Sprachtechnologie

7. Übung

Prof. Tobias Scheffer
Paul Prasse
Michael Großhans
Sebastian Arzt
Philipp Schmidt

Sommer 2014

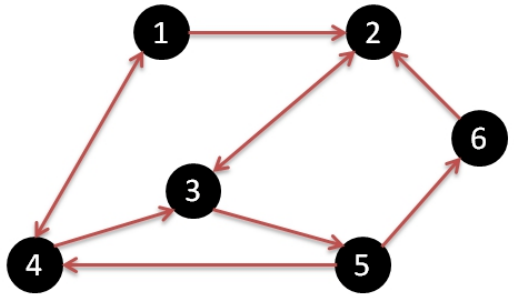
Ausgabe am: 02.06.14
Besprechung am: 16.06.14

Aufgabe 1 *Edit-Distanz*

Bestimmen Sie mit dem in der Vorlesung „Indexieren und Suchen“ vorgestellten Algorithmus die Edit-Distanz der Wörter *abloratu* und *abitur*. Geben Sie die Transformationsregeln an, um ausgehend von dem ersten Wort, das zweite Wort mit möglichst wenigen Transformationen zu erzeugen.

Aufgabe 2 *HITS*

Angenommen, Sie haben eine Linkstruktur zwischen Dokumenten (1 bis 6) bestimmt (siehe Abbildung), wobei alle Kantengwichte auf 1 festgelegt sind. Bestimmen Sie mit Hilfe des HITS-Algorithmus aus der Vorlesung das relevanteste Dokument. Dabei enthalten nur die Dokumente $1, 3, 4, 5$ relevante Schlüsselwörter. Führen sie dafür die Authority und Hub Update Regel $k = 5$ mal aus.



Aufgabe 3 *PageRank*

Bestimmen Sie mit dem PageRank Algorithmus die Scores für alle Dokumente für den Graphen aus Aufgabe 2. Welches Dokument erhält den höchsten Ranking-Score?