Query Answering with Transitive & Linear-Ordered Data

Antoine Amarilli[1], Michael Benedikt[2],
Pierre Bourhis[3], Michael Vanden Boom[2]

1: LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay; Paris, France
2: University of Oxford; Oxford, United Kingdom
3: CNRS CRIStAL, Université Lille 1, INRIA Lille; Lille, France

Query Answering (or query entailment)

Initial set of facts $\mathcal{F}$ (or A-box)

Constraints Σ: logical rules (or T-box)

Boolean query Q: CQ or UCQ

QA problem: does $\mathcal{F} \wedge \Sigma$ entail Q?

Equivalently:
- is Q certain given $\mathcal{F}$ and Σ?
- is $\mathcal{F} \wedge \Sigma \wedge \neg Q$ unsatisfiable?

Our goal: identify and study constraint languages for which QA is decidable, even when some relations are restricted to be **transitive** or to be **linear orders**.

Our approach

Fix relational signature $\sigma := \sigma_B \sqcup \sigma_D$ where
- σ_D: **distinguished** binary relations
- σ_B: **base** relations

We consider query answering with three different special interpretations for the distinguished relations:

- **QAtr**: each $R \in \sigma_D$ is transitively closed
- **QAtc**: each $R^+ \in \sigma_D$ is the transitive closure of $R \in \sigma_B$
- **QAlin**: each $R \in \sigma_D$ is a linear order

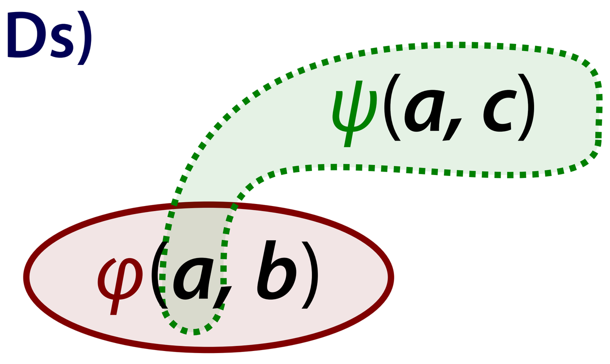
We introduce **base-frontier-guarded** and **base-covered** constraint languages that disallow the use of distinguished relations as guards.

Constraint languages

Tuple-generating dependencies (TGDs)
(or existential rules)

$$\forall x\, y\; \varphi(x, y) \rightarrow \exists z\; \psi(x, z)$$

where φ (body) and ψ (head) are conjunctions of atoms



Frontier-guarded TGDs (FGTGDs)

$$\forall x\, y\; \varphi(x, y) \wedge G(x) \rightarrow \exists z\; \psi(x, z)$$

Atom using variables x is called a **guard** for x

Guarded Negation Fragment (GNF)
rules built up from atoms using
- disjunction
- guarded negation $G(x) \wedge \neg\psi(x)$
- existential quantification

BaseFGTGDs
FGTGDs where guards for frontier variables are from σ_B

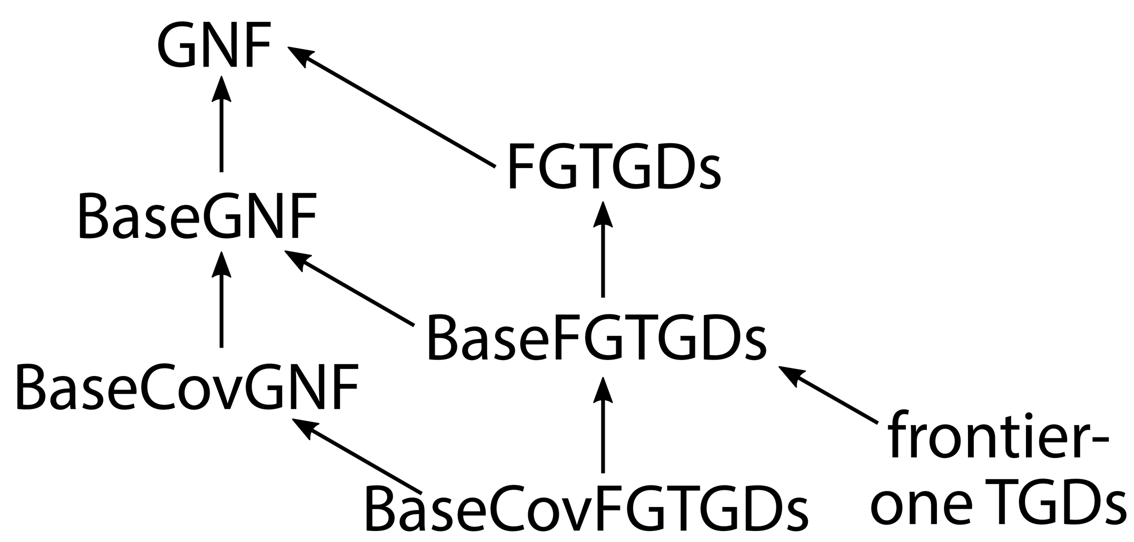
e.g. $\forall x\, y_1\, y_2\; R(x,y_1) \wedge R(x,y_2) \wedge S(y_1,y_2) \rightarrow \exists z\; R(y_2,z) \wedge T(y_1)$
where $\sigma_D = \{R\}$ and $\sigma_B = \{S, T\}$

