Goethe-Universität Frankfurt am Main
Institut für Informatik
Theorie komplexer Systeme
Prof. Dr. Nicole Schweikardt

11. Dezember 2012

Logik und Datenbanken

Wintersemester 2012/13

Übungsblatt 7

Zu bearbeiten bis Donnerstag, 20. Dezember 2012

Aufgabe 1: (30 Punkte)

(a) Betrachten Sie das Datalog-Programm P, welches aus folgenden Regeln besteht.

$$\begin{aligned} R(x) &\leftarrow S(x) \\ B(x) &\leftarrow R(y), E(x,y) \\ R(x) &\leftarrow B(y), E(x,y) \\ F(x) &\leftarrow B(x), R(x) \end{aligned}$$

In einer Datenbankinstanz $\mathbf{I}$ vom Schema $edb(P)$ sei ein nicht leerer, ungerichter Graph $G = (V,E)$ modelliert, derart das $\{u,v\} \in V(G)$ genau dann, wenn die Tupel $(u,v),(v,u)$ in $\mathbf{I}(E)$ und $\emptyset \neq \mathbf{I}(S) \subseteq V$.

Was bedeutet es für einen zusammenhängenden Graphen, falls $[\![(P,F)]\!](\mathbf{I}) = \emptyset$?

(b) 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 $\mathbf{R} := \mathit{edb}(P_1) = \mathit{edb}(P_2)$ gilt:

(i) Es gibt eine Datenbank $\mathbf{J} \in \mathit{inst}(\mathbf{R})$ so dass $[\![P_1]\!](\mathbf{J}) \not\subseteq [\![P_2]\!](\mathbf{J})$.
(D.h.: $P_1 \not\subseteq P_2$ im Sinne von “uniformem Containment” von Datalog-Programmen)

(ii) Es gibt ein $R \in \mathit{idb}(P_1)$, so dass für alle $\mathbf{I} \in \mathit{inst}(\mathbf{R})$ gilt: $[\![(P_1,R)]\!](\mathbf{I}) \subseteq [\![(P_2,R)]\!](\mathbf{I})$.
(D.h.: $(P_1,R) \subseteq (P_2,R)$ im Sinne von Query Containment von Datalog-Anfragen.

Aufgabe 2: (30 Punkte)

(a) Zeigen Sie,

(i) für jedes Datalog-Programm P ist der T_P-Operator monoton. (Lemma 3.3)

(ii) jede Datalog-Anfrage ist monoton.

(b) Finden Sie einen Algorithmus, der bei Eingabe einer Dataloganfrage $Q = (P,R)$ entscheidet, ob Q erfüllbar ist.

Aufgabe 3: (20 Punkte)

Zeigen Sie das Lemma (Δ) aus der Vorlesung, d.h. zeigen Sie:

Sei $\Sigma \subseteq \mathbf{dom}$. Sei $G = (\Sigma, V, S, P)$ eine kontextfreie Grammatik, für die gilt:

(i) Es gibt keine Produktion der Form $X \to \epsilon$, für $X \in V$,

(ii) Es gibt keine Produktion auf deren rechter Seite das Startsymbol S steht.

Sei P_G das Datalog-Programm, welches für jede Produktion $A \to B_1 \cdots B_n$ aus G die Regel

$$R_A(x_1, x_{n+1}) \leftarrow \tilde{B}_1, \ldots, \tilde{B}_n \quad \text{mit} \quad \tilde{B}_i := \begin{cases} E(x_i, a, x_{i+1}) & \text{falls } B_i = a \in \Sigma \\ R_X(x_i, x_{i+1}) & \text{falls } B_i = X \in V \end{cases}$$

enthält. Sei $m \geq 1$ und seien $a_1, \ldots, a_m, b_1, \ldots, b_{m-1} \in \mathbf{dom}$. Dann gilt:

$b_1 \cdots b_{m-1} \in L(G) \quad \Leftrightarrow \quad$ Es gibt einen Beweisbaum für das Faktum $R_S(a_1, a_m)$ bzgl P_G, dessen Blätter mit den Fakten $E(a_1, b_1, a_2), E(a_2, b_2, a_3), \ldots, E(a_{m-1}, b_{m-1}, a_m)$ markiert sind.

Aufgabe 4: (20 Punkte)

Zeigen Sie, dass das folgende AUSWERTUNGSPROBLEM FÜR BOOLSCHE DATALOG-ANFRAGEN (kombinierte Komplexität) EXPTIME-vollständig ist.

AUSWERTUNGSPROBLEM FÜR BOOLSCHE DATALOG-ANFRAGEN	
Eingabe:	Datalog-Anfrage $Q = (P, R)$, Datenbank **I**.
Frage:	Ist $[\![Q]\!](\mathbf{I}) \neq \emptyset$?

Hierbei ist:

$$\text{EXPTIME} := \bigcup_{k \in \mathbb{N}} \text{DTIME}(2^{(n^k)}),$$

wobei $\text{DTIME}(2^{(n^k)})$ die Klasse aller Entscheidungsprobleme ist, die von einer deterministischen Turing-Maschine in Zeit $2^{(n^k)}$ gelöst werden können.
Hinweise zur Lösung der Aufgabe finden Sie auf Seite 387 in **[Datalog]**.