

Übungsblatt 1

5. November 2008

Hinweis: Für jede korrekt bearbeitete Teilaufgabe gibt es einen Punkt. Die Abgabe der Hausübungen ist bis spätestens zum Beginn der nächsten Vorlesung möglich – entweder persönlich direkt vor der Vorlesung oder per Einwurf in den Briefkasten des Instituts (Informatikzentrum, zweiter Stock, vor Raum 236).

Aufgabe 1 (Klassifikationssysteme)

- Die Association for Computing Machinery (ACM) pflegt ein eigenes System zur Klassifikation wissenschaftlicher Dokumente aus dem Bereich der Informatik. Die letzte Revision des ACM Computing Classification System erfolgte 1998. Wie würde man ein Buch über das Boolesche Retrieval-Modell mit diesem System klassifizieren?
- Wie würde man dieses Buch gemäß dem Klassifikationssystem der amerikanischen Library of Congress einordnen? (Es reicht aus, wenn Sie die ersten vier Buchstaben/Ziffern des Klassifikationszeichens angeben.)

Aufgabe 2 (Boolesches Retrieval-Modell)

Eine Dokumentenkollektion enthält die Dokumente d_1 bis d_6 . Diese sind wie folgt durch die Indexterme t_1 bis t_8 repräsentiert:

$$\begin{array}{lll} d_1: \{t_1, t_4, t_6, t_7\}, & d_2: \{t_2, t_4, t_8\}, & d_3: \{t_1, t_3, t_4\}, \\ d_4: \{t_2, t_6, t_7\}, & d_5: \{t_1, t_4\}, & d_6: \{t_1, t_3, t_6\}. \end{array}$$

- Bestimmen Sie den zu dieser Kollektion gehörigen invertierten Index.
- Welche Treffer liefert das Boolesche Modell für die folgenden Anfragen?

$$\text{„}(t_3 \text{ OR } t_4) \text{ BUT NOT } t_6\text{“} \qquad \text{„}(t_1 \text{ AND } t_6) \text{ OR NOT } t_6\text{“}$$

- Formen Sie die Anfragen aus dem vorherigen Aufgabenteil jeweils zu äquivalenten Anfragen in konjunktiver und disjunktiver Normalform um.
- Erläutern Sie wie die Auswertung der im vorherigen Aufgabenteil erhaltenen Normalform-Anfragen erfolgt. Geben Sie dazu für jede der vier Anfragen die Folge der ausgeführten Mengenoperationen sowie die zugehörigen Zwischenergebnisse an.
- Geben Sie zu jeder der folgenden Treffermengen eine (möglichst kurze) Anfrage an, die die jeweilige Treffermenge liefert.

$$\{d_2\}$$

$$\{d_6\}$$

$$\{d_3, d_5\}$$