Herbrand Property, Finite Quasi-Herbrand Models, and a Chandra-Merlin Theorem for Quantified Conjunctive Queries

Fabio Mogavero
(joint work with Simone Bova)

Università degli Studi di Napoli Federico II

68NQRT Seminar - IRISA/INRIA
December 9, 2021

An Interesting & Important Question

The Question

Why is modal logic so robustly decidable?

The Model-Theoretic Answer

Bounded-tree model property + MSOL Encoding [Vardi, 1996].

Few Further Answers

Encoding in:

- *two-variable fragment* of FOL (FOL[2VAR]) [Mortimer, 1975];
- *guarded fragment* of FOL (FOL[GF]) [Andréka et al., 1998];
- *unary-negation fragment* of FOL (FOL[UN]) [ten Cate and Segoufin, 2011].

Analysis of Fragments

FOL[2VAR] is not robust, since simple extensions become undecidable.

FOL[GF] and FOL[UN] are robust and well behaved, since, as modal logic, they enjoy good model-theoretic and algorithmic properties.

	FOL[GF]	FOL[UN]
Model Checking	PTime	$\text{PTime}^{\text{NPTime}}$
Satisfiability	2ExpTime-complete	2ExpTime-complete
Finite-Model Property	✓	✓
Craig's Interpolation	×	✓
Beth's Definability	✓	✓

FOL[2VAR], FOL[GF], and FOL[UN] are orthogonal *w.r.t.* expressiveness.

Another Significant Question

The Question

Why are extensions of modal logic, as $\mathrm{ATL}^{\star}$ and $\mathrm{SL}[1\mathrm{G}]$, decidable?

The High-Level Answer

Bounded-tree model property + MSOL Encoding [Schewe, 2008, M., Murano, Perelli, Vardi, 2012].

A Specific Problem

Why do they enjoy the bounded-tree model property in the first place?

A First-Order Encoding

Consider the ATL* formula $[\![a, b, c]\!]\neg\psi$ over a game with the four agents a, b, c, and d, where ψ is an LTL formula.

It asserts that agent d has a strategy, which depends upon those chosen by the other three ones, ensuring that the property ψ does not hold.

We can encode this property by means of the FOL sentence $\forall a \forall b \forall c \exists d \neg R_\psi(a, b, c, d)$, where R_ψ is the characteristic relation for ψ, *i.e.*, it is true *iff* the strategy profile (a, b, c, d) satisfies the property.

Classic Constraint Violation

The FOL sentence $\varphi = \forall a \forall b \forall c \exists d \neg R_{\psi}(a, b, c, d)$...

- has more than two variables: $\varphi \notin \mathrm{FOL}[\textsc{2var}]$;
- quantifications are not guarded: $\varphi \notin \mathrm{FOL}[\textsc{gf}]$;
- negation is applied to more than one free variable: $\varphi \notin \mathrm{FOL}[\textsc{un}]$.

We cannot derive the good algorithmic properties of $\mathrm{ATL}^{\star}$ and $\mathrm{SL}[\textsc{1g}]$ from known fragments of FOL.

φ has quantification prefix $\forall^3\exists$, so it does not even belong to any of the decidable prefix-vocabulary classes [Börger et al., 1997].

Our Contributions

1. Introduction of a new way to classify FOL fragments.
2. Definition of a new model-theoretic property which allows to generalise the concept of Herbrand model.

The fragments are based on the binding forms admitted in a sentence, *i.e.*, on the way the arguments of a relation can be bound to a variable.

This study allows to answer:

3. the question about the decidability of ATL* and SL[1G];
4. an open problem in database theory.

A Tale of Logic Connections

1. The *Herbrand Property* of functional first-order models.
2. The decidability of *Conjunctive Binding Logic* satisfiability.
3. The decidability of *Quantified Conjunctive Query* containment.
4. A *Chandra-Merlin* characterisation of QCQ containment.

A Tale of Logic Connections

1. The *Herbrand Property* of functional first-order models.
   - A structure $\mathcal{A}$ enjoys the Herbrand Property if:
     - two terms (*semantically*) *equalize* over $\mathcal{A}$ *iff* they (*syntactically*) *unify*.
2. The decidability of *Conjunctive Binding Logic* satisfiability.
3. The decidability of *Quantified Conjunctive Query* containment.
4. A *Chandra-Merlin* characterisation of QCQ containment.

