

WINTER CONTEST 2021

Lưu ý: Các thí sinh có thể nộp bằng ngôn ngữ Pascal, C++ hoặc Java tùy theo ngôn ngữ nào thí sinh thấy thuận tiện. Tên các bài là Tenbai.*, trong đó * tương ứng với pas, cpp hay java. Tenbai là tên của bài sẽ được quy định trong cụ thể từng bài.

Bài 4 (WC4.*):

Bạn được cho 2 chuỗi ký tự s1 và s2 có cùng số lượng ký tự là N. Bạn được cho thêm một chuỗi s có độ dài là M. Các chuỗi ký tự đều được tạo nên từ các chữ cái trong bảng chữ cái Latin in thường ('a' .. 'z'). Bạn hãy đếm xem có bao nhiêu chuỗi st thỏa mãn:

- Gồm N ký tự chữ cái latin in thường
- Chuỗi st không nhỏ hơn s1 về mặt thứ tự từ điển và không lớn hơn s2 về mặt thứ tự từ điển. Một chuỗi a được gọi là có thứ tự từ điển lớn hơn chuỗi b nếu như tồn tại một vị trí i sao cho ký tự thứ i của chuỗi a lớn hơn ký tự thứ i của chuỗi b, đồng thời các ký tự từ 1 đến i - 1 của chuỗi a và chuỗi b bằng nhau. Ví dụ chuỗi a = "abcd" là chuỗi có thứ tự từ điển đứng sau chuỗi b = "abbd", vì tồn tại vị trí i = 3, có a[3] = 'c' > b[3] = 'b', đồng thời các ký tự từ vị trí 1 -> vị trí 2 của 2 chuỗi bằng nhau. Ví dụ: Chuỗi s1 = "abc", chuỗi s2 = "aca", một số chuỗi st có thứ tự từ điển lớn hơn hoặc bằng s1 và nhỏ hơn hoặc bằng s2 là: st = "abc", "abd", "abe", "abf", ..., "abz", "aca".
- Không nhận ký tự s làm một chuỗi con. Một chuỗi st được gọi là "chuỗi con" của chuỗi s nếu như st là một tập hợp gồm các ký tự LIÊN TỤC trong chuỗi s. Ví dụ: s = "wintercontest" thì một số chuỗi con là "winter", "contest", "inter", "conte", "r",

- **Dữ liệu đầu vào (WC4.INP):**

- Đọc từ file WC4.INP
- Dòng đầu tiên gồm một số nguyên dương T ($1 \leq T \leq 40$) là số lượng bộ dữ liệu của đề bài.
- T bộ dữ liệu tiếp theo, mỗi bộ dữ liệu được tổ chức gồm 3 dòng như sau:
 - Dòng đầu tiên chứa một chuỗi ký tự s1
 - Dòng thứ hai chứa một chuỗi ký tự s2
 - Dòng thứ ba chứa một chuỗi ký tự s
 - Dữ liệu input bảo đảm rằng 2 chuỗi s1 và s2 có số lượng ký tự bằng nhau, đồng thời chuỗi s có số lượng ký tự không vượt quá số lượng ký tự của 2 chuỗi s1 và s2

- **Dữ liệu đầu ra (WC4.OUT):**

- Ghi ra file WC4.OUT
- Gồm T dòng, mỗi dòng gồm một số nguyên duy nhất là số lượng chuỗi st thỏa mãn đề bài. Vì kết quả rất lớn, nên ta có thể in ra phần dư của kết quả cho số 1000000007.

