

# NTU PK 2017

National Taiwan University

2018/01/06

Language	Version	Compile Flags	Extensions
C	gcc 5.4.0	-g -O2 -std=gnu11 -static -lm	.c
C++	g++ 5.4.0	-g -O2 -std=gnu++14 -static -lm	.cc, .cpp
Python 2	2.7.12		.py
Python 3	3.5.2		.py

Problem	Problem Name	Time Limit	Memory Limit
A	Almost String Matching	3 s	256 MB
B	BWT	5 s	256 MB
C	Omlan Union	30 s	256 MB
D	Guess the Omlan Union	1 s	256 MB
E	Olan Union	6 s	256 MB
F	Almost GCD	1 s	256 MB
G	Valid Parenthesis Substrings	1 s	256 MB
H	Mouse Trap	3 s	256 MB
I	LCM and GCD	5 s	256 MB
J	String Game	1 s	256 MB
K	Let's PK!	1 s	256 MB

*This page is intentionally left blank.*

# A. Almost String Matching

Problem ID: match

Fcrh is the god of string. He can solve string matching with many different algorithms: KMP, Gusfield, Boyer, Suffix Array, Suffix Automaton, Hash, etc. Just string matching is way too easy and way too boring for him. So he try to challenge you with “Almost String Matching”.

We say two characters are almost same if they are same or adjacent english letters. For example, **a** and **b** are almost same, **t** and **s** are almost same, **g** and **g** are almost same, but **a** and **z** are not. We say two strings  $a, b$  are almost same if  $|a| = |b|$  and all  $a_i, b_i$  are almost same for all  $1 \leq i \leq |a|$ . For example, **any** and **boy** are almost same, **stride** and **rushed** are almost same, but **shik** and **fcrh** are not (The later is much stronger!).

To solve “Almost String Matching”, you need to find all occurrences of a pattern  $s$  within a text  $t$  as an almost same substring.

## Input

The input contains 2 lines. The first line is the pattern  $s$ , and the second line is the text  $t$ .

- $1 \leq |s| \leq |t| \leq 250000$
- The strings  $s$  and  $t$  are consisting of lowercase English letters.

## Output

Please output one or two lines. The first line is the number of possible match positions. If there is any match position, output one more line with match positions in ascending order and separated by a space.

**Sample Input 1**

any amyisaboy	2 1 7
------------------	----------

**Sample Output 1****Sample Input 2**

fcrh ismuchstrongerthanshik	0
--------------------------------	---

**Sample Output 2**

## B. BWT

Problem ID: bwt

Hrcf used to be a ICPC competitor in NTU (NTU Tmt514 University), and he hates BWT (Burrows-Wheeler transform).

It all began on a windy night in 2015, when he was still a ICPC competitor. Hrcf have been dreaming of going to ICPC World Final since he's in senior high. And that contest, NTU PK 2015, is the last chance to accomplish his dream as he was about to graduate from NTU.

He spent all of his time learning algorithms, participating contests on codeforces, solving problems on OJs, optimizing workflow. Despite all the hardwork, he failed. That's all because Hrcf was not familiar with BWT and got his AC a bit slower than a rival, Fcrh.

From then on, he has been trying to alter all words that seems like a BWT result as a revenge. Unfortunately, there're too many words on the world to be altered, even **"ICPC"** could be a BWT result of **"PICC"**.

Being Hrcf's best friend, you decide to write a program to help Hrcf. So that he only needs to replace minimal amount of characters to make a word not a result of BWT of any string.

### Note

As a friend of Hrcf, you're not familiar with BWT, either. Luckily, there's a note from Fcrh.

A BWT of a string is to sort all rotations of such string in lexical order and collect the last character of each rotation in order. For example, after sorting **"PICC"**'s all rotations, you'll get **"CCPI"**, **"CPIC"**, **"ICCP"**, **"PICC"**. Hence, the BWT result will be **"ICPC"**.

### Input

There're multiple test cases in one file, one for each line. In each test case, you'll be given a string  $S$  consist of lower case alphabets.

- $1 \leq |S| \leq 10^6$
- $\sum |S| \leq 10^7$

## Output

Output a **ibwtified** string based on  $S$  or "**life is hard**" if that's impossible. You should modify as less character as possible and select the result with minimal lexical order under such constraint.

