



# Conservative Extensions in Expressive Ontology Languages

Thomas Schneider

AG Theorie der künstlichen Intelligenz (TdKI), Universität Bremen

Joint work with

Jean Christoph Jung, Carsten Lutz, Mauricio Martel, Frank Wolter



# What is a conservative extension?

**Fundamental notion in mathematical logic** to relate theories

Useful in **computer science**:

- for formalising modularity in **software specification**
- for composing subgoals in **higher-order theorem proving**
- for formalising various notions in **ontology engineering**

Remarkably **positive results** for **description logics** and **modal logics**:

- ✓ turn out to be **decidable in many relevant cases**
- ✓ often have natural and insightful **model-theoretic characterisations**

**Natural question:** How far do these extend?

To  $FO^2$ ? Guarded fragment? Existential rules (aka Datalog $^\pm$ )?



# Conservative extensions for ontology design



## W3C Web Ontology Language **OWL**

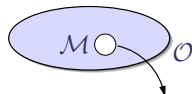
- ... is based on an expressive description logic (*SROIQ*)
- ... admits the design of **ontologies**,  
e.g., SNOMED CT, NCI Thesaurus, FMA (**100,000s of logical axioms**)
- Standard reasoning problems (e.g., SAT) well-understood reasoners: Racer, FaCT++, Pellet, Hermit, Konclude, ...

## **Challenges** for designing/using large ontologies:

- ▶ Navigation, comprehension
- ▶ Efficient (incremental) reasoning
- ▶ Efficient reuse **Efficient reuse**
- ▶ Versioning and more ...



# A reuse scenario



Assume that ...

- you want to buy **a subset of** a medical ontology  $\mathcal{O}$  from me that covers the subdomain of, say, diseases
- I offer two subsets  $\mathcal{M}_1$  and  $\mathcal{M}_2$

**Q:** which one do you choose?

**A:** the one that “knows more” about diseases!

**Q:** which is the **best subset** I can offer?

**A:** a subset  $\mathcal{M} \subseteq \mathcal{O}$  that is

- ... **indistinguishable** from  $\mathcal{O}$  w.r.t. all terms relevant for diseases
  - ... as small as possible
- signature  $\Sigma$



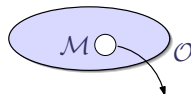
# Inseparability, $\Sigma$ -entailment, conservativity

Let  $\varphi_1, \varphi_2$  be sentences and  $\Sigma$  a **signature** (set of symbols).

- $\varphi_1$   **$\Sigma$ -entails**  $\varphi_2$ , written  $\varphi_1 \models_{\Sigma} \varphi_2$ ,  
if  $\varphi_2 \models \psi$  and  $\text{sig}(\psi) \subseteq \Sigma$  implies  $\varphi_1 \models \psi$ .
- $\varphi_1$  and  $\varphi_2$  are  **$\Sigma$ -inseparable**, written  $\varphi_1 \equiv_{\Sigma} \varphi_2$ ,  
if  $\varphi_1 \models_{\Sigma} \varphi_2$  and  $\varphi_2 \models_{\Sigma} \varphi_1$ .
- $\varphi_1 \wedge \varphi_2$  is a **conservative extension** of  $\varphi_1$   
if  $\varphi_1 \equiv_{\Sigma} \varphi_1 \wedge \varphi_2$  for  $\Sigma = \text{sig}(\varphi_1)$ .

In the reuse scenario, you should want to ...

- buy some  $\mathcal{M} \subseteq \mathcal{O}$  with  $\mathcal{M} \equiv_{\Sigma} \mathcal{O}$  ( $\mathcal{M} \models_{\Sigma} \mathcal{O}$  suffices)



# A closer look at $\Sigma$ -entailment

$\Sigma$ -entailment is the most general notion of the previous 3.

Two variants:

- **Deductive:**  $\varphi_1$   $\Sigma$ -entails  $\varphi_2$   
if  $\varphi_2 \models \psi$  and  $\text{sig}(\psi) \subseteq \Sigma$  implies  $\varphi_1 \models \psi$ .
- **Model-theoretic:**  $\varphi_1$   $\Sigma$ -entails  $\varphi_2$   
if every model of  $\varphi_1$  can be extended to a model of  $\varphi_2$  without changing the interpretation of the  $\Sigma$ -symbols.

