Universität Freiburg
Institut für Informatik
Prof. Dr. G. Lausen
Michael Schmidt

Georges-Köhler Allee, Geb. 51
D-79110 Freiburg
Tel. (0761) 203-8120
Tel. (0761) 203-8127

Formale Grundlagen von Informationssystemen
Sommersemester 2009
05.05.2009

3. Übungsblatt: Chase und Terminierung

Übung 9 (Chase Anwendung, 1+1+3=5 Punkte)
Betrachten Sie das Schema aus Übung 8 (Übungsblatt 2) und die Constraintmenge $\Sigma := \{\alpha_1, \alpha_2, \alpha_3\}$ mit

$$\alpha_1 := \forall\, c_1, c_2, c_3, d_1, d_2\ (\texttt{rail}(c_1, c_2, d_1), \texttt{rail}(c_2, c_3, d_2) \rightarrow \exists\, d_3\ \texttt{rail}(c_1, c_3, d_3))$$
$$\alpha_2 := \forall\, c_1, c_2, d_1, d_2\ (\texttt{fly}(c_1, c_2, d_1) \wedge \texttt{fly}(c_2, c_1, d_2) \rightarrow d_1 = d_2)$$
$$\alpha_3 := \forall\, c_1, c_2, d_1\ (\texttt{fly}(c_1, c_2, d_1) \rightarrow \exists\, d_2\ \texttt{fly}(c_2, c_1, d_2))$$

und die Konjunktive Anfrage

$$Q: \quad \texttt{ans}(C_3) \leftarrow \texttt{rail}(\mathit{Freiburg}, C_1, D_1), \texttt{rail}(C_1, C_2, D_2), \texttt{fly}(C_2, C_3, D_3).$$

a) Beschreiben Sie die Constraints und die Anfrage in Worten.
b) Welche Constraints aus Σ erfüllt $\mathit{body}(Q)$? Erfüllt $\mathit{body}(Q)$ die gesamte Constraintmenge Σ?
c) Chasen Sie Q mit Σ. Geben Sie alle Zwischenergebnisse (= Chase-Schritte) an. Gilt $\mathit{body}(Q^\Sigma) \models \Sigma$?

Übung 10 (Chase und Minimierung, 1+2+3=6 Punkte)
Betrachten Sie das Datenbankschema mit den Relationen

$\texttt{Person}(\mathit{SSN}, \mathit{Name})$
$\texttt{Professor}(\mathit{SSN}, \mathit{Name})$
$\texttt{Course}(\mathit{CourseName}, \mathit{SSN})$
$\texttt{Enrolled}(\mathit{CourseName}, \mathit{Participant})$

wobei `Person` Personen mit Sozialversicherungsnummer (SSN) und Name, `Professor` Professoren mit Sozialversicherungsnummer und Name, `Course` Namen angebotener Vorlesungen und die SSN des Referenten, sowie `Enrolled` Teilnehmer für die entsprechended Kurse enthalten. Weiterhin sei die Constraintmenge $\Sigma := \{\beta_1, \beta_2, \beta_3\}$ mit

$$\beta_1 := \forall\, s, n\ (\texttt{Professor}(s,n) \rightarrow \texttt{Person}(s,n))$$
$$\beta_2 := \forall\, c, s, n\ (\texttt{Course}(c,s) \wedge \texttt{Person}(s,n) \rightarrow \texttt{Professor}(s,n))$$
$$\beta_3 := \forall\, c, s\ (\texttt{Course}(c,s) \rightarrow \exists\, p\ \texttt{Enrolled}(c,p))$$

gegeben. Betrachten Sie nun die Konjunktive Anfrage

Q: $\texttt{ans}(C,N) \leftarrow \texttt{Professor}(S,N), \texttt{Course}(C,S)$

a) Beschreiben Sie die Constraints in Worten.

b) Berechnen Sie Q^{Σ}.

c) Berechnen Sie – ausgehend von Q^{Σ} – die Menge aller minimalen Σ-äquivalenten Anfragen.

Übung 11 (Chase Terminierung, 1 Punkt)
Gegeben sei das Datenbankschema $\texttt{E}(src,dest)$, das die Kantenrelation eines Graphen speichert, und die Konjunktive Anfrage Q: $\texttt{ans}(X) \leftarrow \texttt{E}(X,Y)$. Geben Sie eine Tuple-generating Dependency α an, so dass der Chase für Q mit $\Sigma := \{\alpha\}$ nicht terminiert.

