4. Incomplete Databases and Tables

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.
To benefit as much as possible from the exercises, try to solve them using only the hints provided on the sheet.

These exercises are about Codd tables, v-tables, and c-tables as introduced in the lecture. Clearly, v-tables are a special case of c-tables, since they can be seen as c-tables where every row has the condition `true`. Codd tables are a special case of v-tables, since they are v-tables where each variable occurs at most once. When we talk about tables in these exercises, we refer to all three kinds of tables, or, equivalently, to the class of c-tables.

1. Comparison between *rep*$(\cdot)$ and *Rep*$(\cdot)$

What are the differences between *rep*$(\cdot)$ and *Rep*$(\cdot)$?

1. Is there a table $\mathbf{T}$ such that *rep*$(\mathbf{T}) \neq$ *Rep*$(\mathbf{T})$?
2. Let $\mathbf{T}_1$, $\mathbf{T}_2$ be tables.
   - Does *rep*$(\mathbf{T}_1) =$ *rep*$(\mathbf{T}_2)$ imply that *Rep*$(\mathbf{T}_1) =$ *Rep*$(\mathbf{T}_2)$?
   - Does *Rep*$(\mathbf{T}_1) =$ *Rep*$(\mathbf{T}_2)$ imply that *rep*$(\mathbf{T}_1) =$ *rep*$(\mathbf{T}_2)$?

If your answer to one of these questions is "yes", sketch a proof. If your answer is "no", give a counterexample.

(6 Points)

2. The Membership Problem for Tables

Let $\mathcal{T}$ be a class of tables. The *rep*$(\cdot)$-membership problem for $\mathcal{T}$ is defined as follows:

Given: a table $\mathbf{T} \in \mathcal{T}$, an instance I

Question: is $I \in \mathit{rep}(\mathbf{T})$?

Task: For each of the three classes defined in the lecture (Codd tables, v-tables, c-tables), assess the difficulty of the *rep*$(\cdot)$-membership problem.

(12 Points)

Hint: The upper bound for the membership problem for v-tables and c-tables should be straightforward. For the lower bounds, make use of the reductions that you know from the hardness results for query evaluation and containment.
To obtain an upper bound for the class of Codd tables, have a look at the "Maximum Matching Problem" for bipartite graphs (Google will lead you to all you need to know for this exercise).

3. The Containment Problem for Codd-Tables and v-Tables

Let $\mathcal{T}$ be a class of tables. The *rep*$(\cdot)$-containment problem for $\mathcal{T}$ is defined as follows:

Given: tables $\mathbf{T}_1, \mathbf{T}_2 \in \mathcal{T}$

Question: is $\mathit{rep}(\mathbf{T}_1) \subseteq \mathit{rep}(\mathbf{T})_2$?

Task: Assess the difficulty of the *rep*$(\cdot)$-containment problem for the classes of Codd tables and v-tables.

(12 Points)

Hint: First, give a characterization of containment. Then find out the difficulty by reusing results from Exercise 2.

Submission: 31 May, 2011, 10:30 am, at the lecture