- **Ví dụ:**

WC4.INP	WC4.OUT
6	24
abc	0
aca	0
r	0
aaa	390625
aab	974821414
aa	
aba	

aaa b nguyenngoctu nguyenquocan nguyen aaaa zzzz u awqntqowtiuannagnaosgmqgqlgqoiwhqiwtfagna lgneugbqeughpqnkeptjiqe qpiwppqwnnaowngawgnangladgnoagnqitqiwtqowt qnwtadgmzngdakgnoqngqgnqoi jtgwjtgwit	
---	--

- **Giải thích test ví dụ:**

- Với ví dụ 1: Các chuỗi thỏa mãn là: “abc”, “abd”, “abe”, “abf”, “abg”, “abh”, “abi”, “abj”, “abk”, “abl”, “abm”, “abn”, “abo”, “abp”, “abq”, “abs”, “abt”, “abu”, “abv”, “abw”, “abx”, “aby”, “abz”, “aca”. Chuỗi “abr” không thỏa vì có chuỗi con là chuỗi “r”
- Với ví dụ 2: Chỉ có 2 chuỗi có thứ tự từ điển lớn hơn hoặc bằng “aaa” và nhỏ hơn hoặc bằng “aab” là 2 chuỗi “aaa” và “aab”. Tuy nhiên cả 2 chuỗi đều không thỏa do có chứa chuỗi “aa” như chuỗi con.
- Với ví dụ 3: Chuỗi s1 có thứ tự từ điển lớn hơn chuỗi s2 nên không có chuỗi nào thỏa.
- Với ví dụ 4: Do 2 chuỗi s1 và s2 cùng có tiền tố 6 chữ cái đầu tiên là “nguyen”, nên các chuỗi có thứ tự từ điển lớn hơn hoặc bằng s1 và nhỏ hơn hoặc bằng s2 cũng sẽ có tiền tố “nguyen” đứng đầu. Nhưng do chuỗi s cũng là “nguyen” nên không có chuỗi nào thỏa mãn.
- Với ví dụ 5: Tất cả các chuỗi có 4 chữ cái đều thỏa. Tuy nhiên do chuỗi s = “u” nên tất cả ký tự của các chuỗi đều khác ký tự ‘u’. Do đó, với mỗi ký tự sẽ có 25 sự lựa chọn (ký tự ‘a’ -> ‘z’, ngoại trừ ký tự ‘u’). Do đó, số cách chọn là: $25 * 25 * 25 * 25 = 390625$. Kết quả là $390625 \bmod 1000000007 = 390625$.
- Với ví dụ 6: Dùng để bạn đọc kiểm tra lời giải của mình đã chính xác hay chưa.

- **Bài toán có 2 Dataset:**

- Small Dataset: Tất cả ký tự trong chuỗi s1 là ‘a’, tất cả ký tự trong chuỗi s2 là ‘z’. $|s| = 1$ với $|s|$ là độ dài của chuỗi s.
- Large Dataset: Không có ràng buộc gì thêm.
- Trong các Dataset: $1 \leq N \leq 100$, $1 \leq |s| \leq N$. Ngoại trừ Small Dataset có $|s| = 1$.

- **Giới hạn thời gian và bộ nhớ:**

- 10s / Dataset.
- 512Mb / Dataset.

Đối với học sinh lớp 10, ta cần có thêm một số công cụ hỗ trợ kiến thức như sau:

- Phép lấy phần dư trong phép chia của số a cho số b được định nghĩa là số r nếu số r thỏa mãn: Tồn tại một số nguyên d sao cho $d * b + r = a$. Ví dụ: $5 \bmod 2 = 1$, $9 \bmod 7 = 2$.
 - Trong c++, toán tử lấy phần dư được ký hiệu là %. Ví dụ:

```
int a = 9; int b = 6;
int c = a % b;
cout << c;
```

Đoạn code trên sẽ xuất số 3 ra màn hình, vì $9 \% 6 = 3$.

