Foundations of Databases

Exercises – Datalog with Negation

March 29, 2005

1. Show that for the $\mathbf{T}_P$ operator extended to negation, it holds for every datalog$^\neg$ program P that $\mathbf{K} \in inst(sch(P))$ is a model of Σ_P if and only if $\mathbf{T}_P(\mathbf{K}) \subseteq \mathbf{K}$.

2. Show that for datalog$^\neg$ programs P, a minimal fixpoint of $\mathbf{T}_P$ is not necessarily a minimal model of Σ_P and conversely, a minimal model of Σ_P is not necessarily a minimal fixpoint of $\mathbf{T}_P$.

3. Exhibit a datalog$^\neg$ program P which uses negation at least once such that $\mathbf{T}_P$ is monotonic.

4. Show that it is decidable whether for a given datalog$^\neg$ program P the operator $\mathbf{T}_P$ is monotonic.

5. Exhibit a datalog$^\neg$ program P and an instance $\mathbf{K} \in inst(sch(P))$ such that $\mathbf{K}$ is a model of Σ_P but not a fixpoint of $\mathbf{T}_P$.

6. Consider a database for metro and bus stations, with two relations `metro:station, next_station` and `bus:station, next_station`. Write stratifiable datalog$^\neg$ programs to answer the following queries:

   (a) Find the pairs (a, b) of stations such that one can go from a to b by metro but not by bus.

   (b) Find the pairs (a, b) of stations such that there is a pure bus route from a to b. A pure bus route is a path $a = s_0, s_1, \ldots, s_n = b$, $n > 0$, such that each (s_i, s_{i+1}) is in bus and there is no path from s_i to s_{i+1} by metro only.

   (c) Find the pairs (a, b) of stations such b can be reached from a by some combination of metro and/or bus, but not by metro or bus alone.

7. Consider an extension of the database in the previous exercise, where both relations have an additional attribute `duration`, such that tuples (a, b, d) represent that going from a to the next station b takes d minutes. Write a stratified datalog$^\neg$ program, using a built-in predicate $<$ on durations, which computes in a 3-ary relation `long_reach` all tuples (a, b, d) such that b is directly reached from a in time d but not shorter, and d is extremal (maximal) over all b, i.e., the shortest time to reach b from anywhere else is largest.

8. Show that for each stratifiable datalog$^\neg$ program P and $\mathbf{I} \in inst(edb(P)$, $P_{strat}(\mathbf{I})$ is a fixpoint of $\mathbf{T}_P$ whose restriction to $edb(P)$ coincides with $\mathbf{I}$. Extra: Show that $P_{strat}(\mathbf{I})$ is in fact a minimal such fixpoint.

9. Describe a datalog$^\neg$ program which computes under inflationary semantics the complement of the transitive closure of a directed graph $G = (V, N)$, whose arcs N are given by a binary relation `arc` and the nodes V are implicit by N.

10. Give an example of a datalog$^\neg$ program such that for some input $\mathbf{I}$, $P_{inf}(\mathbf{I})$ is not a minimal fixpoint of $\mathbf{T}_P^+$ containing $\mathbf{I}$. Can this be sharpened to a program P such that for no input $\mathbf{I}$, $P_{inf}(\mathbf{I})$ is a minimal fixpoint of $\mathbf{T}_P^+$ containing $\mathbf{I}$?