ mỗi trạm phát sóng không vượt quá 1000;
- 40% số test còn lại ứng với 40% số điểm của bài không có ràng buộc gì thêm.

**Ví dụ:**

| INTERNET . INP                           | INTERNET . OUT | Giải thích  |
|--|----------------|---|
| 5 4 1<br>1 1<br>3 4<br>8 5<br>5 5<br>7 1 | 5              | <p>Ban đầu có 3 nhóm liên thông là {Trạm 1}, {Trạm 2, Trạm 3, Trạm 4} và {Trạm 5}. Phương án cho chi phí nhỏ nhất để giảm bớt 1 đường truyền cung cấp internet là nối cáp giữa Trạm 3 và Trạm 5 cho chi phí bằng <math> 8-7 + 5-1 =5</math>.</p>  |

----- HẾT -----

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

**HƯỚNG DẪN CHẤM THI**Môn: **TIN HỌC****Đề thi chính thức**Thời gian: **180 phút** (*không kể thời gian giao đề*)Ngày thi: **04/3/2022**

Hướng dẫn chấm thi gồm 03 trang

**I. Hướng dẫn chung**

1. Bài thi của thí sinh được chấm trên máy tính bằng phần mềm chấm thi Themis (*bản quyền của Cục Quản lý chất lượng*) sử dụng bộ test do Hội đồng chấm thi phê duyệt, đúng với đáp án, biểu điểm của Bộ Giáo dục và Đào tạo.
2. Tổ trưởng Tổ chấm thi phân công giám khảo thực hiện quy trình chấm thi trên máy tính, bảo đảm các yêu cầu của Quy chế thi.
3. Điểm bài thi được xuất từ phần mềm chấm thi; không quy tròn điểm thành phần của từng câu, điểm của bài thi.

**II. Đáp án và biểu điểm****Bài 1. Chọn cặp (7 điểm)****Phân bổ điểm**

- Có 20% số test ứng với 20% số điểm của bài thỏa mãn:  $n \leq 10$ ;  $d = 0$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $n \leq 200$ ;  $d = 0$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $n \leq 2000$ ;  $d = 0$ ;
- 30% số test khác ứng với 30% số điểm của bài thỏa mãn:  $n \leq 200$ ;  $d = 1$ ;
- 10% số test còn lại ứng với 10% số điểm của bài thỏa mãn:  $n \leq 2000$ ;  $d = 1$ .

**Thuật toán**

*Nhận xét:* Xét dãy  $a$  đã được sắp xếp tăng dần. Với mỗi giá trị  $S$  là tổng hai phần tử của dãy  $a$ , ta xác định nhiều nhất các cặp có tổng hai số bằng  $S$  hoặc bằng  $S + 1$ . Giả sử số lượng nhiều nhất đạt được là  $k$ , ta sẽ có cách chọn các cặp là  $(i_1, j_1), (i_2, j_2), \dots, (i_k, j_k)$  thỏa mãn:  $i_1 < i_2 < \dots < i_k < j_k < \dots < j_2 < j_1$ .

Từ nhận xét trên ta có thuật toán tham lam như sau:

Bước 1: Sắp xếp dãy  $a$  tăng dần. Độ phức tạp bước này là  $O(n \log n)$ .

Bước 2: Xây dựng danh sách  $L$  lưu trữ các bộ  $(a_i + a_j, i, j)$  được sắp xếp theo thứ tự ưu tiên:

- 1) tổng  $(a_i + a_j)$  từ nhỏ đến lớn;
- 2) chỉ số  $i$  từ nhỏ đến lớn;
- 3) chỉ số  $j$  từ lớn đến nhỏ.

Do đó với mỗi  $S$ , các phần tử  $(a_i + a_j = S, i, j)$  trong danh sách  $L$  là liên tiếp nhau.

Độ phức tạp bước này là  $O(n^2 \log n)$ .

Bước 3: Sử dụng kĩ thuật tịnh tiến hai con trỏ (two pointers) để chọn ra được nhiều nhất các cặp thoả mãn đề bài. Độ phức tạp bước này là  $O(n^2)$ .

Độ phức tạp thuật toán là:  $O(n^2 \log n)$ .

