1. Consider a relational database schema consisting of a unary relational schema BOTTOM and a binary relational schema NEXT-LEVEL. This schema describes a hierarchical organization in which the members of BOTTOM are at the bottom of the hierarchy, while NEXT-LEVEL(a,b) holds if b is one level higher than a.

(a) Give a Datalog program \( \pi \) that defines the query SAME-LEVEL, so that SAME-LEVEL(c,d) holds if c and d are at the same level of the hierarchy. If your program has more than one recursive predicates, then indicate which one is used to define SAME-LEVEL.

(b) Give a recursive equation in relational algebra (or a system of recursive equations in relational algebra if your program has more than one recursive predicates) whose smallest solution is the declarative semantics of your Datalog program.

(c) Give the induction for the bottom-up evaluation of your Datalog program and then derive a relational calculus expression that defines the result of the second iteration of this evaluation (recall that at the 0-th iteration all recursive predicates are set to \( \emptyset \). Refer to the slides of Lectures 12 and 13 for the definition of the bottom-up evaluation).

(d) Can you translate your Datalog program to SQL? Justify briefly your answer.

2. Let PARENT be a binary relational schema such that PARENT(a,b) means that a is a parent of b.

(a) Give a Datalog program that expresses the ternary query \( Q(x, y, z) \) that asserts that x is a descendant of both y and z. If your program has more than one recursive predicates, then indicate which one is used to define \( Q \).

(b) Give a recursive equation in relational algebra (or a system of recursive equations in relational algebra if your program has more than one recursive predicates) whose smallest solution is the declarative semantics of your Datalog program.

(c) Give a recursive equation in relational calculus (or a system of recursive equations in relational calculus if your program has more than one recursive predicates) whose smallest solution is the declarative semantics of your Datalog program.

(d) Express the above query \( Q \) in SQL.

3. A Datalog program \( \pi \) with a single recursive predicate is said to be bounded if there is a positive integer \( n_0 \) such that, on every database instance \( I \), the bottom-up evaluation of \( \pi \) terminates within at most \( n_0 \) steps. Otherwise, \( \pi \) is said to be unbounded.

(a) Consider the Datalog program for the TRANSITIVE CLOSURE query:

\[
\begin{align*}
T(x, y) & : \leftarrow E(x, y) \\
T(x, y) & : \leftarrow E(x, z), T(z, y).
\end{align*}
\]

Show that this program is unbounded. Justify carefully your answer.
(b) Answer the same question for the Datalog program for Cook’s Path Systems query:

\[
\begin{align*}
T(x) & : - A(x) \\
T(x) & : - R(x, y, z), T(y), T(z).
\end{align*}
\]

(c) Give an example of a bounded Datalog program with at least one recursive predicate. Justify why the program you gave is bounded.

4. Let \( R(A, B, C, D, E) \) be a relation schema with the indicated attributes and assume that the following dependencies hold: \( A \rightarrow B, BC \rightarrow E, ED \rightarrow A \).

Find all candidate keys of \( R \) (and justify your answers).

5. Let \( R(A, B, C, D, E) \) be a relation schema with the indicated attributes and let \( F = \{ AB \rightarrow CE, D \rightarrow CE \} \).

(a) Using only the definitions of functional dependency and logical implication, determine whether or not \( F \models ABCE \rightarrow D \).

(b) Use the closure algorithm to determine whether or not \( F \models ABCE \rightarrow D \).

6. Prove the soundness of the closure algorithm, i.e., show, by induction on \( n \), that \( X_n \subseteq X^+ \), for every \( n \geq 0 \). Refer to Lecture 14 for the definitions and the notation.

Note 1: All written assignments, term project report or term paper, and final examination are to be submitted in a typeset format (preferably in \LaTeX{}).

Note 2: The purpose of the homework assignments is to help you develop a better understanding of the material covered in class and also to expose you to some material that lack of time prevents us from covering in class. You should try to work out these problems on your own using your notes from the class. If you discuss any of the problems with other students in the class, you are still expected to produce your own write-up of the assignment; moreover, you are expected to state the names of all other students with whom you have shared ideas about the problems. Finally, please state what published sources you used, if any, in the solutions of the problems.