### Sample Input 1

icpc ntu	acpc ntu
-------------	-------------

### Sample Output 1

## C. Omilan Union

Problem ID: omilan

In the Omilan country, there are  $n$  cities and  $m$  undirected roads. The  $i$ -th road connects cities  $u_i$  and  $v_i$ . Each road  $(u_i, v_i)$  satisfies  $0 < |u_i - v_i| \leq 12$ .

Kai-Chi chooses a non-empty subset of cities to form an Omilan union. For each two cities  $a$  and  $b$  in the Omilan union, there must exist a path from  $a$  to  $b$  passing through no cities outside the Omilan union. In other words, the Omilan union must be connected.

Kai-Chi would like to know how many ways there are to choose such a subset, but he is afraid of large numbers. Therefore, he just wants to find this number modulo 2.

### Input

The first line contains two integers  $n, m$ , indicating the number of cities and the number of roads. Each of the next  $m$  lines contains two integers  $u_i, v_i$ , indicating that there's a road connects cities  $u_i$  and  $v_i$ .

- $1 \leq n \leq 1000$
- $0 \leq m \leq \frac{n(n-1)}{2}$
- $1 \leq u_i < v_i \leq n$
- $0 < |u_i - v_i| \leq 12$
- The graph has no self-loops
- There is at most one road between each pair of cities

### Output

Output an integer which denotes the number of possible subsets modulo 2.

#### Sample Input 1

3 2 1 2 2 3	Sample Output 1 0
-------------------	----------------------

Sample Input 2	Sample Output 2
3 3 1 2 2 3 1 3	1



## D. Guess the Omilan Union

Problem ID: guess

In the Omilan Union problem, the answer of each test case is 0 or 1.

Shik hacked into the judge system with some incredible SQL injection tricks. Although Shik couldn't download all the test cases due to permission denied, he got something useful. There are only  $n$  test cases in the problem! And he also got an oracle which could tell him the number of test cases he guessed were correct.

For example, if there are 5 test cases and the answers are 00101, the oracle will return 2 if Shik asked 01110 (the first and the third ones are correct).

Since the contest would be over soon, Shik could only guess at most 300 times. Could you help Shik to find the answers of the Omilan Union problem?

### Interaction

First, your program should read an integer  $n$  in one line, indicating the number of test cases in the Omilan Union problem.

Then, your program should start guessing the answers by output a string in one line. The length of the string should be  $n$  and the string should consist with "0" and "1".

After making each guess, you need to make sure to flush standard output (like `fflush(stdout)` in C/C++).

After each guess, there will be a result from standard input. The result would be a line with an integer, indicating the output of the oracle. When the result is  $n$ , your program should terminate the execution.

- $1 \leq n \leq 1000$

Sample Input 1	Sample Output 1
5 2 5	01110 00101

*This page is intentionally left blank.*

## E. Olan Union

Problem ID: olan

In the Olan country, there are  $n$  cities and  $m$  undirected roads. The  $i$ -th road connects cities  $u_i$  and  $v_i$ .

Han-Han chooses a **possibly empty** subset of roads to form an Olan union. An Olan union must have an Olan cycle, which is a (not necessarily simple) cycle passing through each of the roads exactly once, and must not contain any roads outside the union. (Due to differences in translation, some people also called it an Eulerian cycle.)

Among many different possible Olan unions, Han-Han prefers larger ones. That is, ones with greater *Haoness*. If an Olan union contains  $x$  edges, its *Haoness* is simply  $x^2$ . Unfortunately, the Olan country is so large that finding the maximum *Haoness* Olan union is not an easy job. So, Han-Han would rather choose a random Olan union uniformly from all valid Olan unions. Can you help him find the **expected** *Haoness* of the chosen Olan union?

### Input

The first line contains two integers  $n, m$ , indicating the number of cities and the number of roads. Each of the next  $m$  lines contains two integers  $u_i, v_i$ , indicating that there's a road connecting cities  $u_i$  and  $v_i$ .

