

Logic, Data, and Incomplete Information

The Second Tsinghua Logic Summer School

Homework Assignment 3

(due Friday, July 1, 2022 - before class)

Problem 1: Let \mathbf{R} be a relation schema with attributes A, B, C, D, E and let R be the following instance of \mathbf{R}

A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
a_1	b_2	c_2	d_2	d_1
a_2	b_1	c_3	d_3	e_1
a_2	b_1	c_4	d_3	e_1
a_3	b_2	c_5	d_1	e_1

Consider the following functional dependencies:

$$A \rightarrow D, AB \rightarrow D, C \rightarrow BDE, E \rightarrow A, B \rightarrow A.$$

Which of these dependencies does R satisfy and which it does not? Explain why.

Problem 2: Let Σ be a fixed finite set of functional dependencies on a relational schema \mathbf{S} . Describe a polynomial-time algorithm for the following decision problem: given two databases I and J over \mathbf{S} , is J a repair for I ? Explain why your algorithm runs in time bounded by a polynomial in the sizes $|I|$ and $|J|$ of the databases I and J .

Problem 3: Let q_1 be the “path” query $\exists x, y, z(R(x, y) \wedge S(y, z))$. Show that $\text{CERTAINTY}(q_1, \Sigma)$ is FO-rewritable as

$$q^* \equiv \exists x, y, z(R(x, y) \wedge S(y, z) \wedge \forall y'(R(x, y') \rightarrow \exists z'S(y', z'))).$$

In other words, show that for every database I , we have that

$$\text{CONS}(q_1, I, \Sigma) = 1 \text{ if and only if } q^*(I) = 1.$$