## Bài 2. Đặc trưng đồ thị (7 điểm)

### Phân bổ điểm

- Có 25% số test ứng với 25% số điểm của bài thoả mãn:  $n \leq 6$ ;
- 30% số test khác ứng với 30% số điểm của bài thoả mãn:  $n \leq 200$  và chỉ có duy nhất một ô bị khuyết;
- 25% số test khác ứng với 25% số điểm của bài thoả mãn:  $n \leq 200$ ;
- 20% số test còn lại ứng với 20% số điểm của bài thoả mãn:  $n \leq 2000$ .

### Thuật toán

Với  $f = f_1, f_2, \dots, f_n$  là một dãy nhận được bằng cách điền vào các vị trí bị khuyết của dãy  $a$ , gọi  $s_i$  là số lượng phần tử khác 0 trong dãy  $f_{i+1}, f_{i+2}, \dots, f_n$ .

*Nhận xét:* Dãy  $f$  thoả mãn khi và chỉ khi  $f_i \leq \min(i - 1, b - s_i), \forall i = \overline{1, n}$ .

Từ nhận xét trên ta có cách xây dựng công thức quy hoạch động như sau: Gọi  $dp(i, k)$  là số lượng dãy  $f_{i+1}, f_{i+2}, \dots, f_n$  nhận được bằng cách điền vào các vị trí bị khuyết của dãy  $a$ , sao cho  $s_i = k$  và  $f_j \leq \min(j - 1, b - s_j), \forall j = \overline{i + 1, n}$ . Khi đó:

$$dp(n, 0) = 1;$$

$$dp(i, k) = \sum_{t=0}^{\min(i, b-k)} dp(i+1, k - [t \neq 0]), \text{ nếu } a_{i+1} = -1;$$

$$dp(i, k) = dp(i+1, k - [a_{i+1} \neq 0]), \text{ nếu } a_{i+1} \neq -1 \text{ và } a_{i+1} \leq b - k;$$

$$dp(i, k) = 0 \text{ các trường hợp còn lại.}$$

$$(\text{trong đó } [x \neq 0] \text{ bằng } 0 \text{ nếu } x = 0 \text{ và } [x \neq 0] \text{ bằng } 1 \text{ nếu } x \neq 0).$$

Độ phức tạp thuật toán:  $O(n^2)$ .

## Bài 3. Kết nối Internet (6 điểm)

### Phân bổ điểm

- Có 20% số test ứng với 20% số điểm của bài thoả mãn:  $n \leq 1000$ ;
- 20% số test khác ứng với 20% số điểm của bài thoả mãn toạ độ tất cả các trạm phát sóng đều nằm trên một đường thẳng;
- 20% số test khác ứng với 20% số điểm của bài thoả mãn giá trị toạ độ mỗi trạm phát sóng không vượt quá 1000;
- 40% số test còn lại ứng với 40% số điểm của bài không có ràng buộc gì thêm.

### Thuật toán

1. Xoay trục toạ độ  $Oxy$  thành trục toạ độ  $Ox^*y^*$  với  $x^* = x + y, y^* = x - y$ . Khi đó khoảng cách giữa hai điểm  $i$  và  $j$  trên trục toạ độ  $Oxy$  là  $|x_i - x_j| + |y_i - y_j|$  tương

ứng với khoảng cách trên hệ trục tọa độ  $Ox^*y^*$  là  $\max(|x_i^* - x_j^*|, |y_i^* - y_j^*|)$ . Độ phức tạp bước này là  $O(n)$ .

2. Sử dụng thuật toán chia để trị trên hệ trục tọa độ  $Ox^*y^*$  tương tự bài toán cơ bản “Cặp điểm gần nhất” (Closest Pair) để tìm các thành phần liên thông với  $s$  cho ban đầu. Độ phức tạp bước này là  $O(n \log n)$ .
3. Thực hiện  $k$  lần tìm khoảng cách nhỏ nhất để kết nối hai thành phần liên thông vào với nhau. Để tìm được khoảng cách nhỏ nhất cần giải quyết bài toán tìm hai điểm gần nhất ở hai thành phần liên thông khác nhau bằng cách tiếp tục sử dụng thuật toán chia để trị tương tự bài toán “Cặp điểm gần nhất”. Độ phức tạp bước này là  $O(kn \log n)$ .

Độ phức tạp thuật toán là:  $O(kn \log n)$ .

