HackerRank

Byte The Correct Apple

The word "Apple" could generally refer to one of these two:

- (a) Apple Inc., the great Computer giant.
- (b) Apple, the fruit

You are provided a text file, with a number of lines. Each line contains either a sentence or a paragraph or a text snippet which could either be related to Apple, the computer company, or the apple, the fruit. Your task is to perform disambiguation between these two groups and identify which one is being referred to. It is possible that the plural or the possessive form of Apple might exist in some of the tests (apples, Apple's).

Training Data

You are provided with two text files, which contain near-complete text from the Wikipedia for Apple Inc. as well as apple the fruit. For offline inspection and access, you could access these two files here:

Text from Wikipedia entry on Apple-Computers Text from Wikipedia entry on Apple the fruit

Also, when you submit your program, you can assume that these two text files are available in the directory where your program is run, and their names are "apple-computers.txt" and "apple-fruit.txt".

Input Format

```
An Integer N, no more than 100.

line_1
line_2
line_3
line_4
...
line_N
```

Constraints

```
N <= 100 
 Each line will have not more than 1000 characters in it. 
 Assume that the encoding is UTF-8.
```

Output Format

```
computer-company
fruit
computer-company
fruit
..
..
N lines of output
```

Sample Input

```
Apple already plans to buy back `$100 billion in shares, including $`16 billion worth last quarter. Icahn probably pounded the dinner table he and Cook shared recently for their much-reported bread-breaking at Icahn's New York apartment. Apple's cash stash currently sits at a whopping `$145 billion but only $`43 billion is in the U.S., which is why Icahn wants to float bonds to cover a buy back. Fortunately, there are "low-chill" apple varieties for temperate climates. (Chilling hours are defined as nonconsecutive hours of winter temperatures below 45 degrees.) As a general guide, if you live on or near the coast, your garden gets only 100 to 200 chilling hours. Inland San Diego gardens get about 400 to 500 chilling hours — still considered "low chill."

If this seems a bit like déjà vu, you'll recall that Apple just held an event to unveil two new iPhone models — the 5c and 5s — back on September 10.

.
```

Sample Output

```
computer-company
fruit
computer-company
.
.
```

Explanation

```
In the first chunk of text, Apple Inc. is being referred to.
In the second chunk, Apple the fruit is being referred to.
In the last chunk of text, again, it is Apple Inc. which is being referred to.
```

A note on the test cases

The sample test case has 10 tests and the hidden test case has 100.

To develop a better model or algorithm you are encouraged to create tests of your own using text available online.

Scoring

Score for a test case will be M * (c-w)/N. Where, M is the maximum score assigned for the test case, c is the number of correct answers, w is the number of incorrect answers, and N is the total number of tests in the file.

In case w > c (i.e. if more predictions are incorrect than correct) a zero score will be assigned. Score will only be based on the hidden test case.