A Tale of Logic Connections

1. The *Herbrand Property* of functional first-order models.
2. The decidability of *Conjunctive Binding Logic* satisfiability.
   - ⋇ CBL is a fragment of First-Order Logic:
     - quantifications of *conjunctions* of formulas on the *same bindings*;
     - example: $\forall x \exists y \forall z \,.\, R(x, y) \wedge (R(y, z) \vee \neg P(y, z))$.
3. The decidability of *Quantified Conjunctive Query* containment.
4. A *Chandra-Merlin* characterisation of QCQ containment.

A Tale of Logic Connections

1. The *Herbrand Property* of functional first-order models.

2. The decidability of *Conjunctive Binding Logic* satisfiability.

3. The decidability of *Quantified Conjunctive Query* containment.

   ⁜ QCQs are an extension of classic Conjunctive Queries:
   - *arbitrary quantifications* of conjunctions of *positive atoms*;
   - fragment of CBL with *no disjunction or negation*;
   - example: $\forall x \exists y \forall z \,.\, R(x, y) \wedge R(y, z) \wedge P(z, y, x)$.

4. A *Chandra-Merlin* characterisation of QCQ containment.

A Tale of Logic Connections

1. The *Herbrand Property* of functional first-order models.
2. The decidability of *Conjunctive Binding Logic* satisfiability.
3. The decidability of *Quantified Conjunctive Query* containment.
4. A *Chandra-Merlin* characterisation of QCQ containment.
   - ⋇ Generalisation of CQ containment via homomorphism:
     - $\varphi_1 \models \varphi_2$ *iff* there is an *"homomorphism"* from φ_2 to φ_1 that *complies with* the *quantification prefixes* of both φ_1 and φ_2.

Quantified Conjunctive Queries

Quantified Conjunctive Queries

Relational fragment of first-order logic obtained by using $\forall$, $\exists$, and $\wedge$.

Example

$\forall x \exists y \forall z.(Px \wedge Rxy \wedge Ryz \wedge Qz)$

- *Relational positive Herbrand logic* in mathematical logic.
- *Quantified constraint satisfaction problems* in constraint satisfaction.

Conjunctive queries (CQs) are QCQs without $\forall$.

Entailment/Containment of QCQs

The Problem

Instance: A pair (φ_1, φ_2) of QCQs.

Question: $\varphi_1 \models \varphi_2$?

Chen, Madelaine, & Martin (LICS'08 & LMCS'15)

Question: Is entailment of QCQs decidable?

Answer: Yes it is, belongs to 3ExpTime, but NPTime-hard!

Question: Do entailment and finite entailment coincide?

Answer: Open problem.

Question: Is finite entailment decidable?

Answer: Open problem.

Entailment/Containment of QCQs

The Problem

Instance: A pair (φ_1, φ_2) of QCQs.

Question: $\varphi_1 \models \varphi_2$?

Chen, Madelaine, & Martin (LICS'08 & LMCS'15)

Question: Is entailment of QCQs decidable?

Answer: Yes it is, belongs to 3ExpTime, but NPTime-hard!

Question: Do entailment and finite entailment coincide?

Answer: Open problem.

Question: Is finite entailment decidable?

Answer: Open problem.

Entailment/Containment of QCQs

The Problem

Instance: A pair (φ_1, φ_2) of QCQs.

Question: $\varphi_1 \models \varphi_2$?

Chen, Madelaine, & Martin (LICS'08 & LMCS'15)

Question: Is entailment of QCQs decidable?

Answer: Yes it is, belongs to 3ExpTime, but NPTime-hard!

Question: Do entailment and finite entailment coincide?

Answer: Open problem.

Question: Is finite entailment decidable?

Answer: Open problem.

Entailment/Containment of QCQs

The Problem

Instance: A pair (φ_1, φ_2) of QCQs.

Question: $\varphi_1 \models \varphi_2$?

Bova & M. (LICS'17)

- Entailment and finite entailment of QCQs coincide.
- NPTime-complete.