Model-theoretic  $\Sigma$ -entailment is **highly undecidable** already for a **very small FO fragment**, that is, the description logic  $\mathcal{EL}$ :

$\forall x\varphi(x)$  with  $\varphi(x)$  built from  $\text{true}$ ,  $\wedge$ ,  $\exists y(Rxy \wedge \varphi(y))$

and when  $\Sigma = \text{sig}(\varphi_1)$ .

[Konev et al., AIJ 2013]



# Examples

- $\varphi_1$   $\Sigma$ -entails  $\varphi_2$   
if  $\varphi_2 \models \psi$  and  $\text{sig}(\psi) \subseteq \Sigma$  implies  $\varphi_1 \models \psi$ .

Examples in the guarded fragment GF of FO (with equality):

$$\varphi ::= x = y \mid R\bar{x} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \exists\bar{y}(R\bar{x}\bar{y} \wedge \varphi(\bar{x}\bar{y}))$$

- $\varphi_1 = \forall x\exists yRxy \not\models_{\{R\}} \varphi_2 = \forall x((\exists yRxy \wedge Ay) \wedge (\exists yRxy \wedge \neg Ay))$   
witnessed by  $\psi = \exists x\exists y(Rxy \wedge x \neq y)$
- $\varphi'_1 = \forall x\exists y(Rxy \wedge x \neq y) \models_{\{R\}} \varphi_2$

Choice of **separating logic** essential:  $\varphi'_1 \not\models_{\{R\}} \varphi_2$  in FO



# Overview



© Arnold Paul, Wikimedia, CC BY-SA 3.0

- 1 An overview of  $\Sigma$ -entailment
- 2  $\Sigma$ -entailment in FO fragments
- 3 Query entailment in expressive Horn description logics
- 4 Outlook





# Next ...

- 1 An overview of  $\Sigma$ -entailment
- 2  $\Sigma$ -entailment in FO fragments
- 3 Query entailment in expressive Horn description logics
- 4 Outlook



# $\Sigma$ -entailment in modal and description logics

Basic description logic  $\mathcal{ALC}$  ( $\approx$  multi-modal logic K):

$\forall x\varphi(x)$  with  $\varphi(x)$  built from true,  $\neg$ ,  $\wedge$ ,  $\exists y(Rxy \wedge \varphi(y))$

Theorem (Lutz & Wolter 2011, “sloppy version”)

$\varphi_1 \models_{\Sigma} \varphi_2$  iff every model of  $\varphi_1$  can be extended to a model of  $\varphi_2$ , up to  **$\Sigma$ -bisimulation**.

( $\hat{=}$  model-theoretic  $\Sigma$ -entailment up to “what the logic can express”)

Enables **decision procedure** using (amorphous) alternating tree automata

Problem is **2ExpTime-complete** (SAT in  $\mathcal{ALC}$ : ExpTime);  
smallest witnesses are triple exponential in the worst case!

Similar characterisations for **many other description and modal logics**



# $\Sigma$ -entailment beyond $\mathcal{ALC}$

- ✓ **Decidable and 2ExpTime-complete, too:** extensions of  $\mathcal{ALC}$  with
- inverse roles (aka past modalities)  $\exists y (Ryx \wedge \varphi(y))$
  - counting (aka graded modalities)

✗ **Undecidable:**

combination of the above two with nominals  $x = c$

[Lutz et al., IJCAI 2007]



# $\Sigma$ -entailment in lightweight description logics



Description logic  $\mathcal{EL}$  (“half  $\mathcal{ALC}$ ”):

$\forall x\varphi(x)$  with  $\varphi(x)$  built from `true`,  $\wedge$ ,  $\exists y(Rxy \wedge \varphi(y))$

- analogous model-theoretic characterisation via **simulations** (“half-bisimulations”)
- **ExpTime-complete** (SAT: in PTime)
- Restriction to acyclic terminologies: in **PTime** (deductive **and** model-theoretic variant)

[Lutz et al., KR 2012; Konev et al., JAIR 2012 & AIJ 2013]



# Further variants of $\Sigma$ -entailment

## $\Sigma$ -query entailment / $\Sigma$ -query inseparability:

separating formulas  $\psi$  are **queries**

Relevant for ontology-based data access (OBDA) aka ontology-mediated querying (OMQ)

Various notions of  $\Sigma$ -query entailment obtained by ...

