

INF 551 – Fall 2018 (Afternoon)

Quiz 8: MapReduce (10 points), 10 minutes

Consider the WordCount example seen in class. Assume that input directory has 2 small files whose contents are shown below. Suppose that NO combiner is used (combiner is not yet covered in class).

File1	File2
dog and cat this cat	cat likes dog this cat and that dog

1. [2 points] How many map tasks will be created?

2

2. [4 points] Show the input and output of each call to the map function for each map task.

1. Input: (0:'dog and cat') Output: ('dog':1), ('and':1), ('cat':1)
2. Input: (12:'this cat') Output: ('this':1), ('cat':1)
3. Input: (0:'cat likes dog') Output: ('cat':1), ('likes':1), ('dog':1)
4. Input: (14:'this cat and that dog') Output: ('this':1), ('cat':1), ('and':1), ('that':1), ('dog':1)

3. [4 points] Assume that **two** reducers are used. One handles all words which start with a letter from 'a' to 'g'; the other handles the rest. Show the input and output of each call to the reduce function for each reduce task.

Reduce task 1:

Input: ('and':[1,1]), ('cat':[1,1,1,1]), ('dog':[1,1,1])

Output: ('and':2), ('cat':4), ('dog':3)

Reduce task 2:

Input: ('this':[1,1]), ('likes':1), ('that':1)

Output: ('this':2), ('likes':1), ('that':1)