

ĐỀ THI CHÍNH THỨC

Môn thi: Tin học

Thời gian: 300 phút

(Không kể thời gian phát đề)

Ngày thi thứ nhất: 09/4/2011

## TỔNG QUAN ĐỀ THI NGÀY THỨ NHẤT

*Đề thi gồm có 4 trang*

TT	Tên bài	Tên file chương trình	Tên file dữ liệu	Tên file kết quả
1	Đua ô tô	RACE.*	RACE.INP	RACE.OUT
2	Phân loại văn bản cổ	PART.*	PART.INP	PART.OUT
3	Bảng thông tin điện tử	TABLE.*	TABLE.INP	TABLE.OUT
4	Phòng chống virus	RECOVER.*	RECOVER.INP	RECOVER.OUT

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

**Hãy lập trình giải các bài toán sau đây:**

**Bài 1. Đua ô tô**

Cuộc thi đua ô tô được tổ chức ở sa mạc Sama. Sa mạc được chia thành lưới ô vuông gồm  $m$  hàng và  $n$  cột. Các hàng của lưới được đánh số từ 1 đến  $m$ , từ trên xuống dưới. Các cột của lưới được đánh số từ 1 đến  $n$ , từ trái sang phải. Ô nằm trên giao của hàng  $i$  và cột  $j$  có tọa độ  $(i, j)$ . Mỗi ô vuông của lưới hoặc chứa vật cản không thể đi qua, hoặc là ô trống có thể đi qua. Riêng ô  $(1, 1)$  và ô  $(m, n)$  luôn là ô trống và luôn có đường đi qua các ô trống chung cạnh từ ô  $(1, 1)$  đến ô  $(m, n)$ .

Xuất phát từ ô  $(1, 1)$  tay đua phải lái ô tô di chuyển đến đích ở ô  $(m, n)$  tuân theo các qui tắc sau:

- Xuất phát từ ô  $(1, 1)$ , tay đua có thể lái ô tô di chuyển sang ô  $(1, 2)$  hoặc ô  $(2, 1)$ , tăng tốc ô tô từ 0 lên  $d$ , và sẽ mất  $1/d$  thời gian để di chuyển qua ô  $(1, 1)$  đến ô tiếp theo.
- Từ một ô, tay đua có thể giữ nguyên hướng để đi thẳng đến ô trước mặt, có thể rẽ trái hoặc rẽ phải sang ô nằm ở bên trái hoặc ô nằm ở bên phải ô hiện tại, nhưng không được di chuyển ra ngoài lưới và việc thay đổi hướng di chuyển (rẽ trái hoặc rẽ phải) chỉ được thực hiện khi tốc độ của ô tô tại ô thực hiện đổi hướng không vượt quá giới hạn cho phép  $v_{limit}$ .
- Khi bắt đầu di chuyển vào một ô mới, tay đua có thể tăng tốc độ của ô tô thêm  $d$ , giảm tốc độ của ô tô đi  $d$ , hoặc giữ nguyên tốc độ để đi qua ô này. Tốc độ tối đa mà ô tô có thể đạt được là  $v_{max}$ . Nếu tốc độ của ô tô đang là  $v$  thì tay đua chỉ được tăng tốc nếu  $v + d \leq v_{max}$  và chỉ được giảm tốc nếu  $v - d > 0$ . Việc di chuyển qua một ô với tốc độ  $v$  đòi hỏi thời gian  $1/v$ .

**Yêu cầu:** Hãy xác định thời gian ngắn nhất để tay đua về đích.

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

- Dòng đầu tiên ghi các số nguyên dương  $m, n, d, v_{limit}, v_{max}$ .  
( $m, n \leq 300$ ;  $d, v_{max} \leq 10^6$ ;  $d \leq v_{limit} \leq 3*d$ .)
- Dòng thứ  $i$  trong số  $m$  dòng tiếp theo chứa một xâu bit độ dài  $n$ , trong đó nếu bit thứ  $j$  của xâu là 0 thì ô  $(i, j)$  là ô trống, còn nếu bit này là 1 thì ô  $(i, j)$  là ô có vật cản.

