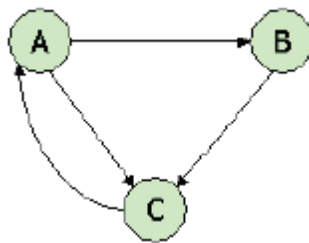




PageRank & MapReduce

Due Wed. Nov. 14th/Thurs. Nov. 15th
(at the beginning of class)

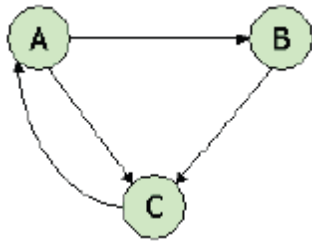
1. What is MapReduce and how it is suitable for processing large datasets?
2. What are the two major components of MapReduce? Explain what each does.
3. Consider three Web pages with the following links:



Assuming no "taxation," compute the PageRanks a , b , and c of the three pages A, B, and C, using iteration, starting with the "0th" iteration where all three pages have rank $a = b = c = 1$. Compute as far as the 5th iteration, and also determine what the PageRanks are in the limit. Then, identify the true statement from the list below.

- a) After iteration 4, $a = 3/2$
- b) In the limit, $a = 6/5$
- c) After iteration 5, $a = 21/16$
- d) After iteration 5, $c = 1$

4. Consider three Web pages with the following links:



Suppose we compute PageRank with $\beta=0.85$. Write the equations for the PageRanks a , b , and c of the three pages A, B, and C, respectively. Then, identify in the list below, one of the equations.

- a) $a = .9c + .05b$
- b) $.95a = .9c + .05b$
- c) $.85b = .575a + .15c$
- d) $c = b + .575a$