2.5 Problem 5

Step 1

Pathway Dissection

This problem is interactive: you can submit either the request to the system or the answer. The communication with the system is done through text files. You are given 7 different tests for the proposed problem in 7 different tabs. Depending on the quality of the answer you will be given some fraction of points on the corresponding test. The description of queries and the score formula are given in the next step.

The points are distributed as follows: for tests 1-3 we give 100 points each, for tests 4-6 we give 250 points each and for test 7 we give 350 points.

For the tests of this problem we took the real life data.

You could download all tests using this link: https://stepik.org/media/attachments/lesson/40926/tests.zip.

Step 2

Pathway Dissection

As John Doe already knows, many reactions could happen inside a cell. He read in the book that each reaction is actually bi-directional, that is could be go in both directions, and can happen only if a cell contains at least one of the proteins that catalyze this reaction. John already have understood which reactions could happen, but he does not know which proteins catalyze which reactions.

To uncover the unknown he set up the following protocol consisting of a series of experiments. In each experiment he prepares up to 96 wells with a cocktail of chemicals and proteins he chooses. After some time in each well all possible reactions happen, and he measures what chemicals are present in each well. His instrument is very accurate and sensitive and all the chemicals that are possible to obtain are measured and only them.

He thinks with this protocol it should be pretty easy to figure out which proteins can catalyze wich reactions. However, he still don't know how to code and asks for your help. Help him design which cocktails should be prepared and then to make a list of reaction catalyzers from experimental results.

Input format

As input you are given two files: a file with reactions and a file with a list of proteins that catalyze at least one reaction.

The first file contains reactions, one per line. A reaction is represented as two lists of chemicals (strings) separated by '<=>', in each list the chemicals are separated by '+'. These reactions are taken from the real life data.

The file with proteins contains a space-separated list of proteins (strings) that carry at least one reaction.

Note. The names of chemicals and proteins do not change from test to test, i.e., the chemical 'c1' represents the same chemical in all the tests.

Example of reactions files:

```
c1 + c2 <=> c3 + c4
c4 + c5 + c1 <=> c3
c1 <=> c4
```

Example of proteins file:

```
p1 p2 p3
```

Requests

You could make two types of requests to the system: either to perform a set of experiments or to check the answer.

Experiments

Request format

The first line of the request should contain '?'. Then at most 96 descriptions of experiments could follow. Each experiment should be described by two lines. The first line should contain a space-separated list of unique chemicals in the environment at the beginning. The second line should contain a space-separated list of unique proteins. Each chemical and protein should appear in the corresponding input files.

```
?
c3 c4
p1
c3
p2
c4
p1 p2
```

Response format

The response is given in the submission outcome. To open the outcome you have to click on 'Show full feedback'. The response contains the results of the experiments, one per line. The result of an experiment consists of a space-separated list of chemicals after performing all possible reactions.

```
c3 c4 c1 c2
c3
c4 c1
```

Answer

When you want to check your answer, your request should be formatted as follows. The first line should contain '+'. Each of the next lines, one per reaction in the input, should contain unordered list of proteins that can catalyze this reaction. If you do not know the list of proteins for the reaction, you could simply put '-'.

```
+
p1 p2
-
p2
```

Score

The score linearly depends on the number of reactions for which your list of proteins is correct.

Step 3

Pathway Dissection (test #1)

Upload your requests on the test #1.

Download the test at https://stepik.org/media/attachments/lesson/40926/1.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/3

Step 4

Pathway Dissection (Test #2)

Upload your requests on the test #2.

Download the test at https://stepik.org/media/attachments/lesson/40926/2.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/4

Step 5

Pathway Dissection (Test #3)

Upload your requests on the test #3.

The test could be downloaded at https://stepik.org/media/attachments/lesson/40926/3.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/5

Step 6

Pathway Dissection (Test #4)

Upload your requests on the test #4.

Download the test at https://stepik.org/media/attachments/lesson/40926/4.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/6

Step 7

Pathway Dissection (Test #5)

Upload your requests on the test #5.

Download the test at https://stepik.org/media/attachments/lesson/40926/5.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/7

Step 8

Pathway Dissection (Test #6)

Upload your requests on the test #6.

Download the test at https://stepik.org/media/attachments/lesson/40926/6.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/8

Step 9

Pathway Dissection (Test #7)

Upload your requests on the test #7.

Download the test at https://stepik.org/media/attachments/lesson/40926/7.zip.

To solve this problem please visit https://stepik.org/lesson/40926/step/9