- varying the query language: CQs, UCQs, PEQs, C2RPQs, ...
- allowing whether  $\varphi_1, \varphi_2$  contain data or not (if not, then relative to all possible instances)

Actively studied for basic and lightweight DLs (we'll get back later)

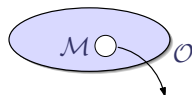


# Relationship with modularity of ontologies

## Back to the reuse scenario:

Given ontology  $\mathcal{O}$  and signature  $\Sigma$ ,  
you want to buy a subset  $\mathcal{M} \subseteq \mathcal{O}$  such that

- (a)  $\mathcal{M} \equiv_{\Sigma} \mathcal{O}$  and
- (b)  $\mathcal{M}$  small (possibly minimal with (a)): a **module of  $\mathcal{O}$  for  $\Sigma$**



Previous results: it is **hard to decide** whether a given  $\mathcal{M} \subseteq \mathcal{O}$  is a module

There are **tractable approximations** guaranteeing (a) but not (b),  
e.g., locality-based modules [Cuenca Grau et al., JAIR 2008]



# Relationship with uniform $\Sigma$ -interpolants

Let  $\varphi$  be a formula and  $\Sigma \subseteq \text{sig}(\varphi)$ .

Formula  $\psi$  is a **uniform  $\Sigma$ -interpolant** of  $\varphi$  if

- 1  $\text{sig}(\psi) \subseteq \Sigma$ ,
- 2  $\psi \equiv_{\Sigma} \varphi$

Relevant for **forgetting**:

eliminate non- $\Sigma$  predicates while preserving  $\Sigma$ -consequences

Applications:

- Ontology reuse
- Predicate hiding
- Ontology summary



# Next ...

- 1 An overview of  $\Sigma$ -entailment
- 2  $\Sigma$ -entailment in FO fragments**
- 3 Query entailment in expressive Horn description logics
- 4 Outlook





# Guarded Fragment and $FO^2$

$GF^k$ ,  $FO^k$ :  $k$ -variable fragment of GF or FO

## Theorem

$\Sigma$ -entailment,  $\Sigma$ -inseparability, and conservative extensions are

- 1 **undecidable** in every logic that **contains**  $GF^3$  or  $FO^2$  (such as the guarded negation fragment **GNF**);
- 2 **2ExpTime-complete** in  $GF^2$ .

(2) is based on a model-theoretic characterisation, but it is **much more complex** than, e.g., for  $\mathcal{ALC}$ .



# Undecidability

## GF<sup>3</sup>

- Reduction from the halting problem of **2-register machines**
- **Crucial:**  
 $\varphi_2$  uses ternary guard **that is not in  $\Sigma$** , thus breaks guardedness
- Little expressive power needed for **separation**: *ALC* suffices

## FO<sup>2</sup>

- Reduction from tiling problem
- Little expressive power needed in  $\varphi_1$  and  $\varphi_2$ : *ALC* suffices
- **Crucial:**  
ability to use **full FO<sup>2</sup> expressive power** in witnessing formula



# Characterisation

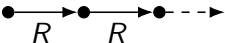
Recall that in  $\mathcal{ALC}$ :

Theorem (Lutz & Wolter 2011, “sloppy version”)

$\varphi_1 \models_{\Sigma} \varphi_2$  iff every model of  $\varphi_1$  can be extended to a model of  $\varphi_2$ , up to  $\Sigma$ -bisimulation.

**Q:** Can't we simply replace bisimulations with  $GF^2$ -bisimulations?

**A:** No!  $\varphi_1 = \exists x Ax \wedge \forall x (Ax \rightarrow \exists y (Rxy \wedge Ay))$      $\Sigma = \{R\}$

“There exists a path ”

$\varphi_2 = \varphi_1 \wedge \exists x (Ax \wedge Bx) \wedge \forall x (Bx \rightarrow \exists y (Ryx \wedge By))$

“There exists a path ”

Then  $\varphi_1 \models_{\Sigma} \varphi_2$ , but  $\Sigma$ - $GF^2$ -bisimulations fail.