Idea Behind the NPTime Upper Bound

Observation

Positive instances of CQ Containment have small refutations!

$$\exists x\,.\,R(x,x) \models \exists y,z,w\,.\,R(y,z) \wedge R(z,w)$$

iff

$R(c_x,c_x) \wedge \forall y,z,w\,.\,\neg R(y,z) \vee \neg R(z,w)$ is unsatisfiable.

$$\dfrac{R(c_x,c_x) \qquad \dfrac{R(c_x,c_x) \qquad \neg R(y,z) \vee \neg R(z,w)}{\neg R(c_x,w)}\ {\scriptstyle y,\,z \,\mapsto\, c_x}}{\bot}\ {\scriptstyle w \,\mapsto\, c_x}$$

From the unification $y,z,w \mapsto c_x$ we can derive the homomorphism $y,z,w \mapsto x$ from $R(y,z), R(z,w)$ to $R(x,x)$.

Idea Behind the NPTime Upper Bound

Observation

Positive instances of CQ Containment have small refutations!

$$\exists x \,.\, R(x,x) \models \exists y,z,w \,.\, R(y,z) \wedge R(z,w)$$

iff

$R(c_x, c_x) \wedge \forall y,z,w \,.\, \neg R(y,z) \vee \neg R(z,w)$ is unsatisfiable.

$$\cfrac{R(c_x,c_x) \qquad \cfrac{R(c_x,c_x) \qquad \neg R(y,z) \vee \neg R(z,w)}{\neg R(c_x,w)}\ {\scriptstyle y,\,z \,\mapsto\, c_x}}{\bot}\ {\scriptstyle w \,\mapsto\, c_x}$$

From the unification $y,z,w \mapsto c_x$ we can derive the homomorphism $y,z,w \mapsto x$ from $R(y,z), R(z,w)$ to $R(x,x)$.

Idea Behind the NPTime Upper Bound

Observation

Positive instances of CQ Containment have small refutations!

$$\exists x\,.\,R(x,x) \models \exists y,z,w\,.\,R(y,z) \wedge R(z,w)$$

iff

$R(c_x,c_x) \wedge \forall y,z,w\,.\,\neg R(y,z) \vee \neg R(z,w)$ is unsatisfiable.

$$\dfrac{R(c_x,c_x) \qquad \dfrac{R(c_x,c_x) \qquad \neg R(y,z) \vee \neg R(z,w)}{\neg R(c_x,w)}\;{\scriptstyle y,\,z \,\mapsto\, c_x}}{\bot}\;{\scriptstyle w \,\mapsto\, c_x}$$

From the unification $y,z,w \mapsto c_x$ we can derive the homomorphism $y,z,w \mapsto x$ from $R(y,z), R(z,w)$ to $R(x,x)$.

Idea Behind the NPTime Upper Bound

Observation

Positive instances of CQ Containment have small refutations!

$$\exists x \,.\, R(x,x) \models \exists y,z,w \,.\, R(y,z) \wedge R(z,w)$$

iff

$R(c_x, c_x) \wedge \forall y,z,w \,.\, \neg R(y,z) \vee \neg R(z,w)$ is unsatisfiable.

$$\dfrac{R(c_x,c_x) \qquad \dfrac{R(c_x,c_x) \qquad \neg R(y,z) \vee \neg R(z,w)}{\neg R(c_x,w)}\ {\scriptstyle y,z \mapsto c_x}}{\perp}\ {\scriptstyle w \mapsto c_x}$$

From the unification $y,z,w \mapsto c_x$ we can derive the homomorphism $y,z,w \mapsto x$ from $R(y,z), R(z,w)$ to $R(x,x)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations as well!

$$\exists x \forall y \,.\, R(x,y) \models \forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$$

iff

$\forall y \,.\, R(c_x, y) \wedge \forall z \,.\, \neg R(z, c_u) \vee \neg R(z, c_v)$ is unsatisfiable.

$$\dfrac{R(c_x, y_2) \qquad \dfrac{R(c_x, y_1) \qquad \neg R(z, c_u) \vee \neg R(z, c_v)}{\neg R(c_x, c_v)}\; {\scriptstyle y_1 \mapsto c_u;\, z \mapsto c_x}}{\bot}\; {\scriptstyle y_2 \mapsto c_v}$$

