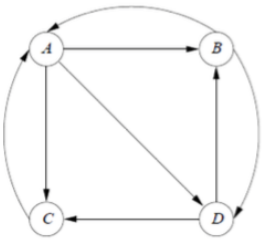
Quiz 9: Link Analysis

1) (1pt) What is the transition matrix for the graph below?

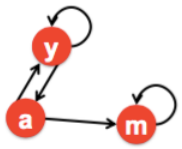


$$M = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1/2 & 1 & 0 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \end{array}$$

2) (2pts) Using power iteration to calculate the PageRank, what is v^1 for the graph in questions 1(v^0 is the initial vector)?

$$V^1 = Mv^0 = \begin{bmatrix} 0 & 1/2 & 1 & 0 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 9/24 \\ 5/24 \\ 5/24 \\ 5/24 \end{bmatrix}$$

3) (2pts) Give a random teleport $\beta = 0.8$ for calculating the PageRank, what is v^1 for the graph below?



M =

	y	a	m
y	½	½	0
a	½	0	0
m	0	½	1

$$0.8 \underbrace{\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 1 \end{bmatrix}}_{M} + 0.2 \underbrace{\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}}_{[1/N]_{NxN}} = \underbrace{\begin{array}{c} \\ y \\ a \\ m \end{array} \begin{array}{c} \begin{array}{ccc} y & a & m \end{array} \\ \begin{bmatrix} 7/15 & 7/15 & 1/15 \\ 7/15 & 1/15 & 1/15 \\ 1/15 & 7/15 & 13/15 \end{bmatrix} \end{array}}_{A}$$

$$v^1 = Av^0 = \begin{bmatrix} 1/3 \\ 1/5 \\ 7/15 \end{bmatrix}$$

OR

$$v' = \beta M v + (1-\beta)e/n$$

$$= 0.8 \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} + 0.2 \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/5 \\ 7/15 \end{bmatrix}$$

4) (2pts) What is the stationary distribution for ransom walks described in class? How do we use the property of stationary distribution to find the page ranks?

The Rank vector(which satisfies P(t+1)=Mp(t)=p(t))

Gaussian elimination and power iteration, we can also use the Eigen vector of M to find the page rank.

5) (1pt) How do we overcome spider traps in calculating page ranks in class? What does it work?
Teleports(taxation)
Because it gives chance to jump to any links, and make the graph strongly connected

6) (2pts) How do we use prune and propagate to overcome dead ends in calculating page ranks in class? What is the weakness of this approach?
Refers to slide "LinkAnalysis" page 79

- **Prune and propagate**
  - Preprocess the graph to eliminate dead-ends
  - This may create new dead-ends
  - Might require multiple passes
  - Compute page rank on reduced graph
  - Approximate values for dead-ends by propagating values from reduced graph

Weakness: may create new dead-ends and require multiple passes