# Bounded bisimulations

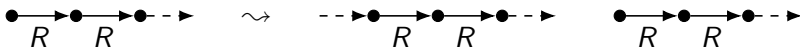
**$k$ -bounded bisimulations:**

$a \sim_{\Sigma}^k b$  iff  $a$  and  $b$  are  $\Sigma$ -GF<sup>2</sup>-bisimilar **up to depth  $k$**

**Theorem**

$\varphi_1 \models_{\Sigma} \varphi_2$  iff for every model  $\mathfrak{A}$  of  $\varphi_1$  of finite outdegree **and every  $k \geq 0$** , there is a model  $\mathfrak{B}$  of  $\varphi_2$  such that:

- ① for every  $a \in A$ , there is  $b \in B$  with  $a \sim_{\Sigma} b$  and
- ② for every  $b \in B$ , there is  $a \in A$  with  $a \sim_{\Sigma}^k b$ .



But it is **not so easy** to deal with **bounded bisimulations** when using tree automata or related techniques!



# “Marker-delimited” bisimulations

**Substitute** for  $k$ -bounded bisimulation between  $\mathfrak{A}$  and  $\mathfrak{B}$ :

- Decorate  $\mathfrak{A}$  with **unary predicate  $X$**  such that
  - on every infinite path there are infinitely many  $X$
  - the distance between two  $X$  is  $\geq k$
- Break off bisimulations at **second  $X$  seen** (both back and forth)

Does **not** travel **exactly  $k$  steps**, but we need it for **every  $k$**  anyway

**Problem:** even in **forest models**, decoration **does not exist** when  $k > 2$

**Solution:** we need **bounded bisimulations only when travelling upwards**, but not when travelling downwards.

Then the distance between markers **only matters on upwards paths**



# Automata-friendly characterisation

## Theorem

$\varphi_1 \models_{\Sigma} \varphi_2$  iff for every **forest** model  $\mathfrak{A}$  of  $\varphi_1$  of finite outdegree **and every marking**  $X \subseteq A$ , there is a model  $\mathfrak{B}$  of  $\varphi_2$  such that:

- ① for every  $a \in A$ , there is  $b \in B$  with  $a \sim_{\Sigma} b$  and
- ② for every  $b \in B$ , there is  $a \in A$  with  $a \sim_{\Sigma}^X b$ .

|  
downwards: unbounded

upwards: stop after seeing second  $X$



# Decision procedure based on automata

**2ATAs:** 2-way alternating tree automata  $\mathcal{A}$

- **Input:** non-empty node-labelled tree  
(unlimited depth, unbounded finite outdegree)
- **Transitions:**  
 $\wedge, \vee$ -formulas with atoms “send copy of  $\mathcal{A}$  in state  $q$  to ...”
  - current node
  - the predecessor node (if exists)
  - some or all successor node(s)

## Theorem

Emptiness for 2ATAs can be solved in time exponential in  $|Q|$ .

**Proof** via reduction to 2ATAs on (exactly)  $k$ -ary trees;  
their emptiness problem: ExpTime-complete [Vardi, ICALP 1998]



# Decision procedure based on automata

Construct 2ATA  $\mathcal{A}$  such that  $L(\mathcal{A}) = \emptyset$  iff  $\varphi_1 \models_{\Sigma} \varphi_2$

with  $|Q|$  polynomial in  $|\varphi_1|$  and exponential in  $|\varphi_2|$

**Input:** labelled tree representing a forest structure;  
labels contain information on unary/binary predicates and markers  $X$

$\mathcal{A}$  consists of three 2ATAs that check whether ...

- the input tree represents a **model**  $\mathfrak{A}$  of  $\varphi_1$
- the **marking** is correct
- there is  $\mathfrak{B} \models \varphi_2$  satisfying the **conditions from the theorem**:
  - ① for every  $a \in A$ , there is  $b \in B$  with  $a \sim_{\Sigma} b$  and
  - ② for every  $b \in B$ , there is  $a \in A$  with  $a \sim_{\Sigma}^X b$ .