From the unification $y_1 \mapsto c_u; y_2 \mapsto c_v; z \mapsto c_x$ we can derive a Skolem homomorphism from $\forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$ to $\exists x \forall y \,.\, R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations as well!

$$\exists x \forall y \,.\, R(x,y) \models \forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$$

iff

$\forall y \,.\, R(c_x,y) \wedge \forall z \,.\, \neg R(z,c_u) \vee \neg R(z,c_v)$ is unsatisfiable.

$$\dfrac{R(c_x,y_2) \qquad \dfrac{R(c_x,y_1) \qquad \neg R(z,c_u) \vee \neg R(z,c_v)}{\neg R(c_x,c_v)}\; y_1 \mapsto c_u;\, z \mapsto c_x}{\bot}\; y_2 \mapsto c_v$$

From the unification $y_1 \mapsto c_u; y_2 \mapsto c_v; z \mapsto c_x$ we can derive a Skolem homomorphism from $\forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$ to $\exists x \forall y \,.\, R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations as well!

$$\exists x \forall y \,.\, R(x,y) \models \forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$$

iff

$\forall y \,.\, R(c_x, y) \wedge \forall z \,.\, \neg R(z, c_u) \vee \neg R(z, c_v)$ is unsatisfiable.

$$\dfrac{R(c_x, y_2) \qquad \dfrac{R(c_x, y_1) \qquad \neg R(z, c_u) \vee \neg R(z, c_v)}{\neg R(c_x, c_v)}\; {\scriptstyle y_1 \mapsto c_u;\, z \mapsto c_x}}{\bot}\; {\scriptstyle y_2 \mapsto c_v}$$

From the unification $y_1 \mapsto c_u; y_2 \mapsto c_v; z \mapsto c_x$ we can derive a Skolem homomorphism from $\forall u,v\, \exists z \,.\, (R(z,u) \wedge R(z,v))$ to $\exists x \forall y \,.\, R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations as well!

$$\exists x \forall y \,.\, R(x,y) \models \forall u,v \,\exists z \,.\, (R(z,u) \wedge R(z,v))$$

iff

$\forall y \,.\, R(c_x, y) \wedge \forall z \,.\, \neg R(z, c_u) \vee \neg R(z, c_v)$ is unsatisfiable.

$$\dfrac{R(c_x, y_2) \qquad \dfrac{R(c_x, y_1) \qquad \neg R(z, c_u) \vee \neg R(z, c_v)}{\neg R(c_x, c_v)}\; y_1 \mapsto c_u; z \mapsto c_x}{\bot}\; y_2 \mapsto c_v$$

From the unification $y_1 \mapsto c_u; y_2 \mapsto c_v; z \mapsto c_x$ we can derive a Skolem homomorphism from $\forall u,v \,\exists z \,.\, (R(z,u) \wedge R(z,v))$ to $\exists x \forall y \,.\, R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations too!

$$\forall y \exists x \,.\, R(x,y) \models \forall u\, \exists z,v \,.\, (R(z,u) \wedge R(v,z))$$

iff

$\forall y \,.\, R(f_x(y),y) \wedge \forall z,v \,.\, \neg R(z,c_u) \vee \neg R(v,z)$ is unsatisfiable.

$$\dfrac{R(f_x(y_2),y_2) \qquad \dfrac{R(f_x(y_1),y_1) \qquad \neg R(z,c_u) \vee \neg R(v,z)}{\neg R(v, f_x(c_u))}\, y_1 \mapsto c_u;\, z \mapsto f_x(c_u)}{\bot}\, y_2 \mapsto f_x(c_u),\, v \mapsto f_x(f_x(c_u))$$

The unification $y_1 \mapsto c_u; y_2, z \mapsto f_x(c_u); v \mapsto f_x(f_x(c_u))$ induces a Skolem homomorphism from $\forall u \exists z,v.(R(z,u) \wedge R(v,z))$ to $\forall y \exists x.R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations too!

