

## I. BÀI TẬP CƠ BẢN

**Bài 1:** Cho 1 xâu, đếm số từ trong xâu

**Bài 2:** Cho 1 xâu có chứa kí tự số. Tách các cụm kí tự số để được các số. Sắp xếp các số tách được theo thứ tự tăng dần

**Bài 3:** Chuẩn hóa xâu

Cho xâu gồm chữ cái, chữ số, dấu cách

- Xóa dấu cách đầu, cuối xâu.
- Giữa các từ cách nhau đúng 1 dấu cách
- Đầu từ viết hoa, phần trong từ viết thường

## II. BÀI TẬP NÂNG CAO

### 1. DANH SÁCH TRƯỜNG HỌC

Ở vòng thi loại Olympic Tin học mọi học sinh đều có quyền đăng ký tham gia. Mục tên trường là một trong những thông tin bắt buộc phải điền trong phiếu đăng ký. Tên trường bao gồm tên thường gọi và số của trường. Số của trường là số nguyên không bắt đầu bằng 0. Không có hai trường nào có số giống nhau. Tên thường gọi của một trường có thể khác nhau tùy cách gọi của người dự thi, nhưng không chứa số. Người đăng ký có thể nêu tên trường theo cách tùy ý, nhưng bắt buộc phải ghi số của trường, ví dụ “*Liceum 9*”, “*School N 9*” hay “*9ya Shkola*”.

Có  $n$  thí sinh đăng ký dự thi. Hãy xác định số lượng trường có số người đăng ký không quá 5 và chỉ ra số của các trường đó.

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

- Dòng đầu tiên chứa số nguyên  $n$  ( $1 \leq n \leq 1000$ ),
- Mỗi dòng trong  $n$  dòng sau chứa xâu độ dài không quá 100 xác định tên trường, xâu chỉ chứa các ký tự la tinh (thường và hoa), ký tự số và dấu cách.

**Kết quả:** Đưa ra file văn bản SCHOOLS.OUT:

- Dòng đầu tiên chứa số nguyên  $m$  là số lượng trường có số người đăng ký không quá 5,
- Mỗi dòng trong  $m$  dòng sau chứa một số nguyên là số của các trường có số người đăng ký không quá 5.

**Ví dụ:**

SCHOOLS.INP	SCHOOLS.OUT
9	2
Physics and Mathematics School 18	18
9ya shkola imeni Pushkina	42
Lyceum 9	
PaMS 18	
Gymnasium 42	
School 9	
Shkola nomer 9	
High school 9	
School N 9	

### 2. QUAN HỆ HỌ HÀNG

Để theo dõi thực trạng đời sống của động vật hoang dã phục vụ cho công tác bảo tồn thiên nhiên người ta gán cho mỗi động vật bắt được một con chip, trong đó có chứa mã phân loại. Theo dõi đường di chuyển và hoạt động hàng ngày của chúng người ta biết được rất nhiều thông tin, trong đó có việc ra đời của các động vật thế hệ F2, F3, ... Mã phân loại là một số thập phân nguyên dương không vượt quá  $10^9$ . Các cấp mã phân loại đảm bảo nếu hai cá thể có quan hệ họ hàng với nhau thì hai mã phải có ít nhất một chữ số giống nhau. Ví dụ 2 con vật có các mã phân loại tương ứng là 47 và 107 có quan hệ họ hàng với nhau, còn hai con vật với các mã 47 và 931 không có quan hệ.

Cho  $n$  mã phân loại khác nhau từng đôi một. Hãy xác định số cặp động vật có quan hệ họ hàng.

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

- Dòng đầu tiên chứa số nguyên  $n$  ( $2 \leq n \leq 5 \times 10^5$ ),
- Dòng thứ 2 chứa  $n$  mã phân loại.

**Kết quả:** Đưa ra file văn bản RELATIVE.OUT số cặp động vật xác định được.

**Ví dụ:**

RELATIVES.INP	RELATIVES.OUT
5 10 74 47 77 301	4