Übung 12 (Chase Terminierung, 1+1+1=3 Punkte)
Entscheiden Sie jeweils für die folgenden Constraintmengen Σ_1, Σ_2 und Σ_3, ob mit Hilfe der *Acyclicity*-Bedingung aus der Vorlesung die Terminierung des Chase Algorithmus für die jeweilige Menge garantiert werden kann. Zeichnen Sie die entsprechenden Relationsgraphen.

$\Sigma_1 := \{\ \forall\, c_1, c_2, c_3, d_1, d_2\ (\texttt{rail}(c_1, c_2, d_1), \texttt{rail}(c_2, c_3, d_2) \rightarrow \exists\, d_3\ \texttt{rail}(c_1, c_3, d_3))\ \}$

$\Sigma_2 := \{\ \forall\, c_1, c_2, d_1, d_2\ (\texttt{fly}(c_1, c_2, d_1) \wedge \texttt{fly}(c_2, c_1, d_2) \rightarrow d_1 = d_2),$
$\forall\, c_1, c_2\ (\texttt{hasAirport}(c_1) \wedge \texttt{hasAirport}(c_2) \rightarrow \exists\, d\ \texttt{fly}(c_1,c_2,d)),$
$\forall\, c_1, c_2, d_1\ (\texttt{fly}(c_1,c_2,d_1) \rightarrow \exists\, d_2\ \texttt{rail}(c_1,c_2,d_2))\ \}$

$\Sigma_3 := \Sigma_2 \cup \{\ \forall\, x_1, x_2, x_3, d_1, d_2\ (\texttt{rail}(x_1,x_2,d_1) \wedge \texttt{fly}(x_2,x_3,d_2) \rightarrow \texttt{hasAirport}(x_2) \wedge \texttt{hasAirport}(x_3)\ \}$

Übung 13 (Chase Terminierung, Bonusaufgabe, 3+2=5 Punkte)
Eine Verbesserung der in der Vorlesung behandelten hinreichenden Chase Terminierungsbedingung *Acyclicity* ist die sogenannte *Weak Acyclicity*. Statt auf kompletten Relationen ist Weak Acyclicity über Positionen in den Relationen definiert. Zum Beispiel hat die Relation $\texttt{fly}(c\_id1,c\_id2,dist)$ drei Positionen, $\texttt{fly}^1$ (Attribut *c_id1*), $\texttt{fly}^2$ (Attribut *c_id2*), $\texttt{fly}^3$ (Attribut *dist*). Basierend auf diesem Begriff ist *Weak Acyclicity* wie folgt definiert:

Definition 1 Sei Σ eine Menge von Tuple- und Equality-generating Dependecies. Der Abhängigkeitsgraph von Σ, $dep(\Sigma) := (V,E)$ ist ein gerichteter Graph der wie folgt definiert ist. V ist gegeben durch die Menge von Positionen die in Tuple-generating Dependencies von Σ vorkommen. Die Kantenrelation E enthält zwei verschiedene Kantentypen und wird wie folgt aufgebaut: Für jede Tuple-generating Dependency

$$\forall\, \overline{x}(\phi(\overline{x}) \rightarrow \exists\, \overline{y}\psi(\overline{x}, \overline{y})) \in \Sigma$$

und jedes x aus $\overline{x}$ das in ψ vorkommt und jedes Vorkommen von x in ϕ in Position π_1

- für jedes Vorkommen von x in ψ in Position π_2, füge eine Kante $\pi_1 \rightarrow \pi_2$ hinzu (wenn sie nicht bereits existiert).
- für jede existentiell quantifizierte Variable y und jedes Vorkommen von y in einer Position π_2, füge eine Spezialkante $\pi_1 \xrightarrow{*} \pi_2$ hinzu (wenn sie nicht bereits existiert).

Eine Constraintmenge Σ heißt *Weakly Acyclic* genau dann wenn ihr Abhängigkeitsgraph $dep(\Sigma)$ keinen Zyklus durch eine Spezialkante besitzt.

Wie auch *Acyclicity* garantiert *Weak Acyclicity* Chase-Terminierung für jede mögliche Datenbankinstanz.

a) Zeichnen Sie die Abhängigkeitsgraphen für die Constraintmengen Σ_1, Σ_2 und Σ_3 aus Aufgabe 12. Sind diese Constraintmengen *Weakly Acyclic*? Können Terminierungsgarantien abgeleitet werden?

b) Geben Sie eine einelementige Constraintmenge über einer Kantenrelation $\texttt{E}(src,dest)$ an, die nicht *Weakly Acyclic* ist.

Abgabe 12.05.2009

Literatur: siehe Vorlesungsfolien