$$\forall y \exists x \,.\, R(x,y) \models \forall u\, \exists z,v \,.\, (R(z,u) \land R(v,z))$$

iff

$\forall y \,.\, R(f_x(y),y) \land \forall z,v \,.\, \neg R(z,c_u) \lor \neg R(v,z)$ is unsatisfiable.

$$\dfrac{R(f_x(y_2),y_2) \qquad \dfrac{R(f_x(y_1),y_1) \qquad \neg R(z,c_u) \lor \neg R(v,z)}{\neg R(v, f_x(c_u))}\; {\scriptstyle y_1 \mapsto c_u;\, z \mapsto f_x(c_u)}}{\bot}\; {\scriptstyle y_2 \mapsto f_x(c_u),\, v \mapsto f_x(f_x(c_u))}$$

The unification $y_1 \mapsto c_u; y_2, z \mapsto f_x(c_u); v \mapsto f_x(f_x(c_u))$ induces a Skolem homomorphism from $\forall u \exists z,v.(R(z,u) \land R(v,z))$ to $\forall y \exists x.R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations too!

$$\forall y \exists x \,.\, R(x,y) \models \forall u\, \exists z,v \,.\, (R(z,u) \wedge R(v,z))$$

iff

$\forall y \,.\, R(f_x(y), y) \wedge \forall z,v \,.\, \neg R(z, c_u) \vee \neg R(v,z)$ is unsatisfiable.

$$\dfrac{R(f_x(y_2), y_2) \qquad \dfrac{R(f_x(y_1), y_1) \qquad \neg R(z, c_u) \vee \neg R(v,z)}{\neg R(v, f_x(c_u))}\; {\scriptstyle y_1 \mapsto c_u;\, z \mapsto f_x(c_u)}}{\bot}\; {\scriptstyle y_2 \mapsto f_x(c_u),\, v \mapsto f_x(f_x(c_u))}$$

The unification $y_1 \mapsto c_u; y_2, z \mapsto f_x(c_u); v \mapsto f_x(f_x(c_u))$ induces a Skolem homomorphism from $\forall u \exists z,v.(R(z,u) \wedge R(v,z))$ to $\forall y \exists x.R(x,y)$.

Idea Behind the NPTime Upper Bound

Result

Positive instances of QCQ Containment have small refutations too!

$$\forall y \exists x \,.\, R(x,y) \models \forall u\, \exists z,v \,.\, (R(z,u) \wedge R(v,z))$$

iff

$$\forall y \,.\, R(f_x(y), y) \wedge \forall z,v \,.\, \neg R(z, c_u) \vee \neg R(v,z) \text{ is unsatisfiable.}$$

$$\cfrac{R(f_x(y_2), y_2) \qquad \cfrac{R(f_x(y_1), y_1) \qquad \neg R(z, c_u) \vee \neg R(v,z)}{\neg R(v, f_x(c_u))}\; {\scriptstyle y_1 \mapsto c_u;\, z \mapsto f_x(c_u)}}{\bot}\; {\scriptstyle y_2 \mapsto f_x(c_u),\, v \mapsto f_x(f_x(c_u))}$$

The unification $y_1 \mapsto c_u; y_2, z \mapsto f_x(c_u); v \mapsto f_x(f_x(c_u))$ induces a Skolem homomorphism from $\forall u \exists z,v.(R(z,u) \wedge R(v,z))$ to $\forall y \exists x.R(x,y)$.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Idea Behind the NPTime Upper Bound

Definition

A *Skolem homomorphism* from $\varphi_2 = \wp_2\chi_2$ to $\varphi_1 = \wp_1\chi_1$ is a substitution σ of variables of $\wp_2$ by terms on the vocabulary of $\mathsf{sk}_{\wp_1}(\chi_1)$ such that:

1. universal variables in $\wp_2$ maps injectively via σ;
2. every existential variable x in $\wp_2$ maps to the term $\sigma(x)$ not containing images of universal variables after x in $\wp_2$;
3. every atom $R(x_1, \ldots, x_k)$ in χ_2 has an atom $R(t_1, \ldots, t_k)$ in $\mathsf{sk}_{\wp_1}(\chi_1)$ such that $R(\sigma(x_1), \ldots, \sigma(x_k))$ and $R(t_1, \ldots, t_k)$ unify.

