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).

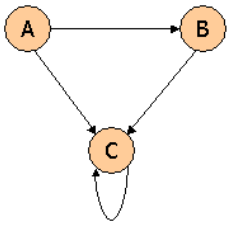
Compute the output, if the input is the set of integers 15, 21, 24, 30, 49.

Ans:

```
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)
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 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.

Ans:

$$a = (1 - 0.7)/3 = 0.1$$

$$b = \beta * a/2 + (1-\beta) * 1/3$$

$$c = \beta(a/2 + c + b) + (1-\beta) * 1/3$$

$$b = 0.7 * 0.1/2 + (1-0.7) * 1/3$$

$$b = 0.135$$

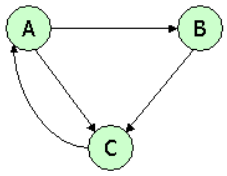
$$c = 0.7 * (0.1/2 + c + 0.135) + (1-0.7) * 1/3$$

$$c = 0.765$$

$$constraint : a + b + c = 3$$

$$=> a * 3 = 0.3, b * 3 = 0.405, c * 3 = 2.295$$

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.

Ans:

$$a = 0.85c + 0.05$$

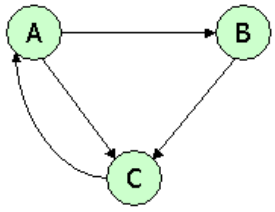
$$b = 0.85 * 0.5a + 0.05$$

$$c = 0.85 * 0.5a + 0.85b + 0.05$$

$$b = 0.425a + 0.05$$

$$c = 0.425a + 0.85b + 0.05$$

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.

Ans:

PageLink matrix

	A	B	C
A	0	0	1
B	1/2	0	0
C	1/2	1	0

Pagerank vector

1
1
1

*Compute pagerank = PageLink matrix * pagerank vector*

Iterations	A	B	C
1	1	0.5	1.5
2	1.5	0.5	1
3	1	0.75	1.25
4	1.25	0.5	1.25
5	1.25	0.625	1.125