----- HẾT -----

**ĐỀ THI CHÍNH THỨC**Môn: **TIN HỌC**Thời gian: **180 phút** (*không kể thời gian giao đề*)Ngày thi: **05/3/2022**

Đề thi gồm 04 trang, 03 bài

**TỔNG QUAN ĐỀ THI**

|       | Tên bài          | File chương trình | File dữ liệu vào | File kết quả |
|-------|------------------|-------------------|------------------|--------------|
| Bài 4 | Đoạn số          | SSEQ.*            | SSEQ.INP         | SSEQ.OUT     |
| Bài 5 | Phân mềm vẽ      | PAINT.*           | PAINT.INP        | PAINT.OUT    |
| Bài 6 | Xây dựng ma trận | MATRIX.*          | MATRIX.INP       | MATRIX.OUT   |

Dấu \* được thay thế bởi PAS hoặc CPP của ngôn ngữ lập trình sử dụng tương ứng là Pascal hoặc C++.

*Hãy lập trình giải các bài toán sau:*

**Bài 4. Đoạn số (7 điểm)**

Là sinh viên ngành Công nghệ thông tin, Nam thường xuyên rèn luyện tư duy và kỹ năng lập trình bằng các bài toán lập trình thi đấu. Một bài toán thú vị mà Nam đang suy nghĩ để giải như sau:

Cho  $n$  đoạn số nguyên trên trục số, đoạn thứ  $k$  ( $1 \leq k \leq n$ ) có đầu mút bên trái là  $L_k$ , đầu mút bên phải là  $R_k$  và có trọng số là  $w_k$ . Với  $a, b$  là hai số nguyên, trọng số của cặp số  $(a, b)$  được tính bằng tổng trọng số của tất cả các đoạn  $t$  mà  $a \leq L_t \leq R_t \leq b$  với  $1 \leq t \leq n$ . Cần tìm cặp số nguyên  $(a, b)$  có trọng số là lớn nhất.

**Yêu cầu:** Cho  $n$  đoạn số, gọi  $S$  là trọng số của cặp số nguyên  $(a, b)$  có trọng số là lớn nhất, hãy giúp Nam xác định giá trị  $S$ .

**Dữ liệu:** Vào từ file văn bản **SSEQ.INP**:

- Dòng đầu chứa số nguyên dương  $n$ ;
- Dòng thứ  $k$  ( $1 \leq k \leq n$ ) trong  $n$  dòng tiếp theo chứa ba số nguyên  $L_k, R_k, w_k$  mô tả đoạn thứ  $k$  ( $1 \leq L_k \leq R_k \leq 10^6$ ;  $|w_k| \leq 10^6$ ).

Các số trên cùng một dòng cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **SSEQ.OUT** một số nguyên  $S$  là trọng số lớn nhất xác định được.

**Ràng buộc:**

- Có 30% số test ứng với 30% số điểm thỏa mãn:  $n \leq 200$ ;
- 30% số test khác ứng với 30% số điểm thỏa mãn:  $n \leq 2000$ ;
- 20% số test khác ứng với 20% số điểm thỏa mãn:  $R_1 - L_1 = R_2 - L_2 = \dots = R_n - L_n$  và  $n \leq 10^5$ ;
- 20% số test còn lại ứng với 20% số điểm thỏa mãn:  $n \leq 10^5$ .



Ví dụ:

| SSEQ. INP                               | SSEQ. OUT | Giải thích  |
|---|-----------|---|
| 4<br>1 2 -5<br>3 5 6<br>3 4 -1<br>4 6 3 | 8         | Trọng số lớn nhất là 8 bằng cách chọn cặp số (3,6). Trọng số của cặp số (3,6) bằng tổng trọng số của ba đoạn [3,5], [3,4], [4,6]. |

### Bài 5. Phần mềm vẽ (7 điểm)

Nam đã xây dựng được một phần mềm vẽ hình với giao diện chính là một bảng vẽ kích thước  $W \times H$ . Bảng vẽ được đặt trên hệ trục chiếm một vùng là một hình chữ nhật có toạ độ trái dưới là  $(0,0)$  và toạ độ phải trên là  $(W,H)$ . Nam đã vẽ  $n$  đa giác lồi trên bảng vẽ để thử nghiệm phần mềm, đa giác nào cũng có đúng  $m$  đỉnh.