- $1 \leq n \leq 200\,000$
- $0 \leq m \leq 340\,215.01$
- $1 \leq u_i < v_i \leq n$
- The graph has no self-loops
- There **may** be multiple roads connecting the same pair of cities

### Output

Output a real number indicating the expected *Haoness* of the chosen Olan union. Your answer will be considered correct if its absolute or relative error doesn't exceed  $10^{-6}$ .

**Sample Input 1**

3 3 1 2 2 3 1 3	4.5000000000
--------------------------	--------------

**Sample Output 1****Sample Input 2**

4 3 1 2 1 3 1 4	0.0000000000
--------------------------	--------------

**Sample Output 2****Sample Input 3**

9 9 1 2 1 3 2 3 2 4 2 4 5 6 5 7 6 7 7 9	21.5000000000
--	---------------

**Sample Output 3****Sample Input 4**

5 10 1 2 1 3 1 4 1 5 2 3 2 4 2 5 3 4 3 5 4 5	27.5000000000
--	---------------

**Sample Output 4**

Sample Input 5	Sample Output 5
2 5 1 2 1 2 1 2 1 2 1 2	7.5000000000

*This page is intentionally left blank.*

## F. Almost GCD

Problem ID: gcd

In mathematics, the greatest common divisor (gcd) of two or more integers, which are not all zero, is the largest positive integer that divides each of the integers. For example, the gcd of 8 and 12 is 4.

Shik is simply good at finding the gcd of a sequence. To be a master of the gcd problems, Shik has to solve the following problem:

Given a sequence of  $n$  integers  $a_1, a_2, \dots, a_n$ . You are allowed to increase or decrease some elements by 1. Your task is to find the maximum gcd.

For example with sequence  $(3, 5, 2, 10)$ , the maximum gcd is 3 because we can make the sequence  $(3, 6, 3, 9)$ .

Shik writes a program to solve the problem. Could you help him to check whether he is correct?

### Input

The first line contains an integer  $T$  indicating the total number of test cases.

For each test case, the first line contains an integer  $n$  indicating the length of the sequence. The second line contains a sequence of  $n$  integers  $a_1, a_2, \dots, a_n$ .

- $2 \leq n \leq 10^5$
- $\sum n \leq 10^5$
- $2 \leq a_i \leq 10^{18}$

### Output

For each test case, output an integer in one line indicating the maximum gcd.

Sample Input 1	Sample Output 1
2	3
4	3
3 5 2 10	
6	
3 14 15 9 2 6	



# G. Valid Parenthesis Substrings

Problem ID: parenthesis

For a given string  $S$  consists of lowercase letters 'a'-'z', we would like to convert each character to either '(' or ')', such that

1. It is a valid (or balanced) parenthesis string.
2. For each pair of corresponding ( and ), they are both from the same lowercase letter.

For example, if  $S = \text{aabaab}$ , then  $()(())$  is a valid conversion. However, neither  $()()()$  nor  $()((($  is a valid conversion.

In this task, you are going to report the number of substrings  $S[i..j]$  of a given string  $S$ , so that  $S[i..j]$  has a valid conversion.

## Input

The input consists of a string  $S$ .

- $1 \leq |S| \leq 1000000$ .

## Output

Output a number indicating the number of substrings that has a valid conversion.

Sample Input 1	Sample Output 1
abcabcabc	0
Sample Input 2	Sample Output 2
aabbcc	6

<b>Sample Input 3</b>	<b>Sample Output 3</b>
aabaab	4

# H. Mouse Trap

Problem ID: concave

Recently, the mouse problem in Bobo's house is getting more serious. Dozens of mice can be seen running inside the living room, the kitchen, and even Bobo's room! Hence, he decided to build a magical mouse trap. The mouse trap is a **\*\*simple\*\*** polygon with vertices chosen from the  $N$  base points. If a mouse enters the polygon, it will be disintegrated by the force field, so Bobo is basically safe there.

Now Bobo wants to maximize the area of the mouse trap, so he will have larger living space. The optimal solution is clearly the convex hull of the  $N$  base points. However, Bobo hates convex polygons, so he will choose the **concave hull** — the maximum area non-convex simple polygon that encloses all  $N$  base points (All  $N$  base points should be inside or on the border of the polygon). Can you help him calculate its area?

