

# Segment the Twitter Hashtags

Given a set of Twitter hashtags, split each hashtag into its constituent words. For example:

- **wearethepeople** is split into **we are the people**
- **mentionyourfaves** is split into **mention your faves**

## Input Format

The first line contains an integer,  $N$ , denoting a number of hashtags.  
Each of the  $N$  subsequent lines contains a single hashtag.

## Dictionaries and Corporuses of Text

We don't strictly prescribe a particular dictionary or corpus or a set of features. To get started, you may find it useful to embed this list of [5000 common words](#) as a dictionary in your program. For more effective segmentation models, you are encouraged to use your own word list, or corpus, or features extracted from a corpus, as required by whatever model you choose. [Project Gutenberg](#) is a good starting point, but keep in mind that language and its usage has evolved and transformed over time.

You may use serialization to build and compress your model offline and to decompress and use it from your program. If you end up with a corpus or model that is too large, you may compress and serialize it, then deserialize it from within your code using zlib (that is in Python) or another tool. This means that your code will contain a compressed string representing the dictionary which will then be de-compressed and used. You can take a look at this code submitted during CodeSprint5 [here](#). For Java users, you might want to look up *java.util.zip.GZIPInputStream* for this purpose.

## Constraints

- $5 \leq N \leq 50$
- The hashtags will *not* contain named entities, other than the names of countries and their abbreviations (e.g.: US, UK, UAE, etc.).}
- The hashtags may occasionally contain slang phrases, such as "faves" (a slang abbreviation for "favorites").

## Scoring

Your score is proportional to the number of hashtags which you split correctly. The final score is computed only on the basis of the hidden test case.

**You may make no more than 15 submissions for this problem, during the contest.**

## Output Format

There should be  $N$  lines of output, where each line  $i$  contains the space-separated set of segmented words corresponding to line  $i$  of the input.

## Sample Input

```
5
wearethepeople
mentionyourfaves
nowplaying
```

thewalkingdead  
followme

## Sample Output

we are the people  
mention your faves  
now playing  
the walking dead  
follow me