Đa giác thứ  $i$  ( $1 \leq i \leq n$ ) được mô tả bằng dãy toạ độ của  $m$  đỉnh liệt kê theo chiều kim đồng hồ tương ứng là  $(x_{i,1}, y_{i,1}), (x_{i,2}, y_{i,2}), \dots, (x_{i,m}, y_{i,m})$ . Hai cạnh bất kì của hai đa giác không có điểm chung và cũng không chạm vào biên của bảng vẽ. Có  $Q$  phương án thử nghiệm, phương án thứ  $t$  ( $1 \leq t \leq Q$ ) sẽ thực hiện bấm tô màu vào điểm  $(x_t, y_t)$  không nằm trên bất kì cạnh nào của các đa giác cũng như biên của bảng vẽ, khi đó toàn bộ vùng chứa điểm  $(x_t, y_t)$  được giới hạn bởi các cạnh của các đa giác và biên của bảng vẽ sẽ được tô, phần mềm sẽ trả về diện tích vùng được tô.

**Yêu cầu:** Cho  $n$  đa giác lồi cùng có  $m$  đỉnh và  $Q$  phương án tô màu, với mỗi phương án hãy xác định diện tích vùng được tô để kiểm tra phần mềm.

**Dữ liệu:** Vào từ file văn bản **PAINT.INP**:

- Dòng đầu chứa năm số nguyên  $W, H, n, m, Q$  ( $W, H \leq 10^8; 3 \leq m \leq 5$ );
- Dòng thứ  $i$  ( $1 \leq i \leq n$ ) trong  $n$  dòng tiếp theo chứa  $2m$  số nguyên dương  $x_{i,1}, y_{i,1}, x_{i,2}, y_{i,2}, \dots, x_{i,m}, y_{i,m}$  lần lượt là toạ độ  $m$  đỉnh được liệt kê theo chiều kim đồng hồ của đa giác thứ  $i$  ( $0 < x_{i,1}, x_{i,2}, \dots, x_{i,m} < W; 0 < y_{i,1}, y_{i,2}, \dots, y_{i,m} < H$ );
- Dòng thứ  $t$  ( $1 \leq t \leq Q$ ) trong  $Q$  dòng tiếp theo chứa hai số nguyên dương  $x_t, y_t$  mô tả phương án thứ  $t$  ( $0 < x_t < W; 0 < y_t < H$ ).

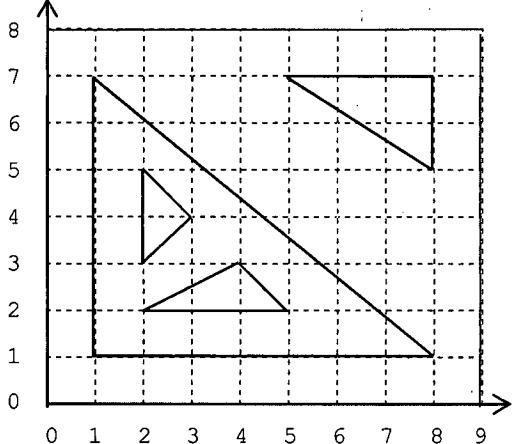
Các số trên cùng một dòng cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **PAINT.OUT** gồm  $Q$  dòng, mỗi dòng chứa một số thực (lấy đúng một chữ số sau dấu chấm thập phân) là diện tích vùng được tô tương ứng phương án trong dữ liệu vào.

**Ràng buộc:**

- Có 30% số test ứng với 30% số điểm của bài thỏa mãn:  $n, Q \leq 20$  và các đa giác là các hình chữ nhật có cạnh song song với một trong hai trục toạ độ;
- 35% số test khác ứng với 35% số điểm của bài thỏa mãn:  $n, Q \leq 2000$ ;
- 25% số test khác ứng với 25% số điểm của bài thỏa mãn:  $n, Q \leq 10^5$  và các đa giác là các hình chữ nhật có cạnh song song với một trong hai trục toạ độ;
- 10% số test còn lại ứng với 10% số điểm của bài thỏa mãn:  $n, Q \leq 10^5$ .