Chandra-Merlin (STOC'77) - Theorem for CQ

$\varphi_1 \models \varphi_2$ *iff* there is an homomorphism from φ_2 to φ_1.

Bova & M. (LICS'17) - Theorem for QCQ

$\varphi_1 \models \varphi_2$ *iff* there is a Skolem homomorphism from φ_2 to φ_1.

Finite Entailment/Containment of QCQs

$\varphi_1 \models_{fin} \varphi_2$ *iff* φ_2 is satisfied on all *finite* models of φ_1.

Finite entailment is the relevant notion in CS applications (*e.g.*, in query optimisation the database is finite).

Proof Approach

We proved the collapse of entailment $\varphi_1 \models \varphi_2$ and finite entailment $\varphi_1 \models_{fin} \varphi_2$ of QCQs via a *domain-preserving* reduction to the satisfiability check of a CBL sentence ψ.

FOL-provers solve (in practice) the finite QCQ-entailment problem!

From Entailment to Finite Entailment

Proof Approach

We proved the collapse of entailment $\varphi_1 \models \varphi_2$ and finite entailment $\varphi_1 \models_{fin} \varphi_2$ of QCQs via a *domain-preserving* reduction to the satisfiability check of a CBL sentence ψ.

Observations

1. The results exploits the finite-model property of CBL.
2. It is more involved than the classic check for the satisfiability of $\varphi_1 \wedge \neg\varphi_2$, as the latter belongs to an undecidable fragment of FOL.

- ψ unsat $\Rightarrow$ $\varphi_1 \models \varphi_2$ $\Rightarrow$ $\varphi_1 \models_{fin} \varphi_2$
- ψ sat $\Rightarrow$ $\mathcal{A} \models_{fin} \psi$ $\Rightarrow$ $\mathcal{A} \models_{fin} \varphi_1 \wedge \neg\varphi_2$ $\Rightarrow$ $\varphi_1 \not\models_{fin} \varphi_2$ $\Rightarrow$ $\varphi_1 \not\models \varphi_2$

From Entailment to Finite Entailment

Proof Approach

We proved the collapse of entailment $\varphi_1 \models \varphi_2$ and finite entailment $\varphi_1 \models_{fin} \varphi_2$ of QCQs via a *domain-preserving* reduction to the satisfiability check of a CBL sentence ψ.

Observations

1. The results exploits the finite-model property of CBL.
2. It is more involved than the classic check for the satisfiability of $\varphi_1 \land \neg\varphi_2$, as the latter belongs to an undecidable fragment of FOL.

- ψ unsat $\Rightarrow$ $\varphi_1 \models \varphi_2$ $\Rightarrow$ $\varphi_1 \models_{fin} \varphi_2$
- ψ sat $\Rightarrow$ $\mathcal{A} \models_{fin} \psi$ $\Rightarrow$ $\mathcal{A} \models_{fin} \varphi_1 \land \neg\varphi_2$ $\Rightarrow$ $\varphi_1 \not\models_{fin} \varphi_2$ $\Rightarrow$ $\varphi_1 \not\models \varphi_2$

From Entailment to Finite Entailment

Proof Approach

We proved the collapse of entailment $\varphi_1 \models \varphi_2$ and finite entailment $\varphi_1 \models_{fin} \varphi_2$ of QCQs via a *domain-preserving* reduction to the satisfiability check of a CBL sentence ψ.

Observations

1. The results exploits the finite-model property of CBL.
2. It is more involved than the classic check for the satisfiability of $\varphi_1 \wedge \neg\varphi_2$, as the latter belongs to an undecidable fragment of FOL.

- ψ unsat $\Rightarrow$ $\varphi_1 \models \varphi_2$ $\Rightarrow$ $\varphi_1 \models_{fin} \varphi_2$
- ψ sat $\Rightarrow$ $\mathcal{A} \models_{fin} \psi$ $\Rightarrow$ $\mathcal{A} \models_{fin} \varphi_1 \wedge \neg\varphi_2$ $\Rightarrow$ $\varphi_1 \not\models_{fin} \varphi_2$ $\Rightarrow$ $\varphi_1 \not\models \varphi_2$

Conjunctive Binding Logic

Conjunctive Binding Logic

CBL Syntax

Positive Boolean combinations of sentences $\wp(\psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_k)$:

- $\wp$ is an arbitrary quantification prefix;
- ψ_i is a Boolean combinations of atoms over the same binding (*i.e.*, same association of variables with positions in a relation).