( $\mathfrak{B}$  is constructed “locally”, memorising only guarded 1-/2-types)





# Harvest



## Theorem

In  $GF^2$ ,  $\varphi_1 \models_{\Sigma} \varphi_2$  can be decided in time single exponential in  $|\varphi_1|$  and double exponential in  $|\varphi_2|$ . The problem is 2ExpTime-complete.

(2ExpTime lower bound via ATM reduction)

## Corollary:

The same holds for  $\Sigma$ -inseparability, conservative extensions, and recognising uniform  $\Sigma$ -interpolants.



# Next ...

- 1 An overview of  $\Sigma$ -entailment
- 2  $\Sigma$ -entailment in FO fragments
- 3 Query entailment in expressive Horn description logics**
- 4 Outlook



# Ontology-mediated querying

## Idea:

- Database stores data – **ABox**  $\mathcal{A}$ , set of ground facts
- Ontology stores domain knowledge – **TBox**  $\mathcal{T}$ , think  $\forall x\varphi(x)$
- Queries  $q(\bar{x})$  are answered over **knowledge base (KB)**  $(\mathcal{T}, \mathcal{A})$

## Standard reasoning task query answering:

given  $(\mathcal{T}, \mathcal{A})$ ,  $q(\bar{x})$ ,  $\bar{a}$ , does  $(\mathcal{T}, \mathcal{A}) \models q(\bar{a})$  hold?

Query answering is well-understood . . .

- for lightweight and “full Boolean” description logics
- and query languages CQs, UCQs, and sometimes PEQs, C2RPQs



# $\Sigma$ -query entailment between TBoxes

As before, relevant for module extraction, versioning, etc.:

Let  $\mathcal{T}_1, \mathcal{T}_2$  be TBoxes,  $\Gamma, \Sigma$  signatures, and  $\mathcal{Q}$  a query language.

$\mathcal{T}_1$  ( $\Gamma, \Sigma$ )-query entails  $\mathcal{T}_2$ , written  $\mathcal{T}_1 \models_{\Gamma, \Sigma}^{\mathcal{Q}} \mathcal{T}_2$ ,

if for all  $\Gamma$ -ABoxes  $\mathcal{A}$ ,  $\Sigma$ -queries  $q(\bar{x}) \in \mathcal{Q}$  and tuples  $\bar{a}$ :

$$(\mathcal{T}_2, \mathcal{A}) \models q(\bar{a}) \quad \text{implies} \quad (\mathcal{T}_1, \mathcal{A}) \models q(\bar{a})$$

Inseparability and conservative extensions are again special cases.

**Variant  $\Sigma$ -query entailment between KBs:**  $(\mathcal{T}_1, \mathcal{A}_1) \models_{\Gamma, \Sigma}^{\mathcal{Q}} (\mathcal{T}_2, \mathcal{A}_2)$



# An overview of $\Sigma$ -query entailment

Analogous model-theoretic characterisations exist for the **KB** variant

- in *ALC*: via **homomorphisms** between tree-shaped models
- in *EL* and Horn-*ALC*: homomorphisms btn. **canonical models**

Useful for the **TBox** variant only if witness ABoxes can be restricted

Theorem (Botoeva et al., IJCAI'16)

$\Sigma$ -CQ entailment is **undecidable** for *ALC* TBoxes and **2ExpTime-complete** for *ALC* KBs.



# An overview of $\Sigma$ -query entailment

In  $\mathcal{EL}$  and Horn- $\mathcal{ALC}$ , the characterisations can be used to show

Theorem (Lutz & Wolter, JSC 2010; Botoeva et al., AIJ 2016)

$\Sigma$ -CQ entailment is ...

- ExpTime-complete for  $\mathcal{EL}$ -TBoxes and in PTime for  $\mathcal{EL}$ -KBs
- 2ExpTime-complete for Horn- $\mathcal{ALC}$  TBoxes and ExpTime-complete for Horn- $\mathcal{ALC}$  KBs

Further results for ...

