



# Cheating Detection (11pts, 20pts)

## Practice Submissions

Attempt 1

S ✓ ✓

Mar 30 2021, 01:28

Last updated: Mar 30 2021, 01:30

PROBLEM

ANALYSIS

## Problem

100 players are competing in a 10000-question trivia tournament; the players are numbered from 1 to 100. Player  $i$  has a skill level of  $S_i$  and question  $j$  has a difficulty level of  $Q_j$ . Each skill level and each question difficulty are chosen uniformly at random from the range  $[-3.00, 3.00]$ , and independently of all other choices. For example, a player can have a skill level of 2.47853 and a question can have a difficulty level of  $-1.4172$ .

When player  $i$  tries to answer question  $j$ , the probability that they answer it correctly is  $f(S_i - Q_j)$ , where  $f$  is the [sigmoid function](#):

$$f(x) = \frac{1}{1 + e^{-x}}$$

where  $e$  is [Euler's number](#) (approximately 2.718...), the mathematical constant. Notice that  $0 < f(x) < 1$  for all  $x$ , so  $f(S_i - Q_j)$  is always a valid probability. Each of these answer attempts is chosen at random independently of all other choices.

There is one exception: exactly one of the players is a cheater! The cheater is chosen uniformly at random from among all players, and independently of all other choices. The cheater behaves as follows: before answering each question, they flip a fair coin. If it comes up heads, they do not cheat and the rules work as normal. If it comes up tails, they secretly look up the answer on the Internet and answer the question correctly. Formally, they decide whether to cheat at random with 0.5 probability for each question, independently of all other choices.

The results of a tournament consist of only the per-question results (correct or incorrect) for each player. Apart from the general description above, you do not know anything about the skill levels of the players or the difficulties of the questions.

You must correctly identify the cheater in at least  $P$  percent of the test cases. That is, you must succeed in at least  $P \cdot T/100$  out of  $T$  cases.

## Input

The first line of the input gives the number of test cases,  $T$ . The second line of the input gives the percentage of test cases,  $P$ , that you must answer correctly for your solution to be considered correct.  $T$  test cases follow. Each case consists of 100 lines of 10000 characters each. The  $j$ -th character on the  $i$ -th line is 1 if the  $i$ -th player answered the  $j$ -th question correctly, or 0 if they answered it incorrectly.

## Output

For each test case, output one line containing Case # $x$ :  $y$ , where  $x$  is the test case number (starting from 1) and  $y$  is the number of the cheater (with player numbers starting from 1).

## Limits

Time limit: 60 seconds.

Memory limit: 1 GB.

 $T = 50$ .

## Test Set 1 (Visible Verdict)

 $P = 10$ .

## Test Set 2 (Visible Verdict)

 $P = 86$ .

## Sample

### Sample Input

```
1
0
001110100010101000001000000100010101000000011000110101
-----
99 lines of input omitted.
```

### Sample Output

```
Case #1: 59
```