

INF 553 – Spring 2017

Quiz 2: Matrix Multiplication via Hadoop MapReduce (10 points)

15 minutes

Consider the **two-phase** approach to computing $C = AB$ using Hadoop MapReduce, where

1,1,1	1,1,1
1,2,2	2,2,1
2,1,1	3,1,2
2,3,1	

Matrix A

Matrix B

$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}$. Each matrix is stored in separate file in sparse format. See above. Recall

that there are two map tasks, each handling one of the matrices. For example, map task for A processes matrix file for A. And for each line, it generates a key-value pair (key is offset, value is line content) and calls the map function with the pair.

1. [3 points] List all key-value pairs produced by **map** tasks for matrix A and B in Phase 1.

A[ij]: emit (j, ('A', i, A[i,j]))
 B[jk]: emit (j, ('B', k, B[j,k]))
 (1, ('A', 1, 1)), (2, ('A', 1, 2)), (1, ('A', 2, 1)), (3, ('A', 2, 1))
 (1, ('B', 1, 1)), (2, ('B', 2, 1)), (3, ('B', 1, 2))
 If zero is not deleted, you will lose 1 point.

2. [3 points] List all key-value pairs produced by **reduce** task(s) in Phase 1.

emit ((i,k), A[i,j]* B[j,k])
 j=1: ((1,1), 1), ((2,1), 1)
 j=2: ((1,2), 2)
 j=3: ((2,1), 2)
 Reduce task will complete A[i,j]* B[j,k]. (1.5 points)
 Right value. (1 point)
 Key-value pair format. (0.5 point)

3. [3 points] List all **input** (key, value-list) pairs to reduce function of reduce task(s) in Phase 2.

((1,1), [1]), ((1,2), [2]), ((2,1), [1, 2])
 1 error (2points); 2 errors (1.5 points)
 (key, value-list) format 0.5 point

4. [1 point] List all key-value pairs **output** from reduce task(s) of Phase 2.

((1,1), 1), ((1,2), 2), ((2,1), 3)
 Not all pairs correct. 0.5 point
 Completely correct. 1 point