

2.1 Problem 1

Step 1

Gene Expression

In this problem you will need to submit your solution in any of the supported programming languages. The code will be tested on all the prepared inputs that will be not shown to you. Fortunately, this is the only problem in the round with hidden tests ;-)

This problem has two different versions with different constraints. If your solution passes all the tests of the easy version you will get 450 points. If your solution passes all the tests of the hard version you will get 750 points. In total you can get 1200 points at maximum. You need to submit correct solution to both versions to get the full score.

The problem is described in the next step. Constraints for each version are also written in the corresponding steps.

Step 2

Gene Expression

A scientist John Doe decided to engineer a bacteria with a minimal possible genome. His idea is that if he removes all the genes that are lowly expressed something magical will happen. For this he carried out an RNA-sequencing experiment to measure gene expression levels.

For this experiment he prepared the genome reference and the genome annotation with description of genes' exonic structure (can someone tell John that what he has is definitely not a bacteria?). After the RNA-sequencing was done he aligned the reads to the reference and got genomic intervals where reads align.

After almost all the work has been done he stuck with a problem of estimating gene expression from this data. His working idea is that gene expression should be estimated by a number of reads that intersect only with this gene. Sadly, he skipped his Python course and can't implement this method himself. Help him calculate the gene expression levels!

Input format

The first line of the input contains two integers n and m ($1 \leq n, m$) – the number of genes and the number of reads.

Each of the next n lines contains a description of a gene. It consists of a list of exon intervals, separated by a space. Each such interval is closed and represents the first and the last covered nucleobases. The list of intervals is represented as a non-decreasing list of positive integers with an even length. Elements $2i - 1$ and $2i$ of the list represent the i -th exon of the gene and they could be equal, that is when the exon consists of one nucleobase. Two exons from the same list do not intersect. The elements of the list are separated by spaces. The integers do not exceed 10^9 .

Each of the next m lines describes an alignment of the read to the genome in a form of list of genomic intervals. Each such interval is closed and represents the first and the last covered nucleobases. The list of intervals is represented as a non-decreasing list of integers with an even length. Elements $2i - 1$ and $2i$ of the list represent the i -th alignment interval of the read and they could be equal, that is when interval consists of one nucleobase. Two intervals from the same list do not intersect. The elements of the list are separated by spaces. The integers do not exceed 10^9 .

Constraints for the easy version: $n, m \leq 300$ and the total number of intervals throughout all genes and reads does not exceed 10^5 .
Constraints for the hard version: the total number of intervals throughout all genes and reads does not exceed $2 \cdot 10^5$.

Output format

The i -th line of the output should contain the value of the expression of the i -th gene in the input order.

Example

Sample input 1:

```
2 4
1000 1200 2000 2400
3000 3300
1100 1160
1190 1200 2000 2050
3280 3300 3500 3550
1500 1560
```

Sample output 1:

```
2
1
```

Sample input 2:

```
2 1
1000 2000 3000 4000
1500 3500
2000 3000
```

Sample output 2:

```
0
0
```

Step 3

Gene Expression (Easy version)

Constraints for the easy version. The number of genes and the number of reads do no exceed 300. Total number of intervals throughout all genes and reads does not exceed 10^5 .

Sample Input 1:

```
2 4
1100 1100 2000 2400
3000 3300
1100 1160
1190 1200 2000 2050
3280 3300 3500 3550
1500 1560
```

Sample Output 1:

```
2
1
```

Sample Input 2:

```
2 1
1000 2000 3000 4000
1500 3500
2000 3000
```

Sample Output 2:

```
0
0
```

To solve this problem please visit <https://stepik.org/lesson/32224/step/3>

Step 4

Gene Expression (Hard version)

Constraints for the hard version. Total number of intervals throughout all genes and reads does not exceed $2 \cdot 10^5$.

Sample Input 1:

```
2 4
1100 1100 2000 2400
3000 3300
1100 1160
1190 1200 2000 2050
3280 3300 3500 3550
1500 1560
```

Sample Output 1:

```
2
1
```

Sample Input 2:

```
2 1
1000 2000 3000 4000
1500 3500
2000 3000
```

Sample Output 2:

```
0
0
```

To solve this problem please visit <https://stepik.org/lesson/32224/step/4>

