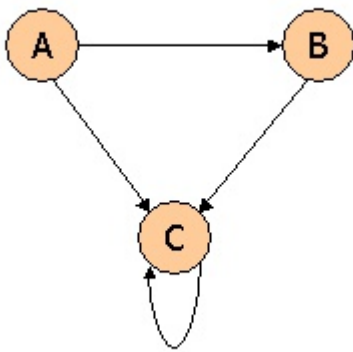


BDDL_HW1

Big Data & Deep Learning Homework1

Problem 1

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, C, respectively. Then, identify from the list below, the true statement.

Solution:

```
"""
Created on Mon Apr 06 16:31:56 2015

@author: jszhujun2010's PC
"""

import numpy as np

#Initialize variables
r=np.matrix([float(1)/3, float(1)/3, float(1)/3]).T
```

```

beta=0.7

M=np.matrix([[0, 0, 0], [0.5, 0, 0], [0.5, 1, 1]])

S=np.matrix([(1-beta)/3, (1-beta)/3, (1-beta)/3]).T

t=0

while True:
    t+=1
    r_p=r[:]
    r=beta*M*r+S
    diff=abs(r_p-r)
    if diff.sum() < 0.0001:
        break

print("The final r is", 3*r)
print("The total iterations is", t)

```

The final answer is $3*r$ (for the purpose of sum of them is 3), which we can see from console:

```

The final r is [[0.3]
 [0.405]
 [2.295]]

The total iterations is 3

```

From the information above, we can conclude that:

Problem 2

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, $\text{map}(12) = [(2, 12), (3, 12)]$.

The reduce function is addition. That is, $\text{reduce}(p, [i_1, i_2, \dots, i_k])$ is $(p, i_1 + i_2 + \dots + i_k)$.

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

Solution:

This is an easy problem! We can get the result after the map operation:

```
map (15) [] = [] [(3, 15), (5, 15)]
map (30) [] = [] [(3, 21), (7, 21)]
map (24) [] = [] [(2, 24), (3, 24)]
map (30) [] = [] [(2, 30), (3, 30), (5, 30)]
map (49) [] = [] [(7, 49)]
```

Then, we can sort them by keys and group them:

```
(2, 24), (2, 30)
(3, 15), (3, 21), (3, 24), (3, 30)
(5, 15), (5, 30)
(7, 21), (7, 49)
```

Finally is the reduce operation:

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

And that is the final answer.