INFORMATION RETRIEVAL – SHORT EXERCISES III – EVALUATION IN INFORMATION RETRIEVAL AND PAGERANK

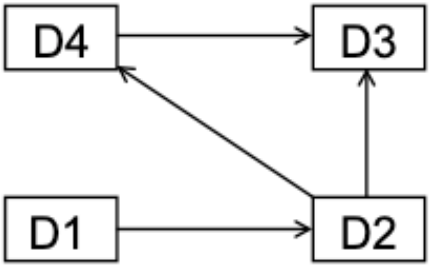
I. Consider an information need for which there are 4 relevant documents in the collection. A system run on this collection returned the top 10 results for which the relevance is judged as follows (R – relevant; N – non-relevant):

R N R N N N N N R R

What is the recall at 6 (R@6)? Answer:

What is the Mean Average Precision? Answer:

II.Consider the web graph presented below to the left. It involves four pages D1-D4 and four links. Fill in the stochastic matrix M given to the right.



0		0	0
1		0	0
0		0	1
0		0	0

Write the equation for PR(D3) without dumping factor *q*? Answer: PR(D3) =

Which page has the greatest PageRank (without computing the exact PR values)? Answer:

An oracle has evaluated D2 as trusted and D4 as spam. What is the starting vector d for TrustRank?

Answer: d = [, , ,]