**Unit 3 Algorithmics**

**Submit Task – Week 10**

AE

BE

DE

CE

Consider the four pages above. The arrows indicate that e.g. A links to B, but B does not link to A.

1. If d = 0.85, what does this signify in the PageRank algorithm?

That’s the damping factor, it signified that there’s an 85% probability that a user will continue to follow links from a webpage.

1. Calculate the PageRank for these pages for the first four iterations. What simple check can you do to make sure your numbers are correct?

Iteration 1:

A: (0.15/4) + 0.85 \* (0.25/2) = 0.13125

B: 0.85 \* (0.25/2) + 0.85 \* (0.25/2) = 0.31875

C: 0.85 \* (0.25/2) + (0.15/2) + 0.85 \* (0.25/2) = 0.31875

D: (0.15/4) + 0.85 \* (0.25/2) = 0.13125

Iteration 2:

A: (0.15/4) + 0.85 \* (0.31875/2) = 0.116015625

B: 0.85 \* (0.13125/1) + 0.85 \* (0.31875/2) = 0.395703125

C: 0.85 \* (0.13125/1) + (0.15/2) + 0.85 \* (0.31875/2) = 0.329296875

D: (0.15/4) + 0.85 \* (0.31875/2) = 0.116015625

Iteration 3:

A: (0.15/4) + 0.85 \* (0.395703125/1) = 0.13389609375

B: 0.85 \* (0.116015625/1) + 0.85 \* (0.395703125/1) = 0.3628515625

C: 0.85 \* (0.116015625/1) + (0.15/2) + 0.85 \* (0.395703125/1) = 0.3628515625

D: (0.15/4) + 0.85 \* (0.395703125/1) = 0.13389609375

Iteration 4:

A: (0.15/4) + 0.85 \* (0.3628515625/1) = 0.132428515625

B: 0.85 \* (0.13389609375/1) + 0.85 \* (0.3628515625/1) = 0.3796697265625

C: 0.85 \* (0.13389609375/1) + (0.15/2) + 0.85 \* (0.3628515625/1) = 0.3796697265625

D: (0.15/4) + 0.85 \* (0.3628515625/1) = 0.132428515625

A good check for PageRank calculations is to verify that the sum of all PageRanks for all pages adds up to 1 (or very close to 1 due to rounding errors).

Sum of probabilities = 0.132428515625 + 0.3796697265625 + 0.3796697265625 + 0.132428515625 ≈ 1

1. Why does the PageRank for page ‘C’ stabilise so quickly? Its cause its in a loop with page ‘B’ and both pages only point to each other which creates a closed system where the influence of external pages becomes minimal over iterations, causeing it to stabilize
2. If we add another page ‘E’ which is linked from C and does not link to any other pages, explain what special procedure exists for such a node.

Instead of the usual 0.85 \* (PageRank of linked page / total outgoing links from that linked page), for Page E, it would be x / (total pages) to ensure that Page E’s rank is distributed among all pages