

# Einführung in die Datenbanktheorie

Wintersemester 2016/2017

## Übungsblatt 5

**Bearbeitung:** in den Übungen am 7./8. Dezember 2016

### Aufgabe 1:

(15 + 10 Punkte)

- (a) Finden Sie zu jeder der beiden Semijoin-Anfragen (wobei  $b$  eine Konstante ist)

$$Q_1 := R(x_1, x_2, b) \bowtie \left( S(x_2, x_3, x_2) \bowtie T(x_2, x_4) \right)$$
$$Q_2 := \left( R(x_1, x_2, b) \bowtie T(x_2, x_4) \right) \bowtie \left( S(x_2, x_3, x_2) \bowtie T(x_2, x_4) \right)$$

äquivalente azyklische regelbasierte konjunktive Anfragen  $Q'_1$  und  $Q'_2$  und geben Sie Join-Bäume für  $Q'_1$  und  $Q'_2$  an.

- (b) Beweisen oder widerlegen Sie die folgenden Behauptungen:

Für alle Semijoin-Anfragen  $Q_1, Q_2, Q_3$  gilt:

$$(1) \quad \left( (Q_1 \bowtie Q_2) \bowtie Q_3 \right) \equiv \left( Q_1 \bowtie (Q_2 \bowtie Q_3) \right)$$
$$(2) \quad \left( (Q_1 \bowtie Q_2) \bowtie Q_3 \right) \equiv \left( (Q_1 \bowtie Q_3) \bowtie Q_2 \right)$$

### Aufgabe 2:

(25 Punkte)

Beweisen Sie Lemma 3.44 (a), d.h. finden Sie einen Algorithmus, der bei Eingabe einer Semijoin-Anfrage  $Q$  in Zeit  $\mathcal{O}(\|Q\|)$  eine zu  $Q$  äquivalente regelbasierte konjunktive Anfrage  $Q'$  und einen Join-Baum von  $Q'$  berechnet.

### Aufgabe 3:

(2 \* 11 + 8 Punkte)

- (a) Betrachten Sie die beiden regelbasierten konjunktiven Anfragen (wobei  $a, b$  und  $c$  Konstanten sind)

(i)  $Q_1 := \text{Ans}() \leftarrow R(v, w, y), R(a, w, u), P(a, v), R(y, w, x), P(y, w), R(x, y, z)$

(ii)  $Q_2 := \text{Ans}() \leftarrow R(v, w, y), R(a, w, c), P(a, v), R(b, w, x), P(y, w), R(x, y, z)$

Welche davon ist azyklisch, welche nicht? Geben Sie jeweils einen Join-Baum an oder erklären Sie, warum es keinen solchen geben kann. Wandeln Sie die azyklische Anfrage in eine äquivalente Boolesche Semijoin-Anfrage um.

- (b) Geben Sie eine azyklische regelbasierte konjunktive Anfrage  $Q$  an, zu der keine äquivalente Semijoin-Anfrage  $Q'$  existiert. Beweisen Sie, dass Ihre Antwort korrekt ist.

**Aufgabe 4:****(20 Punkte)**

Arbeiten Sie die Details für einen effizienten Algorithmus aus, der das Auswertungsproblem für azyklische regelbasierte konjunktive Anfragen beliebiger Stelligkeit löst und analysieren Sie dessen Laufzeit (gemessen in der Größe  $k$  der gegebenen Anfrage und der Größe  $n$  der gegebenen Datenbank).