3. Satisfiability and Containment

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. Satisfiability of Relational Calculus Queries

We say that a relational calculus query Q is *satisfiable* if there is a database instance $\mathbf{I}$ such that

$$Q(\mathbf{I}) \neq \emptyset.$$

Task: Show that satisfiability of relational calculus queries is undecidable.

Note: Even a query without free variables returns the empty tuple over an instance $\mathbf{I}$ if its condition is true over $\mathbf{I}$.

Hint: You can take for granted that Trakhtenbrot's Theorem holds:

> Finite satisfiability of first order logic formulas is undecidable.

Trakhtenbrot's Theorem holds even for formulas without function symbols, constants and equality. As discussed in the lecture, the difficult part of this proof is to make the finite domain, which is implicit in Trakhtenbrot's theorem, explicit

(16 Points)

2. 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, for which another binary relation may come in handy.

Submission: Tuesday, 18 May 2010, 2pm at the lecture