3. Exercise Sheet: Conjunctive Queries, Chase & Datalog

Discussion: 17.05.2013

Submission Guidelines: 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 $I$ of the schema from Exercise Sheet 1.

<table>
<thead>
<tr>
<th>Sales</th>
<th>PName</th>
<th>SName</th>
<th>CName</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audi A7</td>
<td>Autohaus Wenz</td>
<td>Meier</td>
<td></td>
</tr>
<tr>
<td>Audi A8</td>
<td>Autohaus Klein</td>
<td>Meier</td>
<td></td>
</tr>
<tr>
<td>Audi A8</td>
<td>Autohaus Wenz</td>
<td>Smith</td>
<td></td>
</tr>
<tr>
<td>Suzuki GSX</td>
<td>Motorsport AG</td>
<td>Hofmann</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part</th>
<th>PName</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audi A8</td>
<td>Auto</td>
<td></td>
</tr>
<tr>
<td>Audi A7</td>
<td>Auto</td>
<td></td>
</tr>
<tr>
<td>Suzuki GSX</td>
<td>Motorrad</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cust</th>
<th>CName</th>
<th>CAddr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meier</td>
<td>Freiburg</td>
<td></td>
</tr>
<tr>
<td>Smith</td>
<td>Freiburg</td>
<td></td>
</tr>
<tr>
<td>Hofmann</td>
<td>Mannheim</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supp</th>
<th>SName</th>
<th>SAddr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autohaus Wenz</td>
<td>Freiburg</td>
<td></td>
</tr>
<tr>
<td>Autohaus Klein</td>
<td>Mannheim</td>
<td></td>
</tr>
<tr>
<td>Motorsport AG</td>
<td>Mannheim</td>
<td></td>
</tr>
</tbody>
</table>

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

a) $q_1: \text{ans}(C) \leftarrow \text{Sales}(P,S,C), \text{Cust}(C,\text{Freiburg}), \text{Supp}(S,\text{Freiburg})$

b) $q_2: \text{ans}(S,P) \leftarrow \text{Sales}(P,S,\text{Meier}), \text{Supp}(S,\text{Mannheim}), \text{Part}(P,\text{Auto})$

c) $q_3: \text{ans}(S,P) \leftarrow \text{Sales}(P,S,\text{Meier}), \text{Supp}(S,\text{Mannheim}), \text{Part}(P,\text{P2, Auto})$

d) $q_4: \text{ans}(C1,C2) \leftarrow \text{Cust}(C1,\text{Freiburg}), \text{Cust}(C2,\text{Freiburg}), \text{Sales}(P1,S1,C1), \text{Sales}(P2,S2,C2), \text{Supp}(S1,X), \text{Supp}(S2,X)$
**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 $Q_i \sqsubseteq Q_j$ for all $i \neq j, 1 \leq i, j \leq 3$. 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 $E(src, dest)$ store the edge relation of a graph and let $Q: ans(X) \leftarrow E(X, Y)$. Find a tuple-generating dependency $\alpha$ 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

- Person(SSN, Name)
- Professor(SSN, Name)
- Course(CourseName, SSN)
- 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.

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

Further consider the Conjunctive Query

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

a) Describe the constraints informally.

b) Compute $Q^\Sigma$.

c) Compute – starting from $Q^\Sigma$ – the set of all minimal $\Sigma$-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:

\[ \Box : \text{win}(X) \leftarrow \text{move}(X, Y), \neg \text{win}(Y) \]

with EDBs:

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

a) Give all stable models.
b) Give the well-founded models.