**Kết quả:** Ghi ra file văn bản RACE.OUT một số thực với 6 chữ số sau dấu chấm là thời gian ngắn nhất để tay đua về đích.

Ví dụ:

RACE . INP	RACE . OUT
3 4 1 2 7	2.916666
0000	
0100	
0100	

**Giải thích:** Bắt đầu từ ô (1, 1) ô tô tăng tốc lên 1 đi qua ô (1, 1) mất thời gian 1. Tiếp tục di chuyển vào ô (1, 2) tăng tốc lên 2, di chuyển qua ô (1, 2) mất thời gian 1/2. Tiếp tục di chuyển vào ô (1, 3) tăng tốc lên 3, di chuyển qua ô (1, 3) mất thời gian 1/3. Tiếp tục di chuyển vào ô (1, 4), giảm tốc còn 2, di chuyển qua ô (1, 4) mất thời gian 1/2. Rồi phải, di chuyển vào ô (2, 4), tăng tốc lên 3, di chuyển qua ô (2, 4) mất thời gian 1/3. Tiếp tục di chuyển vào ô (3, 4), tăng tốc lên 4, di chuyển qua ô (3, 4) để cán đích mất thời gian 1/4. Như vậy tổng thời gian để về đích theo cách thực hiện trên là:  $1 + 1/2 + 1/3 + 1/2 + 1/3 + 1/4 = 2.916666$ .

## Bài 2. Phân loại văn bản cổ

Trong một lần khai quật một khu di tích được xác định là một thư viện cổ cách chúng ta hàng nghìn năm, các nhà khảo cổ phát hiện được  $n$  văn bản cổ. Ký hiệu  $V$  là tập gồm  $n$  văn bản cổ này. Theo những câu truyện truyền thuyết được dân địa phương kể lại, các nhà khảo cổ biết rằng các văn bản cổ này thuộc  $k$  loại ngôn ngữ. Vấn đề đặt ra cho các nhà khảo cổ là phải xác định xem mỗi văn bản thuộc loại ngôn ngữ nào.

Cụ thể, các nhà khảo cổ cần giải bài toán sau đây: Mỗi một văn bản cổ là một xâu gồm không quá 100 ký hiệu. Cho xâu  $x$ , ta gọi xâu con của  $x$  là xâu thu được từ nó bởi việc xóa đi một số ký hiệu và giữ nguyên thứ tự của các ký hiệu còn lại. Một xâu  $z$  được gọi là xâu con chung của hai xâu  $x$  và  $y$  nếu như nó vừa là xâu con của  $x$  vừa là xâu con của  $y$ . Số lượng ký hiệu trong một xâu được gọi là độ dài của nó. Ta gọi độ phân biệt của hai xâu  $x$  và  $y$ , ký hiệu  $\rho(x, y)$ , là hiệu giữa tổng độ dài của hai xâu  $x, y$  và hai lần độ dài của xâu con chung dài nhất của chúng.

Ví dụ: Nếu  $x = \text{'aaa'}$  còn  $y = \text{'aca'}$  thì do độ dài của xâu con chung dài nhất của hai xâu  $x$  và  $y$  là 2, nên ta có  $\rho(x, y) = 3 + 3 - 2 \cdot 2 = 2$ .

Giả sử  $P$  và  $Q$  là hai tập con không giao nhau của tập  $V$  gồm  $n$  văn bản cổ đã cho, ta gọi độ tương đồng của  $P$  và  $Q$  là số

$$d(P, Q) = \min \{ \rho(x, y) : x \in P, y \in Q \}.$$

Rõ ràng có một tương ứng 1-1 giữa một cách phân loại các văn bản đã cho theo  $k$  ngôn ngữ với một cách phân hoạch tập các văn bản cổ đã cho  $V$  ra thành  $k$  tập con:

$$V = V_1 \cup V_2 \cup \dots \cup V_k;$$

$$V_i \cap V_j = \emptyset, i \neq j.$$

Ta gọi độ chính xác của cách phân hoạch này là số

$$\min \{ d(V_i, V_j) : 1 \leq i < j \leq k \}.$$

Bài toán phân loại văn bản cổ đòi hỏi tìm cách phân hoạch tập các văn bản cổ đã cho  $V$  ra thành  $k$  tập con  $V_1, V_2, \dots, V_k$  với độ chính xác lớn nhất.

