Data Integration and Exchange, Homework 2

Problem 1 (20 marks) Suppose we have two relations $R(X, Y)$ and $S(X, Y)$. Give an example of an SQL query over R and S so that:

1. If R and S contain no nulls, it correctly computes $\pi_X(R) - \pi_X(S)$.
2. It does not use the `IS NULL` condition.
3. For some relations R and S so that $\pi_X(R) = \{1, 2\}$ and $\pi_X(S) = \{1\}$ (in particular, there are no nulls in attribute X), it produces an empty table.

Problem 2 (30 marks) Consider a data exchange setting where the source schema contains two relations $S_1(A, B)$ and $S_2(B, C)$ and the target schema contains two relations $T_1(A, D)$ and $T_2(D, C)$. The mapping is given by

$$T_1(x, z),\ T_2(z, y),\ T_2(z, z) \quad :- \quad S_1(x, x'), S_2(x', y)$$

Assume that we have the following source:

S_1:

A	B
1	2
1	3
1	5
2	6

S_2:

B	C
2	2
3	2
7	3
5	1

- (10 marks) Construct the canonical universal solution.
- (10 marks) Consider the query $Q = \pi_{AC}(T_1 \bowtie T_2)$. What are the certain answers to this query for the given source? Can you compute them using the canonical universal solution? If so, how?
- (10 marks) Answer the previous two questions when the source query in the mapping is changed to be $\exists x'\ S_1(x, x') \wedge S_2(x', y)$.

Problem 3 (10 marks) Using the schemas, the mapping, and the instances of Problem 1, give an example of a relational algebra (or calculus) query Q and two universal solutions U_1 and U_2 such that $Q(U_1) \neq Q(U_2)$.

You have to explain why U_1 and U_2 you constructed are universal, and compute both $Q(U_1)$ and $Q(U_2)$.

Problem 4 (40 marks) Consider mappings Σ and Δ shown on slide 14 from the "Schema mappings" set.

Let mapping Γ between σ and ω be given by

$$\forall n\ \exists y\ \forall c\ \big(\text{Takes}(n, c)\ \rightarrow\ \text{Enroll}(y, c)\big)$$

Does Γ define the composition of Σ and Δ?

If it does, prove it; if it does not, give a counterexample.