Ví dụ:

| PAINT . INP  | PAINT . OUT         | Minh họa   |
|--|---------------------|--|
| 9 8 4 3 3<br>8 1 1 1 1 7<br>5 7 8 7 8 5<br>2 3 2 5 3 4<br>2 2 4 3 5 2<br>7 6<br>6 2<br>8 3 | 3.0<br>18.5<br>48.0 |  |

### Bài 6. Xây dựng ma trận (6 điểm)

Để xây dựng thêm chức năng biến đổi ảnh cho phần mềm vẽ, Nam cần xây dựng ma trận  $A$  có tính chất sau:

- Ma trận có  $m$  hàng và  $n$  cột, các hàng được đánh số từ 1 đến  $m$  từ trên xuống dưới, các cột được đánh số từ 1 đến  $n$  từ trái sang phải. Phần tử nằm ở hàng  $i$  ( $1 \leq i \leq m$ ) và cột  $j$  ( $1 \leq j \leq n$ ) kí hiệu là  $A_{i,j}$ , giá trị mỗi phần tử của ma trận đều là các số nguyên dương;
- Hai phần tử cùng hàng nhưng có chỉ số cột là nguyên tố cùng nhau thì giá trị phải khác nhau. Cụ thể: với  $1 \leq j_1, j_2 \leq n$  và ước số chung lớn nhất của  $(j_1, j_2)$  bằng 1 thì  $A_{i,j_1}$  khác  $A_{i,j_2}$  với mọi  $i$  ( $1 \leq i \leq m$ );
- Hai phần tử cùng cột nhưng có chỉ số hàng là nguyên tố cùng nhau thì giá trị phải khác nhau. Cụ thể: với  $1 \leq i_1, i_2 \leq m$  và ước số chung lớn nhất của  $(i_1, i_2)$  bằng 1 thì  $A_{i_1,j}$  khác  $A_{i_2,j}$  với mọi  $j$  ( $1 \leq j \leq n$ );
- Thứ tự từ điển của ma trận là nhỏ nhất có thể.

Ma trận  $A$  được gọi là có thứ tự từ điển nhỏ hơn ma trận  $B$  (hai ma trận cùng có  $m$  hàng và  $n$  cột) nếu lần lượt so sánh từng phần tử theo từng hàng từ trên xuống dưới, trên mỗi hàng theo thứ tự từ trái sang phải để tìm phần tử khác nhau đầu tiên thì tại vị trí đó giá trị phần tử của ma trận  $A$  nhỏ hơn giá trị phần tử của ma trận  $B$ .

**Yêu cầu:** Cho  $m$  và  $n$ , gọi  $S$  là tổng các phần tử của ma trận, hãy tính  $S \% (10^9 + 7)$ , trong đó  $\%$  là phép toán chia lấy dư.

**Dữ liệu:** Vào từ file văn bản **MATRIX . INP** gồm một dòng duy nhất chứa hai số nguyên dương  $m, n$  cách nhau bởi dấu cách.

**Kết quả:** Ghi ra file văn bản **MATRIX . OUT** một số nguyên duy nhất là  $S \% (10^9 + 7)$ .

**Ràng buộc:**

- Có 20% số test ứng với 20% số điểm của bài thỏa mãn:  $m \times n \leq 10^4$ ;

- 30% số test khác ứng với 30% số điểm của bài thỏa mãn:  $m = 1; n \leq 10^9$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $m, n \leq 10^6$ ;
- 30% số test còn lại ứng với 30% số điểm của bài thỏa mãn:  $m, n \leq 10^9$ .

**Ví dụ:**

| MATRIX.INP | MATRIX.OUT | Giải thích  |
|------------|------------|---|
| 1 3        | 6          | Ma trận thỏa mãn cần tìm là:<br>1 2 3                   |
| 3 3        | 21         | Ma trận thỏa mãn cần tìm là:<br>1 2 3<br>2 1 4<br>3 4 1 |

----- HẾT -----

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

**HƯỚNG DẪN CHẤM THI**Môn: **TIN HỌC****Đề thi chính thức**Thời gian: **180 phút** (*không kể thời gian giao đề*)Ngày thi: **05/3/2022**

Hướng dẫn chấm gồm 03 trang

**I. Hướng dẫn chung**

1. Bài thi của thí sinh được chấm trên máy tính bằng phần mềm chấm thi Themis (*bản quyền của Cục Quản lý chất lượng*) sử dụng bộ test do Hội đồng chấm thi phê duyệt, đúng với đáp án, biểu điểm của Bộ Giáo dục và Đào tạo.
2. Tổ trưởng Tổ chấm thi phân công giám khảo thực hiện quy trình chấm thi trên máy tính, bảo đảm các yêu cầu của Quy chế thi.
3. Điểm bài thi được xuất từ phần mềm chấm thi; không quy tròn điểm thành phần của từng câu, điểm của bài thi.