**Yêu cầu:** Cho biết tập  $n$  xâu và số  $k$ , hãy tính độ chính xác lớn nhất trong số các độ chính xác của các cách phân hoạch tập  $n$  xâu đã cho ra thành  $k$  tập con.

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

- Dòng đầu tiên chứa hai số nguyên dương  $n$  và  $k$  được ghi cách nhau bởi ít nhất một dấu cách ( $1 < k \leq n \leq 500$ );

- Mỗi dòng trong số  $n$  dòng tiếp theo chứa một xâu gồm không quá 100 ký hiệu, mỗi ký hiệu là chữ cái la tinh in thường từ 'a' đến 'z'.

**Kết quả:** Ghi ra file văn bản PART.OUT một số nguyên duy nhất là độ chính xác tìm được.

**Ví dụ:**

PART . INP	PART . OUT
3 2 aaa bba aca	4

### Bài 3. Bảng thông tin điện tử

Bảng thông tin điện tử được lắp trên các đường phố để cung cấp ngắn gọn các thông tin quan trọng, các sự kiện, khẩu hiệu ... Công ty điện tử Sáng Sao vừa cho xuất xưởng một bảng thông tin điện tử có dạng một hàng gồm  $n$  vị trí, mỗi vị trí hiển thị một ký tự. Các vị trí được đánh số từ 1 đến  $n$  từ trái qua phải. Các ký tự chạy từ phải qua trái. Cứ mỗi giây ký tự ở vị trí  $i$  chuyển sang vị trí  $i-1$  ( $i = 2, 3, \dots, n$ ) và ký tự mới từ xâu dữ liệu vào được lên bảng ở vị trí  $n$ . Ban đầu, tất cả các vị trí đều chứa dấu cách.

Trong thời gian thử nghiệm, để kiểm tra chất lượng bảng Công ty Sáng Sao cho phát lên bảng xâu  $S$  được tạo thành từ cách viết liên tiếp các số tự nhiên  $1, 2, 3, 4, \dots, 10^9$ , nhưng khi gặp 2 hay nhiều chữ số liên tiếp giống nhau người ta chỉ giữ lại một chữ số. Như vậy, phần đầu của xâu, khi viết đến số 14 sẽ là

12345678910121314

Nếu  $n = 5$  thì ở giây thứ 17 kể từ lúc bắt đầu phát thử nghiệm trên bảng thông tin sẽ có nội dung

2	1	3	1	4
---	---	---	---	---

**Yêu cầu:** Cho xâu  $T$  độ dài  $n$ , chỉ chứa các ký tự số trong phạm vi từ 0 đến 9 và không có hai ký tự liên tiếp nào giống nhau. Hãy xác định thời điểm lần đầu tiên xuất hiện xâu  $T$ , giả thiết là thời điểm bắt đầu phát thử nghiệm là 0.

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

- Dòng đầu tiên chứa số nguyên  $n$  ( $1 \leq n \leq 100$ );
- Dòng thứ 2 chứa xâu  $T$  độ dài  $n$ . Dữ liệu đảm bảo xâu  $T$  xuất hiện ít nhất 1 lần.