A polygon is simple if no two non-adjacent edges intersect, and adjacent edges intersect at exactly the common vertex. A simple polygon is non-convex if at least one of its interior angles is strictly larger than  $180^\circ$ .

## Input

The first line contains an integer  $N$  indicating the number of base points. For the following  $N$  lines, the  $i$ -th line contains 2 integers  $x, y$  separated by space representing the coordinates of the  $i$ -th base point.

- $1 \leq N \leq 500\,000$
- $|x_i|, |y_i| \leq 10^9$
- No two points coincide.

## Output

The output should contain an integer  $2A$ , where  $A$  is the area of the concave hull. Since the coordinates of base points are integers,  $2A$  is guaranteed to be an integer, too. If the concave hull doesn't exist, please output  $-1$ .

**Sample Input 1**

3 0 0 0 1 1 1	-1
------------------------	----

**Sample Input 2**

4 0 0 1 2 1 -2 -2 0	8
---------------------------------	---

**Sample Input 3**

7 0 0 4 0 4 3 3 3 2 2 1 3 0 3	23
--	----

**Sample Input 4**

5 0 0 1 0 2 0 0 2 1 2	-1
--------------------------------------	----

# I. LCM and GCD

Problem ID: lcmgcd

Give you two positive integers  $N$  and  $M$ .

Please calculate  $(\prod_{X \in S} lcm(X_1, X_2, \dots, X_n)^{gcd(X_1, X_2, \dots, X_n)}) \bmod 10^9 + 7$ , where  $S$  is the set contains all sequence which the length is  $N$  and all number in it are positive integers between 1 and  $M$ .

## Input

The first line contains an integer  $T$  indicating the total number of test cases.

For each test case, there is only one line contain two number  $N, M$ .

- $1 \leq T \leq 10^3$
- $1 \leq N \leq 10^9$
- $1 \leq M \leq 2 \times 10^5$
- the summation of  $M$  in all test cases is not larger than  $2 \times 10^5$

## Output

For each test case, output an integer in one line indicating the answer of this test case.

Sample Input 1	Sample Output 1
2 1 3 5 5	108 574847666

*This page is intentionally left blank.*

# J. String Game

Problem ID: string

There is a game about four strings.

Give you four string  $S1$ ,  $T1$ ,  $S2$ ,  $T2$ .

Both length of  $S1$  and  $T1$  are  $N$  and both length of  $S2$  and  $T2$  are  $M$ .

We denote the  $i$ -th character of a string  $s$  as  $s[i]$ .

the target of this game is change the string  $S1$  to  $T1$ .

The game contains two phase.

In the first phase, you can do following operation as many time as you want:

- Choose two distinct integers  $i$  and  $j$  ( $0 \leq i, j < N$ ) and swap  $S1[i]$  and  $S1[j]$ .

Then in the second phase, you can do following operation as many time as you want. But this time, you want to minimize the time you do it.

- Choose two integer  $i$  and  $j$  ( $0 \leq i < N$  and  $0 \leq j < M$ ) satisfying  $S1[i]$  equal to  $T2[j]$  and swap  $S1[i]$  and  $S2[j]$ .

If it's impossible to change  $S1$  to  $T1$ , output  $-1$ . Otherwise, output the minimum times you need for the operations in the second phase.

## Input

There are four lines. Each line contains a string. They are  $S1$ ,  $T1$ ,  $S2$ , and  $T2$  in this order.

- $1 \leq N \leq 2 \times 10^5$
- $1 \leq M \leq 2 \times 10^5$
- all string are consist of lowercase English letters.

## Output

Output an integer in one line indicating the answer.

