VI BXComp

6th Freshmen's Programming Championship of Information Systems 2016

6th Stage – Challenge 3

Sentiment Analysis

Social media have become an important platform of communication; through them, opinions and sentiments are expressed in simple conversations or messages. The number of active users and the volume of data created daily is impressive and, in this context, researchers and companies have been collecting this data for large-scale content analysis – for instance, sentiment analysis.

One of the sentiment analysis' definitions relies on the detection of the author's attitude, emotion or opinion about a given topic expressed in a text. The task of sentiment analysis is especially challenging in the context of analyzing user generated content from the Internet and it is where companies like LexisNexis operate.

They are teaming up people to work in a sentiment analysis' project that aims to provide the general mood of hypothetical tweets. You were invited to participate in the first phase of this project and now you need to develop a program that evaluate sentiments on tweets and classifies the most positive and negative tweets about a particular topic.

The tweet's topic is defined through a hashtag. Each tweet will have *only one* hashtag, that can appears anywhere in the tweet's text, but does not interfere in the tweet's mood – only states the topic! This hashtag is interpreted as a string initialized by the symbol #, without any blank space.

The actual evaluation of sentiments in the tweets considers a simplified heuristic that classifies a set of positive and negative expressions. By this way, the mood of the given tweets will be determined by the difference between the amount of positive and negative words of the tweet. For example, a tweet that has four positive words and two negative words is less positive than another tweet with three positive words and no negative words. In the case of a tie and considering the input as a timeline, the advantage should be given to the one that was read first.

Task

Your task consists of developing a program that, given a heuristic that defines a set of positive and negative expressions, a set of topics represented through hashtags and a set of tweets wrote by Twitter users, outputs the most positive and the most negative tweets for each topic. The heuristic is case insensitive.

Input

The input simulates a Twitter news' feed, presented in the following structure:

- The first two input lines are the problem's heuristics: the first line consists of a set of positive expressions, while the second contains a set of negative expressions – all of these expressions are properly separated by a single blank space.
- On the third line, there is a sequence of hashtags defining the topics, also separated by a single space.
- Then, there is a blank line, followed by a sequence of tweets.
- The end of the feed is represented by the word "END".

Specifications:

- Each tweet is formed by two lines:
  I. the first one identifies the tweet's number and the user that wrote it in the format "Tweet **X** - **@user**", where ***X*** is the number of the tweet and ***user*** is the respective author. There won't be blank spaces in the user's name
  II. the second one contains the tweet itself, with only one hashtag.
- After each tweet there will be a blank line;
- There will be at least two tweets about each topic;
- Each user appears only once in the feed.

Output

For each topic, in alphabetical order, your program must print the result using three lines: the first one containing the topic's name, including the symbol #, while the second and third lines should contain the name of users whose tweets were, respectively, the most positive and the most negative ones about the topic. After each topic's result, there must be a blank line.

Input Example

```
yes go hello parabens happy :)
no don't dark sarcasm goodbye
#song #each

Tweet 01 - @PaulMcCartney
You say yes, I say no You say stop and I say go go go, oh no You say goodbye and I say hello go go #song

Tweet 02 - @petsi
#each Parabens a todos os participantes do BXComp :)

Tweet 03 - @PinkFloyd
We don't need no education #song We don't need no thought control No dark sarcasm in the classroom

Tweet 04 - @UmAlunoDeSI
Hoje teve prova na #each

Tweet 05 - @Pharrell
Because I'm happy Clap along if you feel like a room without a roof Because I'm happy #song

Tweet 06 - @OutroAlunoDeSi
Indo para a #each participar do bxcomp

END
```

Output Example

```
#each
@petsi
@UmAlunoDeSI

#song
@PaulMcCartney
@PinkFloyd
```