MapReduce and PageRank

Question 1:

Suppose our input data to a map-reduce operation consists of integer values (the keys are not important). The map function takes an integer *i* and produces the list of pairs (*p*, *i*) such that *p* is a prime divisor of *i*. For example, map (12) = [(2,12), (3,12)].

The reduce function is addition. That is, reduce (p, [i_1, i_2, ..., i_k]) is (p, i_1+i_2+...+i_k).

Compute the output, if the input is the set of integers 15, 21, 24, 30, 49.

The output of map function is

map (15) = [(3, 15), (5, 15)]

map (21) = [(3, 21), (7, 21)]

map (24) = [(2, 24), (3, 24)]

map (30) = [(2, 30), (3, 30), (5, 30)]

map (49) = [(7, 49)]

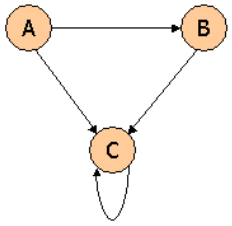
These are the respective prime divisors of inputs

The output of reduce function is

reduce (2, 54), reduce (3,90), reduce (5,45), reduce (7, 70)

Question 2:

Consider three Web pages with the following links:



Suppose we compute PageRank with a β of 0.7, and we introduce the additional constraint that the sum of the Page Ranks of the three pages must be 3, to handle the problem that otherwise any multiple of a solution will also be a solution. Compute the page Ranks *a*, *b*, and *c* of the three pages A, B, and C, respectively.

Value of a, b, or c as we iterate are: a

All PageRank is multiplied by .7 before distribution, and .3 is then added to each new PageRank.

a = $\beta(0) + (1 - \beta) \rightarrow .3$

b = $\beta\ (\frac{a}{2}) + (1 - \beta) \rightarrow .7(\frac{a}{2}) + .3$

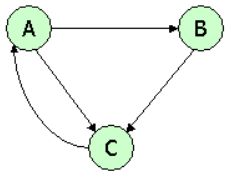
c = $\beta\ (\frac{a}{2} + b + c) + (1 - \beta) \rightarrow .7(\frac{a}{2} + b + c) + .3$

That immediately tells us a = .3. We can then use the second equation to discover b = $.7(.\frac{3}{2}) + .3 = .405$. Finally, the third equation simplifies to c = .7(.555 + c) + .3, or .3c = .6885. From this equation we get c = 2.295

To compute the subs of each two of the variables:

a + b = .705, a + c = 2.595, and b + c = 2.7

Question 3:



Suppose we compute PageRank with β=0.85. Write the equations for the PageRanks *a, b,* and *c* of the three pages A, B, and C, respectively.

Formula:

$$a = \beta * c + (1 - \beta)\frac{1}{3}$$

$$b = \beta * \frac{a}{2} + (1 - \beta)\frac{1}{3}$$

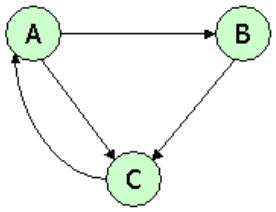
$$c = \beta * (\frac{a}{2} + b) + (1 - \beta)\frac{1}{3}$$

Here $\beta = 0.85$

a = 0.85*c + (1 - 0.85) $\frac{1}{3}$, a = 0.85c + 0.05

b = 0.85*0.5*a + 0.05, b = 0.425a + 0.05
c = 0.85*[0.5*a + b] + 0.05, c = 0.425a + 0.85b +0.05

Question 4:



Assuming no "taxation," compute the page Ranks *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 page Ranks are in the limit.

a = c

b = $\frac{a}{2}$

c = $\frac{a}{2} + b$

At 0th iteration: a = 1; b = 1; c = 1

At 1st iteration: a = c = 1; b = $\frac{1}{2}$; c = $\frac{1}{2}$ + 1 = $\frac{3}{2}$

At 2nd iteration: a = c = $\frac{3}{2}$; $b = \frac{a}{2} = \frac{1}{2}$; $c = \frac{1}{2} + \frac{1}{2} = 1$

At 3rd iteration: a = c = 1; b = $\frac{a}{2} = \frac{\frac{3}{2}}{2} = \frac{3}{4}$; $c = \frac{3}{4} + \frac{1}{2} = \frac{5}{4}$

At 4th iteration: a = c = $\frac{5}{4}$; b = $\frac{a}{2} = \frac{1}{2}$; c = $\frac{5}{4}$

At 5th iteration: a = $\frac{5}{4}$; b = $\frac{5}{8}$; c = $\frac{9}{8}$