Universität Freiburg   Georges-Köhler Allee, Geb. 51
Institut für Informatik   D-79110 Freiburg
Michael Meier   Tel. (0761) 203-8126
Fang Wei   Tel. (0761) 203-8125

Foundations of Query Languages
Summer semester 2010
June 15, 2010

8. Exercise Set: Datalog

Exercise 1
Consider a directed graph which is given by E(X,Y) (edges). Give a Datalog program which computes the following relations:

a) Odd(X,Y), which holds if there is a path with odd length from X to Y.

b) Oddcycle(X), there is a cycle with odd length through X.

c) Evencycle(X), there is cycle with even length through X.

d) Bothcycles(X), there are cycles with even length and cycles with odd length through X.

Exercise 2
parent(X,Y) is a family tree with root p. Please give a Datalog program, which computes the predicates **same-generation(X,Y)**, **sibling(X,Y)** and **cousin(X,Y)**. (**same-generation(X,Y)** holds, if the distance between X and p is the same as the distance between Y and p; **sibling(X,Y)** is true, if X and Y have the same parent; **cousin(X,Y)** holds, if X and Y belong to the same generation but are not siblings).

Hint: You may use negation in your programs.

Exercise 3
Given are the following three Datalog programs to calculate the transitive closure of a graph:

right-recursive
$$C(X,Y) \leftarrow E(X,Y)$$
$$C(X,Y) \leftarrow E(X,Z), C(Z,Y)$$

left-recursive
$$C(X,Y) \leftarrow E(X,Y)$$
$$C(X,Y) \leftarrow C(X,Z), E(Z,Y)$$

double-recursive
$$C(X,Y) \leftarrow E(X,Y)$$
$$C(X,Y) \leftarrow C(X,Z), C(Z,Y)$$

a) Since the naive evaluation algorithm has the drawback that some facts are calculated more than once, it is possible to formulate an optimized version of it (hint: *semi*-naive algorithm). Please specify an optimized form of the naive algorithm. Note that you do not have to hand in a complete algorithm but rather explain the idea and provide an intuition/derivation as to why your algorithm makes sense. You may look into the "Foundations of Databases" and use the seminaive evaluation strategy from the book. In this case you should explain in very detail how this algorithm works.

b) Use your algorithm to calculate the results of the given three programs with respect to the following database:

$$E(1,2), \ E(2,3), \ E(3,4), \ E(4,5), \ E(5,6), \ E(6,7)$$

Due by: June 23, 2010 before the tutorial starts.