BaseCovFGTGDs
BaseFGTGDs where for every σ_D atom in the body using variables $\boldsymbol{v}$, there is a σ_B atom in the body guarding $\boldsymbol{v}$

e.g. $\forall x\, y_1\, y_2\; C(x,y_1) \wedge R(x,y_1) \wedge C(x,y_2) \wedge R(x,y_2) \wedge S(y_1,y_2) \rightarrow \exists z\; R(y_2,z) \wedge T(y_1)$
where $\sigma_D = \{R\}$ and $\sigma_B = \{S, T, C\}$

BaseGNF
GNF where guards for negation are from σ_B

BaseCovGNF
Generalization of BaseCovFGTGDs



Main results

QAtr & **QAtc** are decidable for **BaseGNF** (undecidable for FGTGDs).

QAlin is decidable for **BaseCovGNF** (undecidable for BaseFGTGDs).

Complexity

	QAtr data	QAtr combined	QAtc data	QAtc combined	QAlin data	QAlin combined
BaseGNF	coNP-c	2EXP-c	coNP-c	2EXP-c	undecidable	
BaseCovGNF	coNP-c	2EXP-c	coNP-c	2EXP-c	coNP-c	2EXP-c
BaseFGTGDs	in coNP	2EXP-c	coNP-c	2EXP-c	undecidable	
BaseCovFGTGDs	P-c	2EXP-c	coNP-c	2EXP-c	coNP-c	2EXP-c

Proof ideas

Key property:

For φ in GNF, if φ is satisfiable then it has a **tree-like** witness: a set of facts satisfying φ that has a **tree decomposition** of bounded tree-width.

For **QAtr** & **QAtc** with base-frontier-guarded constraints Σ: reduce to tree automaton emptiness test. $L(\mathcal{A}) = \emptyset$?

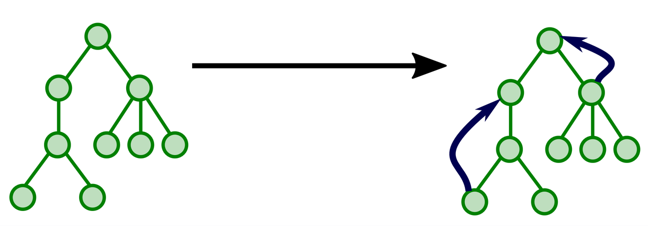
For BaseGNF, there are tree-like witnesses even when each distinguished relation is required to be the transitive closure of some base relation.

Hence it suffices to construct a tree automaton $\mathcal{A}$ that runs on encodings of tree-like sets of facts and checks $\mathcal{F} \wedge \Sigma \wedge \neg Q$.

For **QAtr** & **QAlin** with base-covered constraints Σ: reduce to traditional QA with GNF Σ'.

Cannot axiomatize transitivity or totality using GNF, but can approximate using Σ' in GNF. Key technical result shows that a tree-like approximate witness can be extended to an actual witness respecting special interpretations for σ_D relations.

Tree-like witness for $\mathcal{F} \wedge \Sigma' \wedge \neg Q$



Actual witness for $\mathcal{F} \wedge \Sigma \wedge \neg Q$