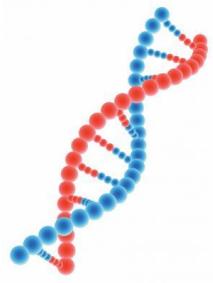


## Problem 23: CSI

Difficulty: Hard

Originally Published: Code Quest 2015



### Problem Background

Last month Bobby reported to the police that someone broke into his house and stole his favorite pair of jeans. He was extremely distraught and asked the police to do whatever it takes to find the burglar. The police looked for DNA samples from around the house. Luckily they found only one, and they believe it belongs to the burglar. Unfortunately, the police's database searching program is on the fritz so they want you to create a program to search their database of all the people in the town and find the closest DNA match.

### Problem Description

DNA samples are rarely perfect – parts of a sample could be missing or contaminated. Therefore, a direct match is not usually found. The police use the “longest common subsequence” method to find the most likely suspect. Two strings share a common subsequence if they have the same set of letters in the same order, but the letters in the subsequence do not necessarily have to be adjacent to one another.

For example:

- The string ABCDEFGHIJKLMNOPQRSTUVWXYZ has subsequences KNOT and MOW, but not PAT.
- The longest common subsequence between the strings ABAFCDEF and BCFEDE is BCDE.

### Sample Input

The first line of your program’s input, received from the standard input channel, will contain a positive integer representing the number of test cases. Each test case will include:

- A string representing the DNA sequence found by the police.
- A positive number **N** representing the number of items in the town’s database.
- **N** lines, containing one item in the town’s DNS database. Each item will be in the form NAME=DNA\_STRING. The DNA sequence letters is made up of the letters A, C, G, and T (representing the four nucleotide bases found in DNA: A: Ademine, C: Cytosine, G: Guanine, T: Thymine).

1

TTTCAGTCTTCGAAACGT

5

David=ATGGCCATCGGGGTCGGCCGTCGCTGGC

Bobby=TTTCAGTCTTCGACGT

Brian=TTGAATGGCGTCTGGCAAACGGCTT

Jose=TTGACCATGACGTGCCACTGGC

Kyle=TTGACCAGGGGAATAAACTTTCT

## Sample Output

Your program's output should display the name of the person whose DNA was the closest match to the sample according to the length of the longest subsequence algorithm. If the length of the longest subsequence is shared by two or more people, output the names in alphabetical order separated by commas.

Bobby