Humboldt-Universität zu Berlin
Lehrstuhl Logik in der Informatik
Dr. André Frochaux 27. Juni 2022

Einführung in die Datenbanktheorie

Sommersemester 2022

Übungsblatt 8

Zu Bearbeiten bis zur Übungsstunde am *4. Juli 2022*

Aufgabe 1:

Finden Sie zwei Datalog-Programme P_1 und P_2 mit $\mathit{edb}(P_1) = \mathit{edb}(P_2)$ und $\mathit{idb}(P_1) = \mathit{idb}(P_2)$, so dass für $\mathbf{S} := \mathit{edb}(P_1) = \mathit{edb}(P_2)$ gilt:

(1) Es gibt eine Datenbank $\mathbf{J} \in \mathit{inst}(\mathbf{S})$ so dass $[\![P_1]\!](\mathbf{J}) \not\subseteq [\![P_2]\!](\mathbf{J})$, und

(2) es gibt ein $R \in \mathit{idb}(P_1)$, so dass für die Anfragen $Q_1 := (P_1, R)$ und $Q_2 := (P_2, R)$, sowie alle $\mathbf{I} \in \mathit{inst}(\mathbf{S})$ gilt:
$$[\![Q_1]\!](\mathbf{I}) \subseteq [\![Q_2]\!](\mathbf{I}).$$

Beachten Sie: Aussage (1) bedeutet, dass das *Uniforme Containment-Problem für Datalog-Programme* bei Eingabe von P_1 und P_2 die Ausgabe "nein" liefert; und Aussage (2) bedeutet, dass das *Query Containment Problem für Datalog-Anfragen* bei Eingabe von Q_1 und Q_2 die Ausgabe "ja" liefert.

Aufgabe 2:

Zeigen Sie die im Beweis von Theorem 4.16 a) noch fehlende Rückrichtung, die Folgendes besagt:

$$Q_G \sqsubseteq Q_{G'} \quad \Longrightarrow \quad L(G) \subseteq L(G').$$

Aufgabe 3:

Betrachten Sie das Relationsschema R mit den Attributen A, B, C und die Anfrage $Q :=$

$$\mathit{Ans}(x_1, z_2) \leftarrow R(x_1, y_1, z_1),\ R(x_2, y_1, z_2),\ R(x_1, y_2, z_3).$$

(a) Stellen Sie Q als Tableau-Anfrage (T, t) dar und finden Sie eine *minimale* zu Q äquivalente Tableau-Anfrage.

(b) Betrachten Sie die Menge $\mathcal{F} := \{A \to B,\ B \to C\}$ funktionaler Abhängigkeiten, berechnen Sie $\mathit{chase}(T, t, \mathcal{F})$ und minimieren Sie das Ergebnis.

Aufgabe 4:

Beweisen Sie Theorem 5.13 (a), d.h. zeigen Sie, dass für alle Tableau-Anfragen $Q_1 = (T_1, t_1)$ und $Q_2 = (T_2, t_2)$ über R und jede Menge $\mathcal{F}$ von FDs über R gilt:

$$Q_1 \sqsubseteq_{\mathcal{F}} Q_2 \quad \Longleftrightarrow \quad \mathit{chase}(T_1, t_1, \mathcal{F}) \sqsubseteq \mathit{chase}(T_2, t_2, \mathcal{F}).$$

Aufgabe 5:

Beweisen Sie Lemma 5.14, d.h. zeigen Sie, dass für jede Tableau-Anfrage $Q = (T, t)$ über R und jeder Menge $\mathcal{F}$ von FDs über R gilt:

$$|\min(\mathit{chase}(T,t,\mathcal{F}))| \quad \leq \quad |\min(T,t)|.$$