Exercise Sheet 2: First-Order-Queries

David Carral, Markus Krötzsch

Database Theory, 24 April 2018, Summer Term 2018

Exercise 2.1. Express the queries from Exercise 1.1 as domain independent FO-queries.

Exercise 2.2. Let R be a table (relational instance) with attributes A and B. Use the construction from the lecture to express the following $\mathsf{RA}_{\text{named}}$ query as a $\mathsf{DI}_{\text{unnamed}}$ query:

$$q[A, B] := \big(\pi_A(R) \bowtie \pi_B(R)\big) - \big(R \bowtie (\delta_{B,A \to A,B}(R))\big)$$

Exercise 2.3. It was stated in the lecture that query mappings under named perspective can be translated into query mappings under unnamed perspective. Specify this translation.

Exercise 2.4. Complete the proof that $\mathsf{RA}_{\text{named}} \sqsubseteq \mathsf{DI}_{\text{unnamed}}$ from the lecture by showing that the results of the transformation are (a) domain independent and (b) equivalent to the input query. In each case, show that the claimed property holds true for each case of the recursive construction under the assumption (induction hypothesis) that it has already been established for all subqueries. Use the mappings from Exercise 2.3 to compare named and unnamed results.

Exercise 2.5. Consider a binary predicate R and the following $\mathsf{AD}_{\text{unnamed}}$ query:

$$\varphi[x, y] = \neg(R(x, y) \wedge R(y, x))$$

Use the construction from the lecture to express it as an $\mathsf{RA}_{\text{named}}$ query.

Exercise 2.6. Complete the constructions for the proof of $\mathsf{AD} \sqsubseteq \mathsf{RA}$ given in the lecture.

1. Define the relational algebra expression $E_{a,\mathbf{adom}}$, such that $E_{a,\mathbf{adom}}(\mathcal{I}) = \{\{a \mapsto c\} \mid c \in \mathbf{adom}(\mathcal{I}, q)\}$ (assume that the query and the database schema are known).
2. Define the expressions E_φ for $\varphi = \varphi_1 \vee \varphi_2$ and $\varphi = \forall y.\psi$ in terms of expressions that have already been defined in the lecture.
3. Give a direct definition for the expression E_φ for $\varphi = \varphi_1 \vee \varphi_2$.

Exercise 2.7. Use the function rr from the lecture to compute the set of range-restricted variables for the following FO queries:

1. $\exists y_{\text{SID}}, y_{\text{Stop}}, y_{\text{To}}.\big(\mathsf{Stops}(y_{\text{SID}}, y_{\text{Stop}}, \texttt{"true"}) \wedge \mathsf{Connect}(y_{\text{SID}}, y_{\text{To}}, x_{\text{Line}})\big)[x_{\text{Line}}]$
2. $\neg\mathsf{Lines}(x, \texttt{"bus"})[x]$
3. $\big(\mathsf{Connect}(x_1, \texttt{"42"}, \texttt{"85"}) \vee \mathsf{Connect}(\texttt{"57"}, x_2, \texttt{"85"})\big)[x_1, x_2]$
4. $\forall y.p(x, y)[x]$
5. $\exists x.(((p(x) \to q(c)) \to p(x)) \to p(x))$

Which of these queries is a safe-range query? Which of the queries is domain independent?