A Wise Person in the Mountain

There used to be a very smart person who also thought they knew everything. As they grew older and got wiser, they realized they can only have some level of confidence in what they think the answer is. This person has moved to a mountain where they could focus on their studies, but every once in a while they would welcome pilgrims and answer their questions.

In your college, you were given an assignment to evaluate the accuracy of the wise person's answers. Your task is to measure their ability to recognize correctness of statements posed as "Yes or No" questions. To this kind of question, wise person would always answer with a "Yes" and the confidence can be anything from 0% to 100%. For example, if wise person is certain the answer is "No", they would say their confidence in "Yes" is 0%.

You've compiled a list of questions and for each question you've noted an answer which is widely considered to be correct. From the confidence provided by wise person, you are forming a prediction. Here is an example:
Question: Is it true that the Earth is flat?
Correct Answer: Yes *(see task #3 of this assignment for more info)*
Wise person's confidence in "Yes": 5%
Prediction: No *(see below for explanation how a prediction is being formed)*

Your task is to aggregate all the answers into a set of measured quantities explained below.

Forming the prediction

To evaluate the correctness of answers wise person has provided, you need to convert their confidence in "Yes" to a definite answer, which we will call prediction. If the confidence for "Yes" is larger or equal to some threshold T, the prediction will be "Yes", otherwise the prediction is a "No". Here are a few examples:

Confidence in Yes	Threshold	Prediction
90%	65%	Yes
30%	50%	No
10%	5%	Yes
80%	95%	No

Measurement quantities

Prediction types

In the dataset you have:

- a number of questions whose correct answer is Yes. We call these Positive questions.
- a number of questions whose correct answer is No. We call these Negative questions.

The types of predictions are summarized in the table:

	Positive (P)	**Negative (N)**
Predicted positive	True positives (TP)	False positives (FP)
Predicted negative	False negatives (FN)	True negatives (TN)

TPR (True Positive Rate) and **FPR** (False Positive Rate) can be computed as follows:

- TPR = TP / P
- FPR = FP / N

EER (Equal Error Rate)

TPR and FPR depend on the threshold you chose. By varying the threshold (T) from 0% to 100% we get different values for TPR and FPR. There is a threshold T for which FPR and (1 - TPR) have the same value. This value is called EER.

Datasets

As most of the students do, you've started working on this assignment a bit too late. Due to this, the correct answers and wise person's answers are all scattered around in different files on your computer. Each answer is in separate file: correct answers are in files named ca#.txt, and wise person's answers are in files named wpa#.txt (where # is a unique number of some question you had prepared). Unfortunately, in all this mess, you've lost some of the data, so not every correct answer has a matching wise person's answer, and vice versa.

There are two data sets:

- *Public data set* is used for developing your solution. After you submit your solution, you will be able to see how well your solution performs against this data set. *Public data set* is not used for calculating the final score. Public data set is available **here**.
- *Private data set* is used for testing your solution. The final score will be measured against this data set. *Private data set* and the final score will be available after the homework finishes. *Private data set* contains different data than the *public data set*, but the type of data (number of questions, depth of folder hierarchy...) is roughly the same.

Input

Path to the input folder will be given through standard console input. If in doubt on how input is parsed, please refer to the solution of the first task (pixel task).

Input folder may contain an arbitrarily deep subfolder tree which contains only files named as ca#.txt and wpa#.txt.

Each ca#.txt contains only one word: **Yes** or **No**. Each wpa#.txt contains the confidence in "Yes" expressed in percents. For example:

```
90%
```

Note, wise person gives only integer confidences.

Output

The results should be printed to standard output in the form of a comma separated string of values.

These values are:

- (a) number of positive and (b) negative questions (this number does not depend on the availability of wise person's answers),
- (c) number of questions which can be used for evaluation (valid questions, they have both correct answer and prediction),
- (d) TPR and (e) FPR for 70% threshold,
- (f) Equal Error Rate.

Example output:

```
420,320,539,0.84,0.08,0.10
```

Note that to compute TPR, FPR and EER, you can only use questions which have both correct answer and prediction.

Scoring

- Each of the correct numbers (a, b, c) will get you 8 points each,
- Each of the numbers (d, e) with error less than 0.005 will get you 4 points each,
- number (f) with error less than 0.01 will get you 8 points.

Hence, Each test case will bring you up to 40 points. Private set has 10 test cases for a total of 400 points for this task.

All the values in the output are evaluated independently. If for some reason you cannot provide some value, just write nothing in between commas. For example, here is expected result and result to be evaluated:

	Positive questions	Negative questions	Valid questions	TPR @ 70%	FPR @ 70%	EER
Expected result	1000	1000	1700	0.778	0.125	0.168
Result	1000	/	1650	0.78	0.08	0.16
Scores	8	0	0	4	0	8

Corresponding output you'd write for the results from the above table is:

```
1000,,1650,0.78,0.08,0.16
```

Constraints

- Time limit per test case is 5s.
- Memory limit per test case is 128MB.