Map Reduce 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.

Map Tasks:

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)]

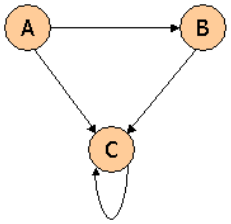
Above outputs will be the inputs to the reduce tasks.

Reduce Tasks:

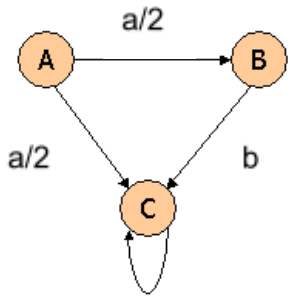
Reduce (3, [15, 21, 24, 30]) = reduce (3, [15+21+24+30]) = (3, 90)
Reduce (2, [24, 30]) = (2, [24+30]) = (2, 54)
Reduce (5, [15, 30]) = (5, 15+30) = (5, 45)
Reduce (7, [21, 49]) = (7, 21+49) = (7, 70)

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 PageRanks 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 PageRanks *a*, *b*, and *c* of the three pages A, B, and C, respectively.



c

Here n = 3, β = 0.7
Page rank formula;

$$r_j = \sum_i^j \frac{r_i}{d_i} \beta + (1 - \beta)(1/n)$$

So, here for node A,

a = 0 + (1 – 0.7) $\frac{1}{3}$

= 0 + (0.3) $\frac{1}{3}$

= 0 + 0.1

a = 0.1

For node B,

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

= ((0.7)*(0.1)*(0.5)) + $\frac{(1-0.7)}{3}$

= 0.035 + $\frac{0.3}{3}$

b = 0.135

For node C,

c = βc + βb + β. $\frac{a}{2}$ + (1 - β). $\frac{1}{3}$

= (0.7)c + (0.7)(0.135) + $\frac{(0.7)(0.1)}{2}$

c = 0.765

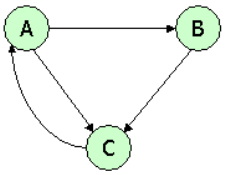
Now, a + b + c = 0.1 + 0.135 + 0.765 = 1
But the constraint is a + b + c = 3, so multiplying all the values with 3.

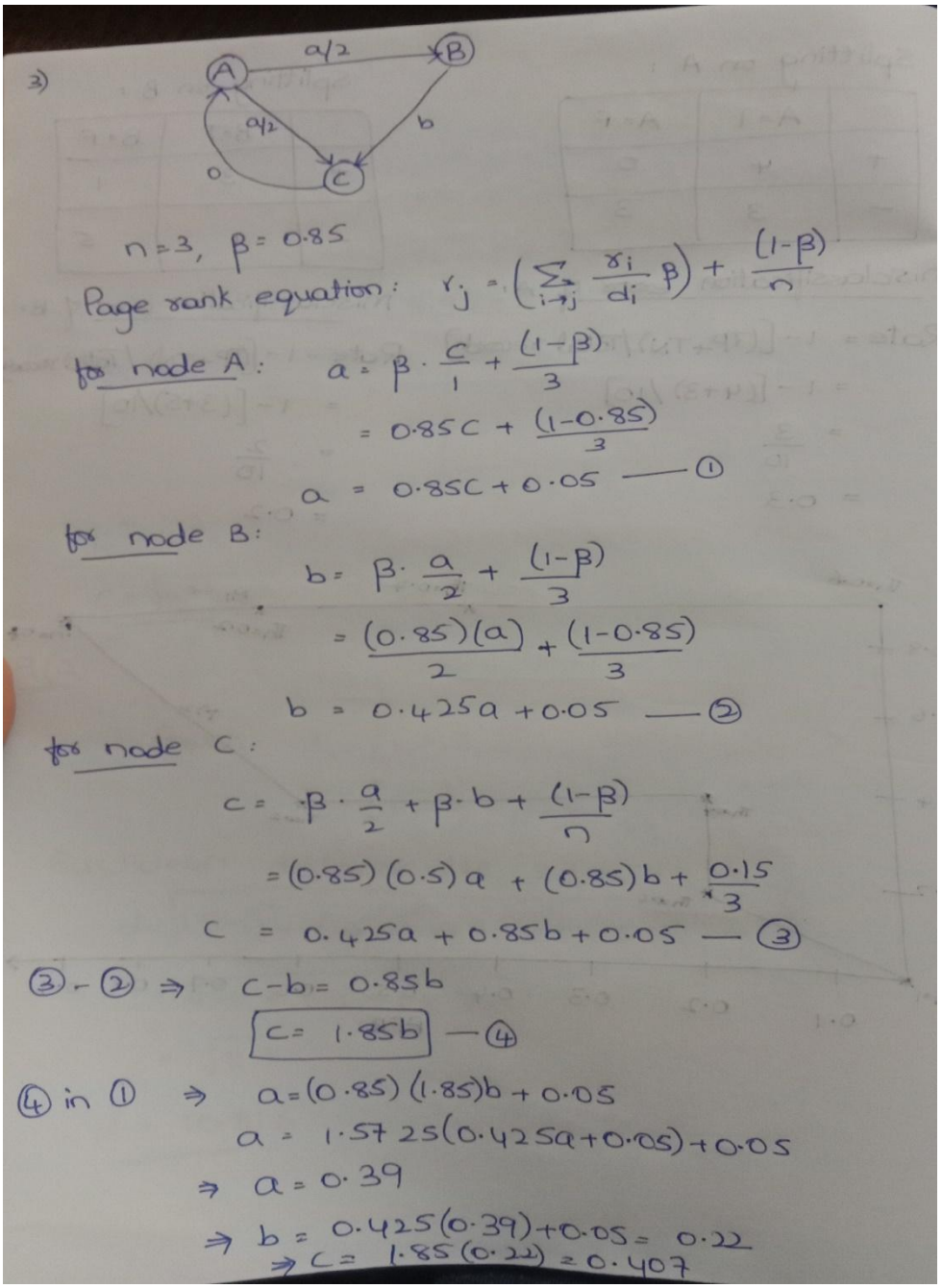
a = 3 * 0.1 = 0.3
b = 3 * 0.135 = 0.405
c = 3 * 0.765 = 2.295

Now the value of, a + b + c = 0.3 + 0.405 + 2.295 = 3

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





Question 4:

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.