9. Übung zur Vorlesung "Objektorientierte und Deduktive Datenbanken" im Wintersemester 2004 – mit Musterlösungen –

Prof. Dr. Gerd Stumme, Dipl.-Inform. Christoph Schmitz

10. Januar 2005

Aufgabe 1

Formulieren Sie die Anfrage "Welche Filme haben Allen als Schauspieler oder Regisseur?" in nr-Datalog, SPCU, SPJRU!

Das Schema sei *Movies(Title, Director, Actor)*.

nr-Datalog

$S(x) \leftarrow Movies(x, allen, z).$

$S(x) \leftarrow Movies(x, y, allen).$

SPCU

$\pi_1(\sigma_{2=allen}(Movies) \cup \sigma_{3=allen}(Movies))$

SPJRU

$\pi_{Title}(\sigma_{Director=allen}(Movies) \cup \sigma_{Actor=allen}(Movies))$

Aufgabe 2

Zeigen Sie, daß SPCU-Algebra und nr-Datalog äquivalent sind.

Tip: Geben Sie eine Konstruktion an, um aus einem Algebra-Ausdruck ein gleichwertiges nr-Datalog-Programm zu machen und umgekehrt!

SPCU nach nr-Datalog

Wir konstruieren Datalog-Regeln induktiv über den Aufbau von Ausdrücken F in SPCU.

Induktionsanfang: $F = R$

Dieser Formel entspricht eine Datalog-Regel der Form

$S(\dots) \leftarrow R(\dots)$

Induktionsschritt

Seien für die Ausdrücke α, β schon Datalog-Regeln der Form $S_\alpha(\dots) \leftarrow \dots$ und $S_\beta(\dots) \leftarrow \dots$ gefunden.

$F = \sigma_{k=j}(\alpha)$

Dieser Formel enspricht die Regel

$S_{\sigma\alpha}(\dots) \leftarrow S_\alpha(\dots, \underset{k}{A}, \dots, \underset{j}{A}, \dots)$

$F = \sigma_{k=a}(\alpha)$

$S_{\sigma\alpha'}(\dots) \leftarrow S_\alpha(\dots, \underset{k}{a}, \dots)$

$F = \pi_{i_1,\dots,i_k}(\alpha)$

$S_{\pi\alpha}(A_{i_1}, \dots, A_{i_k}) \leftarrow S_\alpha(A_1, \dots, A_n)$

$F = \alpha \times \beta$

$S_{\alpha\beta}(A_1, \dots, A_k, B_1, \dots, B_l) \leftarrow S_\alpha(A_1, \dots, A_k), S_\beta(B_1, \dots, B_l)$

$F = \alpha \cup \beta$

$S_{\alpha\beta'}(\dots) \leftarrow S_\alpha(\dots)$

$S_{\alpha\beta'}(\dots) \leftarrow S_\beta(\dots)$

nr-Datalog nach SPCU

Gegeben sei das nr-Datalog-Programm

$$\begin{array}{rcl} S_1 & \leftarrow & body_1 \\ S_2 & \leftarrow & body_2 \\ & \vdots & \\ S_n & \leftarrow & body_n \end{array}$$

Wir konstruieren einen SPCU-Ausdruck α_j für jedes S_j und nehmen dabei an, daß jeweils schon Ausdrücke α_i für die S_i mit $i < j$ gefunden seien – das funktioniert per Induktion, da keine Rekursion zulässig ist und somit kein S_i mit $i \geq j$ vorkommen kann.

$body_j$ habe die Form $R_1(\dots), \dots, R_k(\dots), S_{i_1}(\dots), \dots, S_{i_l}$.

Wir konstruieren daraus einen Algebraausdruck

$\pi_\theta(\sigma_\phi(I(R_1) \times \dots \times I(R_k) \times \alpha_{i_1} \times \dots \times \alpha_{i_l}))$

wobei θ die Positionen der in S_j vorkommenden Variablen sind, und ϕ einen Ausdruck bezeichnet, der entsprechend der Konstanten und mehrfach vorkommenden Variablen in $body_j$ Tupel selektiert.

Für gleiche S_j bis S_l bilden wir die die Vereinigung $\alpha_{jl} := \alpha_j \cup \cdots \cup \alpha_l$, die überall dort eingesetzt wird, wo $S_j = \cdots = S_l$ vorkommt.

Schließlich müssen wir noch den Fall berücksichtigen, daß im Regelkopf Konstanten oder mehrfach vorkommende Variablen auftreten, also

$S(x,x) \leftarrow R(x,y)$ bzw. $S(x,a) \leftarrow R(x,y)$

Den ersten Fall lösen wir durch ein weiteres Kreuzprodukt und entsprechende Selektion; folgender Ausdruck entspricht $S(x,x)$:

$\pi_{1,3}(\sigma_{2=4}(I(R) \times I(R)))$

Im zweiten Fall benötigen wir ein Kreuzprodukt mit der Menge, die nur das Tupel $\langle a \rangle$ enthält: $S(x,a)$ entspricht

$\pi_1(I(R)) \times \{\langle a \rangle\}$

Aufgabe 3

Demonstrieren Sie, daß keine der Operationen $\sigma, \pi, \times, \cup$ der SPCU-Algebra überflüssig ist. Nennen Sie dazu zu jeder Operation ein Schema und eine Anfrage, die zeigt, daß diese Operation sich nicht aus den anderen nachbilden läßt.

Wir betrachten das Schema *Movies(Title, Director, Actor)*.

Selektion

Betrachte z. B. $\sigma_{1=anniehall}(Movies)$ auf der Instanz

$I(Movies) = \{\langle anniehall, allen, allen \rangle, \langle deconstructingharry, allen, allen \rangle\}$

Mit keiner Kombination der anderen Operationen gelingt es, genau das Tupel $\langle anniehall, allen, allen \rangle$ auszuwählen, da $\times, \cup, \pi$ nicht die Anzahl der vorkommenden Konstanten in einem Attribut der Relation verringern können.

Projektion

$\pi_1(Movies)$ hat einelementige Tupel als Resultat. Jeder Ausdruck auf $Movies$ aus den Operationen $\sigma, \times, \cup$ hat aber Tupel als Resultat, die eine Stelligkeit von $3k, k \in \mathbb{N}$ haben. Demnach kann π nicht nachgeildet werden.

Kreuzprodukt

$\times$ ist die einzige der Operationen, die Tupel größerer Stelligkeit als das Argument zum Ergebnis hat. Daher kann sie nicht durch die anderen Operationen ersetzt werden.

Vereinigung

Sei *Documentaries(Title, Director, Actor)* ein zweites Schema. Der Ausdruck $\pi_1(Movies \cup Documentaries)$ enthält dann z. B. die Titel aller Filme und Dokumentationen. Da keine der anderen Operationen Tupel erzeugen kann, deren Werte für ein bestimmtes Attribut aus Instanzen verschiedener Relationen stammen, kann diese Operation nicht nachgebildet werden.