Combined Tractability of Query Evaluation via Tree Automata and Cycluits

Antoine Amarilli, Pierre Bourhis, Mikaël Monet, Pierre Senellart

Problem

Evaluate Datalog queries tractably in parameterized combined complexity

Theorem: Fixed Parameter Tractable

Given a Boolean ICG-Datalog program P of body-size k, and an instance I of treewidth k, we can determine if I satisfies P in FPT-linear complexity: f(k, k, l) |P| • |I|

ICG-Datalog

Intensional Clique-Guarded Datalog:

- Fragment of Datalog with stratified negation
- Clique-guarded: every intensional atom S is guarded by a clique of extensional atoms R.

 $T(x,z) \leftarrow R_1(x, y) \wedge R_2(y, z) \wedge R_3(z, x) \wedge S(x, y, z)$

- Body-size k_D: maximal size of a rule

Languages captured

- α-acylic conjunctive queries (CQs)
- CQs of bounded simplicial width
- Guarded negation fragments
- Monadic Datalog of bounded body-size
- Strongly Acyclic 2RPQs

Approach: Through Provenance

Provenance: Boolean function capturing how the query

result depends on the input database

Input: A subinstance J of I (with the facts as variables)

Output: Does J satisfies P?

Existing representations: formulas, circuits

Cycluits: A New Provenance Representation

Boolean cyclic circuits (cycluits) with stratified negation

Semantics: least fixed-point **Evaluation:** linear time Can be decyclified

Application: Probabilistic Query Evaluation



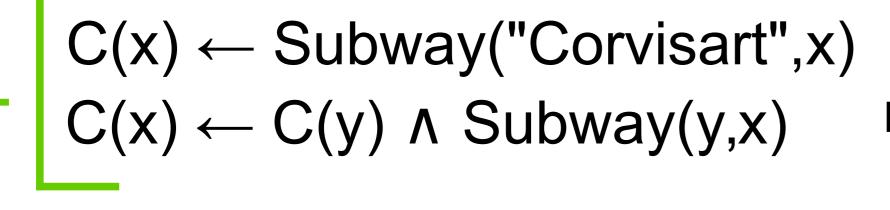






Proof Structure

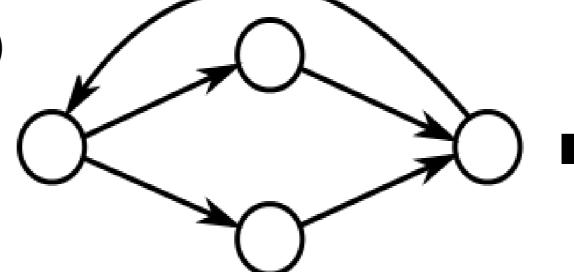
ICG-Datalog program P of body-size ≤ k_D



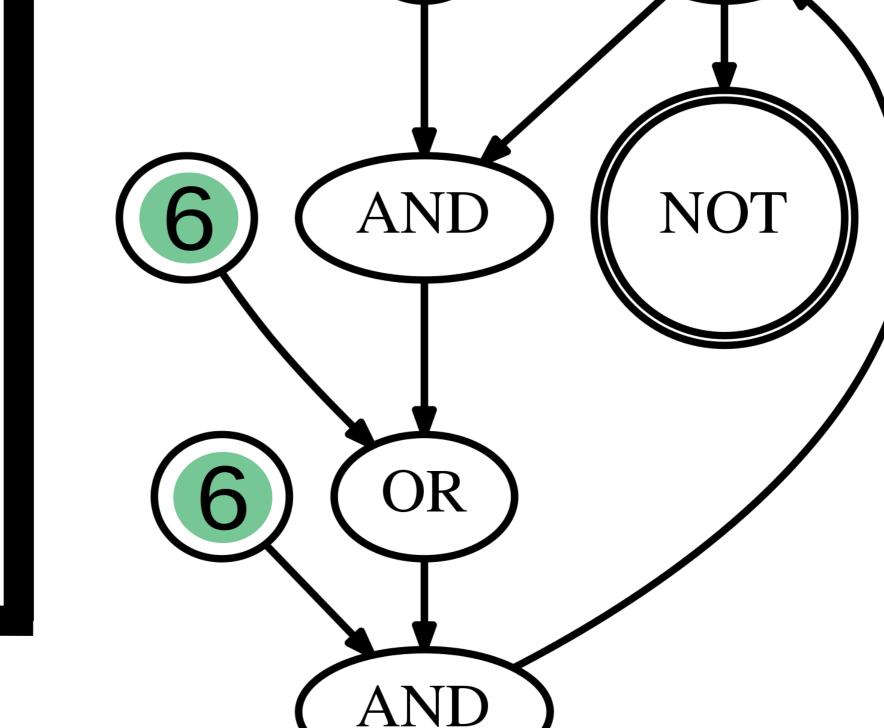
Goal() ← ¬ C("Châtelet")

Two-way Alternating Tree Automaton A

 $O(g(\mathbf{k_P}, \mathbf{k_I})|\mathbf{P}|)$



 $O(|\mathbf{A}| \cdot |\mathbf{E}|)$



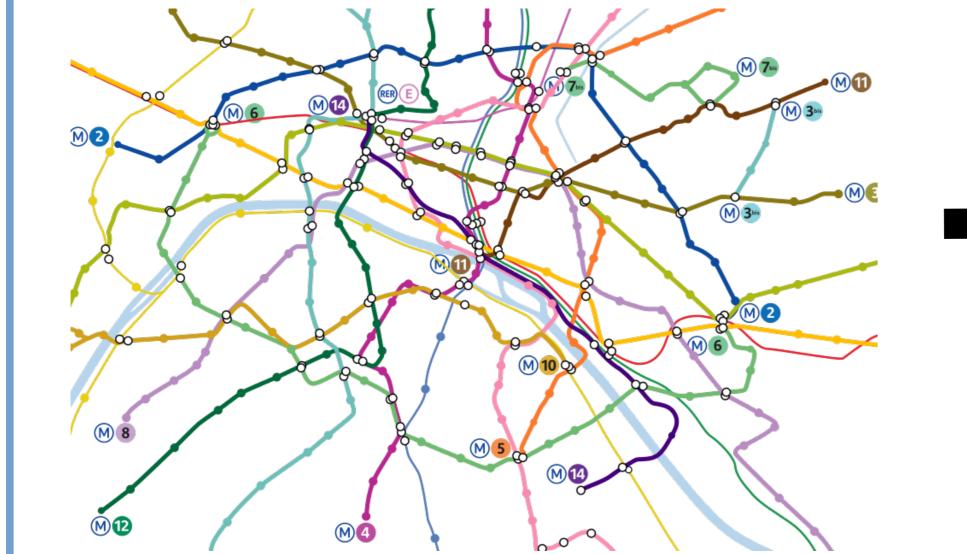
Provenance Cycluit

6

OR

"Under which conditions is it impossible to go from station Corvisart to station Châtelet with the subway?"





(Paris Metro map)

Tree encoding E