**II. Đáp án và biểu điểm****Bài 4. Đoạn số (7 điểm)****Phân bổ điểm**

- Có 30% số test ứng với 30% số điểm thỏa mãn:  $n \leq 200$ ;
- 30% số test khác ứng với 30% số điểm thỏa mãn:  $n \leq 2000$ ;
- 20% số test khác ứng với 20% số điểm thỏa mãn:  $R_1 - L_1 = R_2 - L_2 = \dots = R_n - L_n$  và  $n \leq 10^5$ ;
- 20% số test còn lại ứng với 20% số điểm thỏa mãn:  $n \leq 10^5$ .

**Thuật toán**

*Nhận xét:* Nếu  $w_k \leq 0$  với mọi  $1 \leq k \leq n$  thì kết quả là 0. Ngược lại thì  $a$  là đầu mút bên trái của một đoạn nào đó và  $b$  là đầu mút bên phải của một đoạn nào đó.

Duyệt toàn bộ các cặp  $a, b$  thỏa mãn tính chất trên. Dùng kĩ thuật đường quét (sweep line):

- Xét  $b$  tăng dần, duy trì  $S$  (bằng cấu trúc dữ liệu cây phân đoạn - Segment Tree) là tập các đoạn có đầu mút bên phải không quá  $b$ ;
- Khi  $b$  tăng lên, tập  $S$  được bổ sung thêm các đoạn mới;
- Với mỗi  $b$ , tìm  $a$  sao cho tổng trọng số các đoạn trong  $S$  mà có đầu mút bên trái lớn hơn hoặc bằng  $a$  là lớn nhất.

Độ phức tạp:  $O(n \log n)$

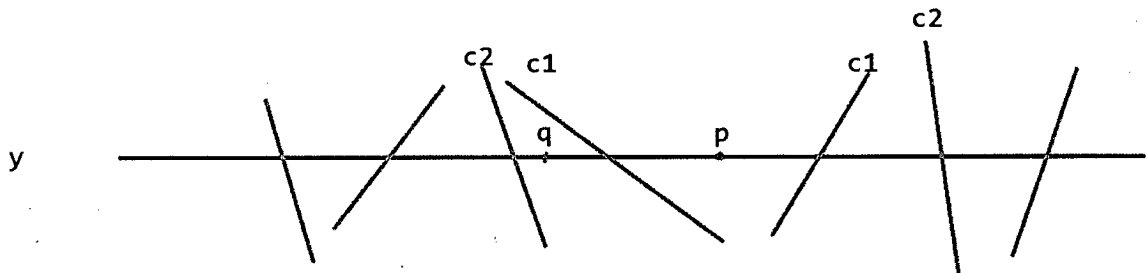
## Bài 5. Phần mềm vẽ (7 điểm)

### Phân bổ điểm

- Có 30% số test ứng với 30% số điểm của bài thỏa mãn:  $n, Q \leq 20$  và các đa giác là các hình chữ nhật có cạnh song song với một trong hai trục tọa độ;
- 35% số test khác ứng với 35% số điểm của bài thỏa mãn:  $n, Q \leq 2000$ ;
- 25% số test khác ứng với 25% số điểm của bài thỏa mãn:  $n, Q \leq 10^5$  và các đa giác là các hình chữ nhật có cạnh song song với một trong hai trục tọa độ;
- 10% số test còn lại ứng với 10% số điểm của bài thỏa mãn:  $n, Q \leq 10^5$ .

