Universität Freiburg
Institut für Informatik
Prof. Dr. G. Lausen
Alexander Schätzle
Martin Przyjaciel-Zablocki

Georges-Köhler Allee, Geb. 51
D-79110 Freiburg
lausen@informatik.uni-freiburg.de
schaetzle@informatik.uni-freiburg.de
zablocki@informatik.uni-freiburg.de

Data Models and Query Languages
Summerterm 2013

3. Exercise Sheet: Conjunctive Queries, Chase & Datalog

Discussion: 17.05.2013

Submission Guidlines: This is a mandatory exercise sheet where you have to get 50% of the points to qualify for the exam! Hand in your solutions at the beginning of the tutorial on 17.05.2013.

Exercise 1 (Evaluation of Conjunctive Queries, 1+1+1+1=4 points)
Consider the following sample instantiation $\mathcal{I}$ of the schema from Exercise Sheet 1.

Sales	*PName*	*SName*	*CName*
	Audi A7	Autohaus Wenz	Meier
	Audi A8	Autohaus Klein	Meier
	Audi A8	Autohaus Wenz	Smith
	Suzuki GSX	Motorsport AG	Hofmann

Part	*PName*	*Type*
	Audi A8	Auto
	Audi A7	Auto
	Suzuki GSX	Motorrad

Cust	*CName*	*CAddr*
	Meier	Freiburg
	Smith	Freiburg
	Hofmann	Mannheim

Supp	*SName*	*SAddr*
	Autohaus Wenz	Freiburg
	Autohaus Klein	Mannheim
	Motorsport AG	Mannheim

Give the evaluation result of the following queries on instance $\mathcal{I}$ and informally describe their meaning. Note that constants inside the queries are distinguished by *italic* font.

a) q_1: ans(C) ← Sales(P,S,C), Cust(C,*Freiburg*), Supp(S,*Freiburg*)

b) q_2: ans(S,P) ← Sales(P,S,*Meier*), Supp(S,*Mannheim*), Part(P,*Auto*)

c) q_3: ans(S,P) ← Sales(P,S,*Meier*), Supp(S,*Mannheim*), Part(P2,*Auto*)

d) q_4: ans(C1,C2) ← Cust(C1,*Freiburg*), Cust(C2,*Freiburg*), Sales(P1,S1,C1), Sales(P2,S2,C2), Supp(S1,X), Supp(S2,X)

Exercise 2 (Containment, 3+2=5 points)
Let $E(src, dest)$ denote the edge relation of a directed graph and consider the following conjunctive queries.

- $Q_1 : ans(X, Y) \leftarrow E(X, Y), E(Y, Z)$
- $Q_2 : ans(X, Y) \leftarrow E(X, W), E(W, Y)$
- $Q_3 : ans(X, Y) \leftarrow E(X, Y), E(X, U), E(U, Y)$

a) Check if $Q_i \sqsubseteq Q_j$ forall $i \neq j$, $1 \leq i, j \leq 3$. Whenever containment does not hold for a pair of queries, provide a sample instance that proves violation.
b) Show that $\{Q_1, Q_2\} \equiv \{Q_1, Q_2, Q_3\}$ holds.

Exercise 3 (Containment in Graph Cycles, 3 points)
Consider the infinite sequence of Conjunctive Queries $Q_1, Q_2, \ldots$, where

$$Q_i : ans(X) \leftarrow E(X, Y_1), E(Y_1, Y_2), \ldots, E(Y_{i-1}, Y_i), E(Y_i, X)$$

Q_i represents a cycle of length $i + 1$ in a directed graph with edge relation $E(src, dest)$.
Which containment relationships exist between the Q_i?

Exercise 4 (Chase Termination, 3 points)
Let $\texttt{E}(src,dest)$ store the edge relation of a graph and let Q: $\texttt{ans}(X) \leftarrow \texttt{E}(X,Y)$. Find a tuple-generating dependency α such that the chase of Q with $\Sigma := \{\alpha\}$ does not terminate.

Exercise 5 (Chase and Minimization, 1+2+1=4 points)
Consider the following database schema with relations

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

where `Person` stores persons including social security number (SSN) and name, `Professor` stores professors including social security number and name, `Course` contains course names and the SSN of the lecturer, and `Enrolled` stores course inscriptions. Further let $\Sigma := \{\alpha_1, \alpha_2, \alpha_3\}$ be the set of the following constraints.

$alpha_1 := \forall\, s, n\ (\texttt{Professor}(s,n) \rightarrow \texttt{Person}(s,n))$
$\alpha_2 := \forall\, c, s, n\ (\texttt{Course}(c,s) \wedge \texttt{Person}(s,n) \rightarrow \texttt{Professor}(s,n))$
$\alpha_3 := \forall\, c, s\ (\texttt{Course}(c,s) \rightarrow \exists\, p\ \texttt{Enrolled}(c,p))$

Further consider the Conjunctive Query

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

a) Describe the constraints informally.
b) Compute Q^{Σ}.
c) Compute – starting from Q^{Σ} – the set of all minimal Σ-equivalent queries.

Exercise 6 (Datalog, Graph Properties, 1+1+1+1=4 points)
Consider a directed graph which is given by $E(src, dest)$. Give a Datalog program which computes the following relations:

a) $Odd(X, Y)$, which holds if there is a path with odd length from X to Y.

b) $Oddcycle(X)$, there is a cycle with odd length through X.

c) $Evencycle(X)$, there is cycle with even length through X.

d) $Bothcycles(X)$, there are cycles with even length and cycles with odd length through X.

Exercise 7 (Datalog, Family Tree, 3 points)
Let $parent(X, Y)$ be a family tree with root p. Give a Datalog program which computes the predicates $same - generation(X, Y)$, $sibling(X, Y)$ and $cousin(X, Y)$.

- $parent(X, Y)$ holds if Y is a parent of X.
- $same - generation(X, Y)$ holds, if distance between X and p is the same as distance between Y and p.
- $sibling(X, Y)$ holds, if X and Y have the same parents.
- $cousin(X, Y)$ holds, if X and Y belong to the same generation but are not siblings.

Exercise 8 (Datalog, Stable Models, 2+2=4 points)
Consider the following Datalog program:

$$\sqcap : \quad win(X) \leftarrow move(X, Y), \neg win(Y)$$

with EDBs:

- $\{move(1, 2), move(2, 3), move(3, 1), move(3, 4)\}$
- $\{move(1, 2), move(2, 3), move(3, 1), move(3, 4), move(4, 5)\}$

a) Give all stable models.

b) Give the well-founded models.