

5. Query Plans and Rewritings

Instructions: Work in groups of 2 students. You can write up your answers by hand (provided your handwriting is legible) or use a word processing system like Latex or Word. However, experience shows that Word is in general difficult to use for this kind of task. Please, include name and email address in your submission. Please, indicate also the sources you have used for your work. If you make significant use of other work without providing a reference, this will be considered an attempt at plagiarism.

Consider a schema \mathcal{S} and let Q be a query over \mathcal{S} . Let $\mathcal{V} = \{V_1, \dots, V_n\}$ be a set of relation symbols, disjoint from \mathcal{S} , and assume that each of the V_i is associated with a relational conjunctive query $V_i(\bar{s}_i) :- L_i$, where L_i is a conjunction of relational atoms. We say that \mathcal{V} is a set of *views*.

A query R over the relation symbols in \mathcal{V} is a (contained) *rewriting of Q using the views \mathcal{V}* if $R \circ \mathcal{V} \subseteq Q$, or in a different notation, if $R \subseteq_{\mathcal{V}} Q$ (see the lecture slides for definitions). We will assume \mathcal{V} to be fixed and talk in the sequel simply about “rewritings of Q ” or just “rewritings.”

Note that containment defines a preorder on any set of queries, e.g., on the rewritings of Q . (A preorder is a relation that is reflexive and transitive.) Let \mathcal{R} be a set of queries. We say that a query $R \in \mathcal{R}$ is *maximal* in \mathcal{R} if it is maximal wrt. “ \subseteq ”, that is, if $R \subseteq R'$ entails $R' \subseteq R$ for any $R' \in \mathcal{R}$.

1. Maximal Rewritings are Small and Simple—in the Relational Case

Let $\text{REW}_{\mathcal{V}}^{\text{rel}}(Q)$ be the set of relational conjunctive queries that are rewritings of Q using \mathcal{V} and let $R \in \text{REW}_{\mathcal{V}}^{\text{rel}}(Q)$ be maximal.

Show the following two claims from the lecture:

1. R has at most as many atoms as Q ;
2. R contains only constants occurring in Q or \mathcal{V} .

(12 Points)

2. Rewriting Conjunctive Queries Using Views with Comparisons

For any query Q , defined as $Q(\bar{s}) :- L, M$, let \bar{Q} be defined as $\bar{Q}(\bar{s}) :- L$, that is, \bar{Q} is obtained from Q by dropping the comparisons.

Consider a set of views \mathcal{W} defined by conjunctive queries possibly containing comparisons. Let $\bar{\mathcal{W}}$ the collection of all views obtained from \mathcal{W} by dropping comparisons.

Let Q be a relational conjunctive query. Show the following claims:

1. For any maximal rewriting $R \in \text{REW}_{\mathcal{W}}^{rel}(Q)$, we have $R \equiv \bar{R}$.
2. $\text{REW}_{\mathcal{W}}^{rel}(Q) = \text{REW}_{\bar{\mathcal{W}}}^{rel}(Q)$.
3. Consider a LAV information integration system with sound mappings, containing a set of views \mathcal{W} , possibly with comparisons. Let Q be a relational conjunctive query. Then $\text{REW}_{\bar{\mathcal{W}}}^{rel}(Q)$ is a complete set of plans.

In other words, maximal rewritings need not have comparisons if the query does not have comparisons. Moreover, one can safely ignore the comparisons in the views when computing rewritings for queries without comparisons. Finally, one can compute all certain answers with rewritings (= plans), provided the query does not contain comparisons.

(18 Points)

Submission: 21 June, 2011, 10:30 am, at the exam.