WOLFGANG K. SEILER

Tel. 2515

December 2, 2023

Mathematics and Information, Exercise sheet 11

Problem 1: (*10 points*)

In the *mini web*

$$
\begin{array}{ccccc}
1 & \leftarrow & 2 & \rightarrow & 3 \\
\downarrow & & \uparrow & \nearrow & \uparrow \\
4 & \leftarrow & 5 & \overset{\leftarrow}{\rightarrow} & 6 \\
\downarrow & \searrow & \downarrow & & \downarrow \\
7 & \rightarrow & 8 & \leftarrow & 9
\end{array}
$$
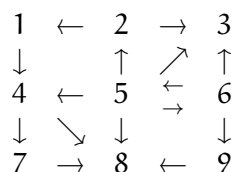
somebody submits a query consisting of only one term, and this term only occurs in document one.

a) Which part of the web will be considered after one, two, and three iterations?

b) Recursively determine the authority and the hub weights of the nine documents for the whole collection!

Problem 2: (*10 points*)

Some collection contains six documents, in which the following words occur with frequencies given in parenthesis:

D_1: *Shannon (7), Entropy (1), Information (2) Cryptanalysis (1)*

D_2: *Shannon (1), Boltzmann (1), Entropy (8), Information (5)*

D_3: *Boltzmann (2), Entropy (4), Clausius (1), Information (3)*

D_4: *Boltzmann (5), Entropy (1), Clausius (1), Heat (5)*

D_5: *Information (4), Shannon (2), Code (6), Cryptanalysis (3), Key (4)*

D_6: *Kelly (5), Shannon (1), Bet (6), Portfolio (8), Information (3)*

D_7: *Las Vegas (5), Shannon (3), Kelly (1)*

D_8: *Shannon (20), Juggling (2), Roboter (4), Unicycle (1)*

a) Determine the global weight of each term using first the entropy based definition, and then the inverse document frequency!

b) Compute the local weight of the term *Shannon* in these eight document, using the different definitions introduced in the course!

c) Determine the first and the last column of the term document matrix for each of these definitions and the entropy based global weight!

d) Determine the relevance of documents one and eight with respect to the queries *Shannon* and *Shannon information!*