1. The code main.py provides a simple Python implementation of counting words in a text file using map reduce mentality. The code does the following:

   (a) Load the file's words into a list

   (b) Map each element of the list to a (key:value) pair

   (c) Group the set of pairs by key

   (d) Reduce the set of pairs by counting unique entries.

   Rather than writing all of the scaffolding code ourselves, some Python libraries which implement Mapreduce already exist, such as mrjob. We'll use mrjob for the rest of this homework. You can learn how to download and install it by referring to the website: `https://mrjob.readthedocs.io/en/latest/`.

   Using mrjob, the code we wrote above can be reduced to what we have in mrjobcode.py, which is not only much easier to read, but also leaves far less room for error in implementation of, e.g., the shuffle process.

   We are trying to count the number of words in Bertrand Russell's book "An Inquiry into Meaning and Truth" (AIIMAT.txt). To run the code, you should do the following from your Termina, Cygwin, or Virtual Machine:

   $ python mrjobcoe.py AIIMAT.txt

   Report the outputs you get in text files or screenshots. (50 pts)

2. We want to use the Map Reduce mentality to solve the problem of computing mutual friend lists. The input data will be of the form:

   1 : 2,3,6,7,8

   2 : 1,3,4,7

   3 : 1,2,5,8

   4 : 2,5,7,8

   5 : 3,4,6

   6 : 1,5,8

   7 : 1,2,4,8

   8 : 1,3,4,6,7

   where the first number on each line corresponds to a particular person, and the list of numbers following the colon : corresponds to their list of friends.

   Your code should output the mutual friends lists in the form

   1,2 : 3,7

   1,3 : 2,8

   1,6 : 8

   1,7 : 2,8

1,8 : 3,6,7

2,3 : 1

2,4 : 7

2,7 : 1,4

3,5 :

3,8 : 1

4,5 :

4,7 : 2,8

4,8 : 7

5,6 :

6,8 : 1

7,8 : 1,4

where the ordered pair preceding the colon keys the ordered list of mutual friends following the colon. You should avoid printing keys corresponding to pairs of persons who are not friends in the first place.

Your code should be written in Python. Use the code template.py for your solution. (70 pts)