

5. Containment of Positive Queries

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. Note that experience has shown that Word is in general difficult to use for this kind of task. If you prefer to write up your solution by hand, submit a scanned electronic version. Please, include name and email address in your submission.

1. Containment of Unions of Conjunctive Queries

We consider again queries without built-in atoms, which we called relational or simple queries.

Let Q_1, \dots, Q_n be conjunctive queries, each defined as $Q_i(\bar{x}) :- L_i$. Then

$$Q := \bigcup_{i=1}^n Q_i$$

defines a new query, the union of the Q_i . Over an instance \mathbf{I} , the query Q returns the result $Q(\mathbf{I}) = \bigcup_{i=1}^n Q_i(\mathbf{I})$. We call such a query a *union of conjunctive queries*. Note that if Q is the union of the conjunctive queries Q_i , then all the Q_i have the same arity, indicated by using the same vector of distinguished variables \bar{x} for all Q_i . Clearly, the arity of the Q_i is also the arity of Q .

Find out how to decide the following variants of the containment problem:

1. “UCQ in CQ”: Given a union of conjunctive queries $Q = \bigcup_{i=1}^n Q_i$ and a conjunctive query Q' , is Q contained in Q' ?
2. “CQ in UCQ”: Given a conjunctive query Q and a union of conjunctive queries $Q' = \bigcup_{j=1}^m Q'_j$, is Q contained in Q' ?
3. “UCQ in UCQ”: Given a union of conjunctive queries $Q = \bigcup_{i=1}^n Q_i$ and union of conjunctive queries $Q' = \bigcup_{j=1}^m Q'_j$, is Q contained in Q' ?

For each case, give a decidable criterion for containment and show that your criterion is correct. Also, assess the complexity of each problem.

Hint: When proving your criterion for Case 2, you may want to take the proof of the Homomorphism Theorem as a starting point for the proof of your new criterion.

(18 Points)

2. Containment of Positive Queries

In previous coursework we have shown how to decide in polynomial time whether a relational (or “simple”) positive query is safe.

- Can one decide containment for safe relational positive queries? If yes, how? If no, why not?

The general definition of containment ($Q \sqsubseteq Q'$ iff $Q(\mathbf{I}) \subseteq Q'(\mathbf{I})$ for all instances \mathbf{I}) applies also to general positive queries that may be unsafe.

- How can one generalize the answer to the question about safe queries to arbitrary positive queries?

(12 Points)

Submission: 12 June, 11:30 pm, by email