- $\forall x \exists y \forall z \,.\, R(x, y) \wedge (R(y, z) \vee \neg P(y, z))$ ✓
- $\forall x \exists y \forall z \,.\, R(x, y) \wedge (R(x, y) \vee \neg P(y, z))$ ×

CBL is incomparable with other FOL fragments (FOL[GF], FOL[UN]).

Main Properties of CBL

Results for CBL

1. Finite-model property.
2. Decidable satisfiability (Σ_3^P-COMPLETE).

Satisfiability Criterion

$\varphi \in$ CBL is sat *iff* there is an implicant where all atoms over the same relation and unifying bindings agree on the polarity.

$\varphi = \exists x \forall y \,.\, R(x, y) \wedge \neg R(y, x)$ is sat *iff* $\forall y \,.\, R(c_x, y) \wedge \neg R(y, c_x)$ is sat:

- (c_x, y) and (y, c_x) unify $\Rightarrow$ φ is unsat.

$\varphi = \forall y \exists x \,.\, R(x, y) \wedge \neg R(y, x)$ is sat *iff* $\forall y \,.\, R(f_x(y), y) \wedge \neg R(y, f_x(y))$ is sat:

- $(f_x(y), y)$ and $(y, f_x(y))$ do not unify $\Rightarrow$ φ is sat.

Herbrand Property (I)

Two terms $s(x_1, \ldots, x_n)$, $t(x_1, \ldots, x_n)$ are *equalizable* on a structure $\mathcal{A}$ if

$$\mathcal{A} \models \exists x_1, \ldots, x_n \,.\, s(x_1, \ldots, x_n) = t(x_1, \ldots, x_n)$$

Herbrand Property of $\mathcal{A}$: two terms *equalize* over $\mathcal{A}$ *iff* they *unify*.

Herbrand Property (II)

Herbrand Property of $\mathcal{A}$: two terms *equalize* over $\mathcal{A}$ *iff* they *unify*.

The Satisfiability Criterion for CBL is based on the Herbrand Property:

- two unifying bindings do equalize (*i.e.*, may assume the same value) on all models;
- there is a (finite) model on which all non-unifying bindings do not equalize.

- Structures satisfying HP are called Quasi-Herbrand Models.
- Standard Herbrand Models satisfy HP.

Conclusion

Summing up

1. Introduction and study of a new decidable fragment of FOL.
2. Identification of a model-theoretic property, namely the Herbrand Property, useful to prove the decidability of fragments of first order logic with functions.
3. Solution of the open problem about QCQ (finite) containment.
4. Discover of another reason why $\text{ATL}^\star$ and $\text{SL}[1\text{G}]$ are decidable.

Thank you!

References I

Andréka, H., van Benthem, J., and Németi, I. (1998).
Modal Languages And Bounded Fragments Of Predicate Logic.
JPL, 27(3):217–274.

Börger, E., Grädel, E., and Gurevich, Y. (1997).
The Classical Decision Problem.
Perspectives in Mathematical Logic. Springer.

Mogavero, F., Murano, A., Perelli, G., and Vardi, M. (2012).
What Makes ATL* Decidable? A Decidable Fragment of Strategy Logic.
In *CONCUR'12*, LNCS 7454, pages 193–208. Springer.

Mortimer, M. (1975).
On Languages with Two Variables.
MLQ, 21(1):135–140.

Schewe, S. (2008).
ATL* Satisfiability is 2ExpTime-Complete.
In *ICALP'08*, LNCS 5126, pages 373–385. Springer.

References II

ten Cate, B. and Segoufin, L. (2011).
Unary Negation.
In *STACS'11*, LIPIcs 9, pages 344–355. Leibniz-Zentrum fuer Informatik.

Vardi, M. (1996).
Why is Modal Logic So Robustly Decidable?
In *DCFM'96*, pages 149–184. AMS.