- Phép lấy phần dư được áp dụng để tính kết quả cho một số phép tính như sau:
 - o $(a + b) \% M = (a \% M + b \% M) \% M$. Ví dụ: tính $(15 + 29) \% 8$
 - Ta có: $15 + 29 = 44$, $44 \% 8 = 4$
 - Ngoài ra, ta có thể tính: $15 \% 8 = 7$, $29 \% 8 = 5$. $(7 + 5) \% 8 = 12 \% 8 = 4$.
 - o $(a - b) \% M = (a \% M - b \% M + M) \% M$. Ví dụ: tính $(15 - 29) \% 8$.
 - Ta có: $15 - 29 = -14$. Phép lấy phần dư trong c++ không thể sử dụng cho số âm, do đó, ta cộng -14 cho $2 * 8 = 2$. Lấy $2 \% 8 = 2$.
 - Ngoài ra, ta có thể tính: $15 \% 8 = 7$, $29 \% 8 = 5$. $(7 - 5 + 8) \% 8 = 2$.
 - o Lưu ý: Hoàn toàn không thể áp dụng phép lấy phần dư vào phép chia.

Tác dụng của việc lấy phần dư là để tránh hiện tượng tràn số. Tràn số là khi ta nhận một con số nhưng vượt ngoài giá trị cho phép của kiểu dữ liệu ta đang sử dụng.

Các em có thể xem giới hạn dữ liệu trong bảng sau:

C type	stdint.h type	Bits	Sign	Range
char	uint8_t	8	Unsigned	0 .. 255
signed char	int8_t	8	Signed	-128 .. 127
unsigned short	uint16_t	16	Unsigned	0 .. 65,535
short	int16_t	16	Signed	-32,768 .. 32,767
unsigned int	uint32_t	32	Unsigned	0 .. 4,294,967,295
int	int32_t	32	Signed	-2,147,483,648 .. 2,147,483,647
unsigned long long	uint64_t	64	Unsigned	0 .. 18,446,744,073,709,551,615
long long	int64_t	64	Signed	-9,223,372,036,854,775,808 .. 9,223,372,036,854,775,807

Bài 5 (WC5.*):

Cho một dãy số nguyên dương a gồm N số. Hãy tính tổng các giá trị $a_i * a_j$ với mọi i nhỏ hơn j.

- **Dữ liệu đầu vào (WC5.INP):**
 - o Đọc từ file WC5.INP.
 - o Dòng đầu tiên gồm một số nguyên dương T ($1 \leq T \leq 40$) là số lượng bộ dữ liệu cần xử lý
 - o T bộ dữ liệu tiếp theo, mỗi bộ dữ liệu được tổ chức như sau:
 - Dòng 1: Gồm 1 số nguyên dương là số lượng phần tử trong dãy a.
 - Dòng 2: Gồm N số nguyên dương, số nguyên dương thứ i có giá trị là a_i .
- **Dữ liệu đầu ra (WC5.OUT):**
 - o Ghi vào file WC5.OUT.
 - o Kết quả trả về cho T bộ test, mỗi bộ test gồm 1 số duy nhất là kết quả của bộ dữ liệu tương ứng.
- **Ví dụ:**

WC5.INP	WC5.OUT
4	10
5	11

1 1 1 1 1 3 1 2 3 6 4 4 4 2 2 2 10 9 5 12 58 1 85 24 90 100 99	132 96986
--	--------------

- **Giải thích test ví dụ:**
 - Ví dụ 1: Có tổng cộng 10 cặp (i, j) sao cho $i < j$ là: (1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5). Do các số trong dãy đều bằng nhau và bằng 1, nên tích của 2 số = 1. Do đó tổng là 10
 - Ví dụ 2: Có những cặp chỉ số (i, j) như sau: (1, 2), (2, 3), (1, 3). Tổng là $1 * 2 + 2 * 3 + 1 * 3 = 11$
 - Ví dụ 3 và 4: Nhằm mục đích để bạn đọc kiểm tra thuật toán đã đúng hay chưa.
- **Bài toán có 2 Dataset:**
 - Small Dataset: $1 \leq N \leq 200$
 - Large Dataset: $1 \leq N \leq 200000$
 - Trong mọi Dataset, ta luôn có $1 \leq a_i \leq 100$.
- **Giới hạn dữ liệu:**
 - 5s / Dataset.
 - 512Mb / Dataset.

