

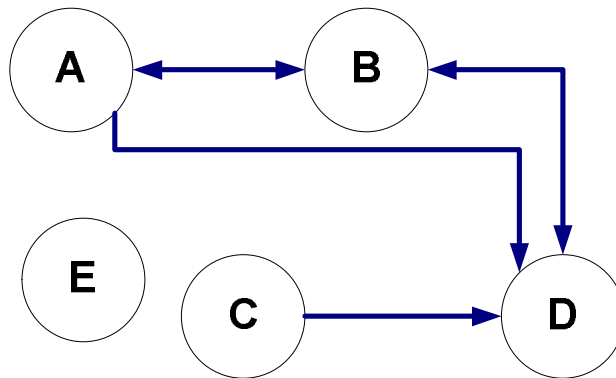
Foundations of Information Retrieval

Exercise 11

Issue date: 31.01.2013
Exercise session on: 31.01.2013
Questions to: Elena Demidova: demidova@l3s.de

Link analysis with PageRank

1. Given is the following graph:



1a. Order the nodes in the graph based on the estimated PageRank values (You do not have to compute PageRank now!). Justify your answer.

1b. Create the link matrix A' with teleportation for this graph. Use the teleportation probability of 25%.

1c. Given the PageRank formula:

$$\vec{x}^{k+1} = (1 - c)\vec{x}^k A + \frac{c}{N}\vec{e}$$

\vec{e} is $\vec{1}$. In \vec{x}_0 the random walk is uniformly distributed. Compute the vector \vec{x} for the first five iterations of the PageRank formula ($k = 0..4$) for this graph. Round to 5 decimal places!

Crawling

1. Sketch the architecture of a Crawler. How will this architecture modified in a distributed system?