**Sample Input 1**

ae  
ed  
bcd  
abc

**Sample Output 1**

3

**Sample Input 2**

a  
b  
ab  
ab

**Sample Output 2**

-1



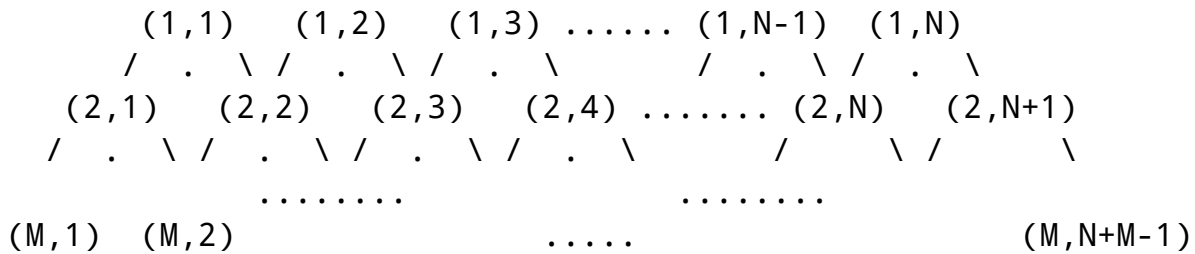
# K. Let's PK!

Problem ID: marble

One day, Hanhan is very boring, so he decides to play marbles in the corner. But later, a chicken, named Bobo comes and laughs at him. "What a poor lonely weird old man!" Hanhan gets angry, and he decides to "PK" with the chicken Bobo.

The chicken Bobo accepts Hanhan's challenge. They agree to play a game similar to pinball and decide who is the winner.

To be more specifically, the binball board which they are going to play look like this.



Where  $(x, y)$  represent the  $y$ -th possible position in  $x$ -th row, and the single dots represent the pin. In the beginning of the challenge, They agree to place  $K$  marbles in the first row. That is, they would put marbles at position  $(1, a_1), (1, a_2), \dots, (1, a_K)$ . Then, due to the same reason why Newton would be hit by an apple, the marbles would start to fall at the same speed. That is, the marbles initially in row 1 will start to fall to row 2 and then 3, 4 ... Also, because of the pin, the marble would have equal probabilities to fall to the left and to the right. So, a marble at position  $(1, 2)$  will fall to  $(2, 2)$  or  $(2, 3)$  both with probability 0.5. Eventually, all the marbles will fall to  $M$ -th row and stop.

At the beginning, Hanhan will choose  $K$  distinct integer  $b_1, b_2, \dots, b_K$  in  $[1, N + M - 1]$  and bet that the marbles will fall to position  $(M, b_1, b_2, \dots, b_K)$  at last. Similarly, Bobo will choose  $K$  distinct integer  $c_1, c_2, \dots, c_K$  and bet the final position of the marbles. One who makes the correct prediction wins the game. If both of them are right, they will flip a fair coin to determine who wins the game. If both of them are wrong, they will restart the game until someone wins. To restart the game means that they pick up all the marbles from the board, place them to their initial position and wait for the result again without changing  $a_i, b_i, c_i$ .

But there is a small problem here. Those marbles are actually made by a material called Deuterium. If two of the marbles accidentally collide (For example, two marbles at  $(1, 1)$  and  $(1, 2)$

both fall to  $(2, 2)$ ), something similar to  $E = mc^2$  will happen, and Hanhan and Bobo will forget about the game and run for their lives. After they think they have run far enough, they will restart the game.

Just after they have decided the position,  $\chi\chi$  ( $\chi$  is a greek letter which has name “chi” and pronounced like “Kai”, so  $\chi\chi$  should be read as “Kai-chi”) comes and finds it interesting. He decides to calculate the probability  $p$  that Hanhan will win.

Of course,  $\chi\chi$  knows the answer within a second. Now give you which position they bet, could you find out the probability  $p$ ?

## Input

The first line contains an integer  $T$  indicating the total number of test cases. Each test case begin with a line contains 3 integers  $N, M, K$  separated by a space. In the following 3 lines, There are  $K$  integers separated by space each, represent  $a_i, b_i$  and  $c_i$ .

- $1 \leq T \leq 1000$
- $1 \leq N, M \leq 60$
- $1 \leq K \leq 25$
- $a_i, b_i, c_i$  are sorted in ascending order

## Output

Output a positive real number  $p$  represent the probability that Hanhan will win. Solution with absolute or relative error in  $10^{-6}$  would be judged as correct. If the game last forever and no one wins, puts -1 instead.

Sample Input 1	Sample Output 1
2	0.75
2 3 2	0.0
1 2	
2 3	
1 4	
4 3 2	
1 3	
1 6	
2 4	

*This page is intentionally left blank.*