Extra Credit for HW#3

Students who complete the following exercise will have one point added to their score on HW#3

Inverted Index of Bigrams using Map-Reduce

Now that you are familiar with setting up and running Hadoop jobs on GCP, you will now modify your `InvertedIndexJob.java` script to generate an inverted index of bigrams (instead of unigrams).

Your existing Mapper class emits (word, docID) pairs which are then aggregated in the Reducer class. You will have to modify your Mapper class to emit ("word1 word2", docID) pairs instead. The reducer remains unchanged.

Once you modify your class(es), create the jar `invertedindex__bigrams.jar` and dispatch a Hadoop job **on the `devdata/`** in the same manner as before.

The output will look something like this:

```
a buff  5722018235:1 5722018508:1
a buffer    5722018235:12 5722018301:34 5722018101:2 5722018496:1 5722018508:21
a bugs  5722018301:1 5722018235:2 5722018508:1
a bulging   5722018235:2 5722018101:1
a bulldozer 5722018235:2 5722018301:2 5722018101:1 5722018496:1 5722018508:2
a bullet    5722018301:9 5722018235:7 5722018101:1 5722018496:4 5722018508:7
```

Submission Instructions

To get credit for this task, create another text file `index_bigrams.txt` with the index entries for the following bigram phrases (generated from `devdata/`):

1. computer science
2. information retrieval
3. power politics
4. los angeles
5. bruce willis

You can apply grep on the output file in the same way you did for the previous exercises.

Submit this file along with your modified java code (rename it as `InvertedIndexBigrams.java`) as part of the `index.zip` archive.

Note: do not submit anything else for this exercise!