### Thuật toán

*Nhận xét:* Quan sát hình bên dưới thấy rằng, khi bấm vào các điểm  $p, q$  và các cạnh cắt đường thẳng  $y$ , với điểm  $p$  nằm trong hai cạnh của đa giác  $c_1$  thì có thể khẳng định điểm  $p$  nằm trong đa giác  $c_1$ , còn điểm  $q$  nằm giữa hai cạnh của đa giác  $c_1$  và  $c_2$  vì  $c_1$  nằm trong  $c_2$  nên  $q$  sẽ nằm trong đa giác  $c_2$ . Việc kiểm tra một hình nằm trong đa giác vào cũng chính là bài toán kiểm tra một điểm nằm trong đa giác nào, do các đa giác không cắt nhau nên chỉ cần xét một điểm của một đa giác  $c$  và tìm đa giác  $c'$  chứa điểm đó để biết đa giác  $c'$  chứa đa giác  $c$ . Như vậy, có thể sử dụng kỹ thuật đường quét (sweep line) để tìm đa giác chứa một điểm. Với mỗi độ cao  $y$  mà đường quét quét qua duy trì một tập  $S$  là các cạnh cắt qua độ cao  $y$ . Lưu các cạnh trong  $S$  tương ứng là hai điểm của cạnh đó, với hai cạnh sẽ xác định được cạnh nào nằm bên trái cạnh nào nằm bên phải bằng tích có hướng, do đó các cạnh trong  $S$  sẽ được lưu theo thứ tự từ trái sang phải. Sử dụng kỹ thuật tìm kiếm nhị phân để xác định điểm  $p$  nằm giữa hai cạnh nào.



Từ nhận xét trên ta có thuật toán sau:

- Với mỗi hình  $i$ , cần xác định hình  $j$  có diện tích nhỏ nhất chứa  $i$ , khi đó, diện tích sẽ tổng của hình  $j$  sẽ giảm đi một lượng bằng diện tích của hình  $i$ .
- Với mỗi điểm  $p$  được bấm, xác định hình nhỏ nhất chứa điểm  $p$  để trả lời.

Độ phức tạp:  $O((n + Q) \log n)$ .

## Bài 6. Xây dựng ma trận (6 điểm)

### Phân bổ điểm

- Có 20% số test ứng với 20% số điểm của bài thỏa mãn:  $m \times n \leq 10^4$ ;
- 30% số test khác ứng với 30% số điểm của bài thỏa mãn:  $m = 1; n \leq 10^9$ ;
- 20% số test khác ứng với 20% số điểm của bài thỏa mãn:  $m, n \leq 10^6$ ;
- 30% số test còn lại ứng với 30% số điểm của bài thỏa mãn:  $m, n \leq 10^9$ .

## Thuật toán

Gọi  $p_i$  là số nguyên tố lớn thứ  $i$ . Gọi  $G(x)$  là số nguyên dương nhỏ nhất sao cho  $x$  chia hết cho  $p_{G(x)}$ . Quy ước  $p_0 = 1$  và  $G(1) = 0$ .

Nhận xét:  $A_{i,j} - 1 = (A_{i,1} - 1) \text{ XOR } (A_{1,j} - 1) = G(i) \text{ XOR } G(j)$  với mọi  $i, j \geq 1$ .

Đếm  $c_i$  là số lượng số  $x \leq n$  mà  $G(x) = i$  như sau:

- Với  $p_i \leq \sqrt{n}$ :

Gọi  $w(x, i)$  là số lượng số nhỏ hơn hoặc bằng  $x$  mà tất cả các ước nguyên tố của nó đều lớn hơn  $p_i$ . Khi đó:

$$w(x, i) = w(x, i-1) - w\left(\left\lfloor \frac{x}{p_i} \right\rfloor, i-1\right) \text{ và } c_i = w\left(\left\lfloor \frac{n}{p_i} \right\rfloor, i-1\right).$$

Việc tính toán chỉ cần thực hiện với  $x \in \{\left\lfloor \frac{n}{i} \right\rfloor, 1 \leq i \leq n\}$  và  $p_i \leq \sqrt{n}$ , có độ phức tạp cỡ  $O\left(\frac{n}{\log(n)}\right)$  và bộ nhớ  $O(\sqrt{n})$ .

- Với  $p_i > \sqrt{n}$ :

Sau khi loại bỏ hết các số  $x$  mà  $G(x) \leq \sqrt{n}$  thì các số còn lại đều là số nguyên tố. Do đó,  $c_i = 1$  với mọi  $\sqrt{n} < p_i \leq n$ . Số lượng các số nguyên tố này là  $n - \sum_{i=1}^{p_i \leq \sqrt{n}} c_i$ .

Tương tự, tính được  $d_i$  là số lượng các số  $x \leq m$  mà  $G(x) = i$ .

Bài toán đưa về tính tổng tất cả các cặp  $c_i \text{ XOR } d_j$  và có thể tính toán trên từng bit.

Độ phức tạp:  $O\left(\frac{n}{\log(n)}\right)$ .

----- HẾT -----