5. Containment of Conjunctive Queries (2)

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.

1. Conjunctive Queries with Equalities

In this exercise we consider conjunctive queries with equalities (but without disequalities).

Give a characterization of containment between two such queries and find out the complexity of the decision problem.

Hint: Solving the exercise may involve

- slightly generalising the class of queries you want to consider;
- slightly generalizing the concept of a homomorphism;
- defining a normal form for queries and/or tableaux.

(10 Points)

2. Unions of Conjunctive Queries

A *union of conjunctive queries* is a sequence $(q_1, \ldots, q_n)$ of simple conjunctive queries q_i such that all q_i have the same arity. We write unions of conjunctive queries in the form

$$Q = \bigcup_{i=1}^{n} q_i(\vec{x}).$$

A union of conjunctive queries can itself be interpreted as a query by defining

$$Q(\mathbf{I}) = \bigcup_{i=1}^{n} q_i(\mathbf{I}).$$

Develop a characterisation of containment for unions of conjunctive queries and determine the complexity of the decision problem.

Hint: This exercise is pretty easy if you approach it in the right way. Solutions can be quite short.

(8 Points)

3. Reducing the Hamiltonian Path Problem to Containment

Let $G = (V, E)$ be an undirected graph, where V is a finite set, the elements of which are called *vertices,* and $E \subseteq \mathcal{P}_2(V)$ is a collection of two-element subsets of V, the elements of which called the *edges* of G.
A path in G is a sequence $v_1, \ldots, v_n$ of vertices such that $\{ v_i, v_{i+1} \} \in E$ for $i = 1, \ldots, n-1$ (that is, each node is connected to the next by an edge). A path $v_1, \ldots, v_n$ is *Hamiltonian* if in addition we have

1. $v_i \neq v_j$ if $i \neq j$ (that is, all vertices on the path are distinct)
2. $\{ v_1, \ldots, v_n \} = V$ (that is, $v_1, \ldots, v_n$ enumerates all vertices of G).

The Hamiltonian Path Problem is defined as follows:

Given: An undirected graph $G = (V, E)$.

Question: Does there exist a Hamiltonian Path in G?

This problem is known to be NP-complete.

Show that containment of simple conjunctive queries is NP-hard by reducing the Hamiltonian Path Problem to Query Containment.

Proceed in the following three steps:

1. Construct, for each graph G, a pair of queries Q, Q' such that $Q \subseteq Q'$ if and only if there is a Hamiltonian path in G.

   (4 Points)

2. Prove that for a graph G and the corresponding pair of queries Q, Q', the containment $Q \subseteq Q'$ implies the existence of a Hamiltonian path.

   (4 Points)

3. Prove that from the existence of a Hamiltonian path in G one can conclude the containment of $Q \subseteq Q'$.

   (4 Points)

Hints: You may want to encode edges using a binary relation `edge`. Note that you have to make sure that any two nodes of a Hamiltonian path are distinct.

Submission: 20 May 2009, 10:30 am, at the lecture