Assignment 3
Due Dec 22, 2016

Part 1: Core of Universal Solutions

In class, we have seen an example of a solution $J$ and an endomorphism $\mu$ over $J$, such that $\mu(J)$ is not a solution. (“Data Exchange,” Page 35.)

**Question 1.1.** Recall the example, and explain why $\mu(J)$ is indeed not a solution.

**Question 1.2.** In this part you will prove the following theorem, originally established by Fagin et al. [2].

**Theorem 1** Let $(S, T, \Sigma)$ be a schema mapping such that $\Sigma$ consists of $s$-TGDs, $t$-TGDs and $t$-EGDs, let $I$ be a source instance, and let $J$ be target instance. If $J$ is a solution, then every core of $J$ is a solution.

To prove the theorem, we need some standard definitions.

**Definitions.** Let $J$ be a $v$-instance, and let $\mu$ be an endomorphism over $J$. Denote by $\mu(J)$ the image of $J$ under $\mu$, that is, the subinstance of $J$ that is obtained from $J$ by replacing every variable $x$ with $\mu(x)$. Let us say that $\mu$ is a core endomorphism if $\mu(J)$ is a core of $J$. For a natural number $m$, we denote by $\mu^m$ the composition of $\mu$ with itself $m$ times; that is:

$$\mu^m(x) \overset{\text{def}}{=} \begin{cases} \mu(x) & \text{if } m = 1; \\ \mu(\mu^{m-1}(x)) & \text{if } m > 1. \end{cases}$$

A fixed point of $\mu$ is a value $x$ such that $\mu(x) = x$. Finally, we will say that $\mu$ is image preserving if every value in the image of $\mu$ is a fixed point (that is, $\mu$ is the identity function over $\mu(J)$).

Use the following steps to prove Theorem 1.

1. If $\mu$ is a core endomorphism, then $\mu^m$ is a core endomorphism for all $m > 1$.
2. If $\mu$ is a core endomorphism, then $\mu$ is an isomorphism over $\mu(J)$.
3. Suppose that (1) $\mu$ is a core endomorphism, and (2) $\mu$ is not image preserving. There exists some $m > 0$ such that $\mu^m$ has a fixed point that $\mu$ does not have.
4. There exists a core endomorphism over $J$ that is image preserving.
5. If $J$ is a solution and $\mu$ is a core endomorphism that is image preserving, then $\mu(J)$ is a solution.
6. If $K$ is a core of a solution, then $K$ is a solution.

**Note:** You can use the fact that the cores of a target instance are isomorphic to each other.
Part 2: Rehearsal

Following is a special case of the task of reverse query processing [1].

<table>
<thead>
<tr>
<th>Input:</th>
<th>Goal:</th>
</tr>
</thead>
<tbody>
<tr>
<td>◦ A schema $T$ with Functional Dependencies (FDs);</td>
<td>Construct an instance $J$ over $T$ such that $Q(J) = r$, or determine that no such $J$ exists.</td>
</tr>
<tr>
<td>◦ An acyclic CQ $Q(x_1,\ldots,x_k)$, without constants, over $T$;</td>
<td></td>
</tr>
<tr>
<td>◦ A $k$-ary relation $r$.</td>
<td></td>
</tr>
</tbody>
</table>

1. No instance $J$ exists, but if we eliminated the FDs then it would exist.

2. There are no FDs, and yet, no instance $J$ exists.

Question 2.1. Show examples of input where:

- No instance $J$ exists, but if we eliminated the FDs then it would exist.
- There are no FDs, and yet, no instance $J$ exists.

Question 2.2. Devise a polynomial-time algorithm for solving the task. Prove the correctness of your algorithm, and explain why it terminates in polynomial time.

Note: We use combined complexity here, which means that you cannot assume that $S$ and $Q$ are fixed. In particular, the fact that $Q$ is acyclic does not mean that, given a candidate instance $J'$, you can construct $Q(J')$ in polynomial time!

Hint terms: Data exchange, universal solution, chase, polynomial total time.

Good luck!

References
