**W1D4**

Question 1

**Create a Java simulator for WordCount with In mapper combining.**

Create a InMapperWordCount class. You should be able to specify the number of Input-splits (the same as the number of Mappers) and the number of reducers.

You have already created the Mapper class. (See W1D1). Just modify it appropriately. Rest is as in W1D3.

Add necessary output statements so that your output matches one shown in the attached file “InMapperTestDataOutput”.

**Question 2. Illustrate WordCount In-Mapper Combining Algorithm.**

Also assume that there are three input splits:

Input split 0 : [apple lemon mango salmon wheat apple mango]

Input split 1 : [barley salmon apple orange carrot rice salmon]

Input split 2 : [mango carrot lemon carrot apple rice tuna]

Since there are three input splits, there will be three Mappers. Thus, Input split I is handled by Mapper I (I = 0, 1, 2). Assume that there are three reducers. Note that Mapper I and Reducer I run on the same machine (I = 0, 1, 2).

Further, let the partitioner assign all words less than letter ‘k’ to Reducer 0, all words greater than ‘r’ to Reducer 2 and everything else to Reducer 1.