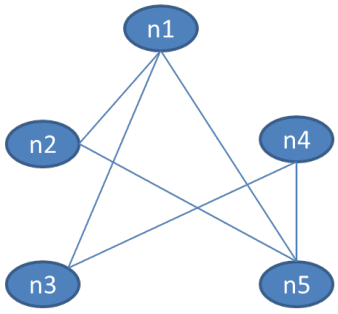
Name: _____ USC ID: _____

INF 553 – Spring 2018

Quiz 1: Hadoop MapReduce (10 points), 15 minutes

Consider computing the degree for each node in a undirected graph using Hadoop MapReduce. The graph information is stored in a text file, where each line records an edge in the graph. For example, here is an example graph and its data file.



Graph

n1,n2
n1,n3
n2,n5
n3,n4
n5,n4
n1,n5

Graph data file

1. [5 points] Write the logic of **map** function of Map task in pseudocode. What will the map function output for the graph above?

   (2 points) Pseudo-code:

   ```
   map(key, value): // key = offset of line; value = line content
   for each word w in value:
   emit(w, 1)
   ```

   (3 points) Output:(n1,1), (n1,1), (n1,1), (n2,1), (n2,1), (n3,1), (n3,1), (n4,1), (n4,1), (n5,1), (n5,1), (n5,1)

2. [5 points] Write the logic of **reduce** function in pseudocode. What will be the **input** and **output** of the reduce function for the graph above?

   (2 points) Pseudo-code:

   ```
   reduce(key, values):
       result = 0
   for each count v in values:
       result +=v
   emit(key, result)
   ```

   (1.5 points) Input:(n1,[1,1,1]), (n2,[1,1]), (n3,[1,1]), (n4,[1,1]), (n5,[1,1,1])

   (1.5 points) Output: (n1,3), (n2,2),(n3,2),(n4,2), (n5,3)