**Kết quả:** Đưa ra file văn bản TABLE.OUT một số nguyên  $k$  là thời điểm lần đầu tiên xuất hiện xâu  $T$ .

**Ví dụ:**

TABLE . INP	TABLE . OUT
5 21314	17

### Bài 4. Phòng chống virus

Tình trạng lây lan virus trên mạng nhiều khi làm tê liệt hoạt động của hệ thống máy tính. Mạng máy tính của công ty gồm  $n$  máy tính được kết nối với nhau bởi  $n-1$  cáp nối giữa các cặp máy sao cho luôn có đường truyền tin từ một máy đến bất cứ máy nào còn lại hoặc theo cáp nối trực tiếp giữa chúng, hoặc qua dây các cáp nối giữa các máy trong mạng. Các máy tính được đánh số từ 1 đến  $n$ , các cáp nối được đánh số từ 1 đến  $n-1$ . Tuấn là quản trị mạng của công ty Sao Sáng được

Giám đốc giao cho nhiệm vụ tìm cách xử lý tối ưu việc ngăn ngừa lây lan của virus theo một kịch bản được giả định sẵn. Kịch bản khắc phục sự cố diễn ra trong  $m$  ngày như sau:

- Tại ngày thứ  $i$  ( $i = 1, 2, \dots, m$ ) máy  $p_i$  bị nhiễm virus, và virus này có thể lây nhiễm sang  $k(i)$  máy  $p_{i_1}, p_{i_2}, \dots, p_{i_{k(i)}}$ .
- Tuấn phải lựa chọn và ngắt hoạt động của một số cáp nối trong mạng sao cho không còn đường truyền tin từ máy  $p_i$  đến các máy  $p_{i_1}, p_{i_2}, \dots, p_{i_{k(i)}}$ .
- Đến cuối ngày  $i$  Tuấn phải khôi phục lại hoạt động của các cáp nối đã bị ngắt.

Biết chi phí để ngắt và khôi phục hoạt động của các đoạn cáp nối, Tuấn cần tìm cách thực hiện nhiệm vụ trong kịch bản với tổng chi phí ngắt và khôi phục hoạt động của các cáp nối trong mỗi ngày là nhỏ nhất.

**Yêu cầu:** Cho trước sơ đồ kết nối của mạng và kịch bản khắc phục sự cố hãy giúp Tuấn xác định chi phí nhỏ nhất để hoàn thành nhiệm vụ đề ra của mỗi ngày trong kịch bản.

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

- Dòng đầu tiên chứa số nguyên dương  $n$  là số lượng máy tính trong mạng;
- Dòng thứ  $j$  trong số trong số  $n-1$  dòng tiếp theo chứa ba số nguyên  $u_j, v_j, s_j$  cho biết cáp nối thứ  $j$  nối máy  $u_j$  với máy  $v_j$  và chi phí để ngắt và khôi phục hoạt động là  $s_j$ ;
- Dòng tiếp theo chứa số nguyên dương  $m$  là số ngày trong kịch bản;
- Dòng thứ  $i$  trong số  $m$  dòng cuối cùng chứa thông tin về nhiệm vụ của ngày  $i$  trong kịch bản gồm các số  $p_i, k(i), p_{i_1}, p_{i_2}, \dots, p_{i_{k(i)}}$ . ( $p_{i_u} \neq p_{i_v}$  nếu  $u \neq v$  và  $p_i \neq p_{i_v}$ , với  $1 \leq v \leq k(i)$ .)

Các số trên cùng một dòng được ghi cách nhau ít nhất một dấu cách.

**Hạn chế:**

- $2 \leq n \leq 250000, m \geq 1, k(1) + k(2) + \dots + k(m) \leq 500000, 1 \leq k(i) \leq n-1, 1 \leq s_i \leq 100000$ ;
- Có ít nhất 50% tests thỏa mãn điều kiện: “Máy bị nhiễm virus  $p_i$  là không thay đổi trong tất cả các ngày diễn ra kịch bản”.

**Kết quả:** Ghi ra file văn bản RECOVER.OUT gồm  $m$  dòng, dòng thứ  $i$  chứa tổng chi phí thực hiện trong ngày  $i$  ( $i = 1, 2, \dots, m$ ).

**Ví dụ:**

RECOVER.INP	RECOVER.OUT	Hình vẽ minh họa
7 1 3 9 2 3 4 3 4 8 3 6 12 2 5 17 5 7 6 2 2 3 1 6 7 2 3 4 5 7	10 21	

----- HẾT -----

- Thí sinh không được sử dụng tài liệu dưới mọi hình thức.
- Giám thị không giải thích gì thêm.