Multimedia Information Extraction and Retrieval
Summer Term 2012
Exercise Sheet 7

Ralf Möller, Karsten Martiny

Exercise Session:
Thursday, June 21, 2012, 8.00-8.45, SBS95-D1025

1. Three documents d_1, d_2, d_3 with the following properties are given:

   - d_1 contains a link to d_2
   - d_2 contains a link to every document in the collection
   - d_3 contains links to d_1 and d_2

   Compute the Pagerank for each document (assuming no teleporting). In order to determine the Pagerank, perform the following steps:

   (a) Construct the Markov chain for this scenario.
   (b) Give the transition matrix P for the corresponding Markov process.
   (c) Check whether the Markov process is ergodic.
   (d) Determine the steady state of the process.
   (e) Give the Pagerank score for each document.

2. Explain the idea of *teleporting* when determining the Pagerank scores. What would change in the above situation, if the system using teleporting with a 10% chance?

3. What are the limitations of the above scoring procedure?