

Einführung in die Datenbanktheorie

Wintersemester 2018/2019

Übungsblatt 9

Bearbeitung: in der Übung am 31. Januar 2019

Aufgabe 1:

(25 Punkte)

Zeigen Sie das Lemma \triangle aus der Vorlesung, d.h. zeigen Sie:

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

- (i) Es gibt keine Produktion der Form $X \rightarrow \epsilon$, 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 $A \rightarrow B_1 \cdots B_n$ aus G die Regel

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

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

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

Aufgabe 2:

(25 Punkte)

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

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

Aufgabe 3:**(25 Punkte)**

Zeigen Sie, dass das

BOUNDEDNESS PROBLEM FÜR DATALOG-PROGRAMME*Eingabe:* Datalog-Programm P *Frage:* Ist P beschränkt, d.h. gibt es ein $d \in \mathbb{N}$ so dass $Ab\text{-Stufe}(P, \mathbf{I}) \leq d$ für alle $\mathbf{I} \in inst(edb(P))$?

unentscheidbar ist.

Aufgabe 4:**(25 Punkte)**In der Literatur wird die Semantik von Datalog oft durch den Fixpunkt der iterativen Anwendung des T_P -Operators definiert. Für jedes $\mathbf{J} \in inst(sch(P))$ ist dabei

$$T_P^0(\mathbf{J}) := \mathbf{J} \quad \text{und} \quad T_P^{i+1}(\mathbf{J}) := T_P(T_P^i(\mathbf{J})).$$

(a) Zeigen Sie, dass für jedes Datalog-Programm P , alle $i \in \mathbb{N}$ und jedes $\mathbf{I} \in inst(edb(P))$ gilt:

$$T_P^i(\hat{\mathbf{I}}) = S_P^i(\hat{\mathbf{I}}).$$

(b) Gilt sogar für alle $\mathbf{J} \in inst(sch(P))$, dass

$$T_P^i(\mathbf{J}) = S_P^i(\mathbf{J}) ?$$

Beweisen Sie, dass Ihre Antwort korrekt ist.