Bài 6 (WC6.*):

Hoán vị của N số tự nhiên đầu tiên được định nghĩa là một dãy số nguyên dương, các phần tử có giá trị nằm trong đoạn từ 1 đến N và các phần tử đều khác nhau.

Ví dụ: Dãy $a = [2, 1, 5, 3, 4]$ là một hoán vị của 5 số tự nhiên đầu tiên. Dãy $a = [1, 2, 3]$ là hoán vị của 3 số tự nhiên đầu tiên.

Bạn được cho một dãy a là một hoán vị của N số tự nhiên đầu tiên. Tuy nhiên, vì một lỗi sơ suất nhỏ của hệ thống tạo dữ liệu, mà dãy a lại có 2 số giống nhau. Chính vì có 2 số giống nhau nên lại có 1 số không xuất hiện. Bạn hãy tìm ra 2 số đó.

- **Dữ liệu đầu vào (WC6.INP):**
 - Dòng đầu tiên gồm một số nguyên dương T ($1 \leq T \leq 20$), là số lượng bộ dữ liệu.
 - T bộ dữ liệu tiếp theo, mỗi bộ dữ liệu được tổ chức dưới dạng như sau:
 - Dòng đầu tiên chứa một số nguyên dương N.
 - Dòng thứ 2 chứa N số nguyên dương, số nguyên dương thứ i có giá trị là a_i . Dữ liệu input bảo đảm rằng các số trong dãy a hoàn toàn khác nhau và giá trị nằm trong đoạn $[1 .. N]$. Tuy nhiên duy nhất chỉ có 2 số giống nhau.
- **Dữ liệu đầu ra (WC6.OUT):**
 - Kết quả cho T bộ dữ liệu, mỗi bộ dữ liệu in ra 2 số nguyên dương là số bị lặp lại 2 lần và số không xuất hiện.

Ví dụ:

WC6.INP	WC6.OUT
3	5 3
6	1 2
5 2 1 5 4 6	4 6
3	
1 1 3	
6	
5 2 4 4 1 3	

- **Giải thích test ví dụ:**
 - Với test ví dụ 1: Dãy a có 2 số 5 giống nhau và số 3 không xuất hiện
 - Với test ví dụ 2: Dãy a có 2 số 1 giống nhau và số 2 không xuất hiện
- **Bài toán sẽ được chia làm 2 Dataset:**
 - Small Dataset: 512Mb
 - Large Dataset: Do từ lúc chấm bài các thí sinh đến giờ, hệ thống chấm đã sử dụng khá nhiều bộ nhớ, nên chỉ có thể sử dụng 6Mb cho Dataset này thôi ^_^.
 - Trong mọi Dataset, ta có ràng buộc $2 \leq N \leq 1000000$.
 - Bộ input/output của cả 2 Dataset là giống nhau. Chúng chỉ khác ở mặt bộ nhớ.
 - Đối với tin học, chúng ta cần phải quan tâm tới 1 điều đó là kích thước bộ nhớ mà chúng ta đã sử dụng trong nguồn tài nguyên của máy tính. Nếu bạn sử dụng 100 biến, mỗi biến có kiểu dữ liệu int (4 byte) thì bạn đã chiếm 400 byte trong bộ nhớ. Giới hạn kích thước các kiểu dữ liệu bạn có thể xem lại ở bảng giới hạn tại bài WC4. Bạn hãy cẩn thận trong việc sử dụng thêm mảng hay những cấu trúc dữ liệu có khả năng lưu trữ dữ liệu lớn nhé :D
- **Giới hạn dữ liệu:**
 - 3s / Dataset.

-----Chúc các bạn làm bài tốt <3 <3 <3 -----