- DL-Lite dialects
- rooted (U)CQs
- data complexity of the KB variant
- $\mathcal{T}_1, \mathcal{T}_2$  of different expressivity (e.g.,  $\mathcal{ALC}$  and  $\mathcal{EL}$ )

↪ Recent survey [Botoeva et al., RW 2016]



# Beyond $\mathcal{EL}$ and Horn- $\mathcal{ALC}$

**Goal:** study  $\Sigma$ -query entailment for Horn-DLs with

( $\mathcal{I}$ ) **inverse roles:**  $\exists y(Ryx \wedge Ay)$

- and more features:

( $\mathcal{F}$ ) **functionality:**  $\forall x_1 x_2 y (Rx_1 y \wedge Rx_2 y \rightarrow x_1 = x_2)$

( $\mathcal{H}$ ) **role hierarchies:**  $\forall xy (Rxy \rightarrow Sxy)$

and establish

- model-theoretic characterisations
- decidability/complexity



# Model-theoretic characterisation

In  $\mathcal{EL}$ :

Theorem (Lutz & Wolter 2010, “sloppy version”)

$\mathcal{T}_1 \models_{\Gamma, \Sigma}^{\text{CQ}} \mathcal{T}_2$  iff for all  $\Gamma$ -ABoxes  $\mathcal{A}$  there is a  **$\Sigma$ -homomorphism** from the the **canonical model** of  $(\mathcal{T}_2, \mathcal{A})$  to that of  $(\mathcal{T}_1, \mathcal{A})$ .

**Fails** in the presence of inverse roles for **very similar reasons**:  
infinite backward path not embeddable into infinite forward path!





# Bounded homomorphisms

**$k$ -bounded homomorphisms:**

$\mathcal{I}_1 \rightarrow_{\Sigma}^k \mathcal{I}_2$  iff  $\mathcal{I}_1$  embeds homomorphically into  $\mathcal{I}_2$  **up to depth  $k$**

Shorthand for the canonical model of  $(\mathcal{T}, \mathcal{A})$ :  $\mathcal{I}_{\mathcal{T}, \mathcal{A}}$

Characterisation for Horn-DLs **with inverse roles**:

Theorem

$\mathcal{T}_1 \models_{\Gamma, \Sigma}^{\text{CQ}} \mathcal{T}_2$  iff for all  $\Gamma$ -ABoxes  $\mathcal{A}$  **and all  $k \geq 0$** :

$$\mathcal{I}_{\mathcal{T}_2, \mathcal{A}} \rightarrow_{\Sigma}^k \mathcal{I}_{\mathcal{T}_1, \mathcal{A}}$$

Again, bounded homomorphisms are **difficult** for tree automata!



# Automata-friendly characterisation

**Lemma (sloppy):** it suffices to consider **tree-shaped** ABoxes and CQs.

## Theorem

$\mathcal{T}_1 \models_{\Gamma, \Sigma}^{\text{CQ}} \mathcal{T}_2$  iff for all **tree-shaped**  $\Gamma$ -ABoxes  $\mathcal{A}$ :

- (1) “the  $\Sigma$ -connected part of  $\mathcal{I}_{\mathcal{T}_2, \mathcal{A}}$ ”  $\rightarrow_{\Sigma} \mathcal{I}_{\mathcal{T}_1, \mathcal{A}}$  and
- (2) For every  $\Sigma$ -subtree  $\mathcal{I}$  in  $\mathcal{I}_{\mathcal{T}_2, \mathcal{A}}$ , one of the following holds:
  - (a)  $\mathcal{I} \rightarrow_{\Sigma} \mathcal{I}_{\mathcal{T}_1, \mathcal{A}}$
  - (b) there is a  $\Sigma$ -subtree of  $\mathcal{I}_2$  rooted in the ABox part such that  $\forall k \geq 0$ , we have  $\mathcal{I} \rightarrow_{\Sigma}^k \mathcal{I}'$

- (1) and (2a) use **unbounded** homomorphisms  
 $\rightsquigarrow$  decide via **2ATAs with counting**
- (2b) uses **bounded** homomorphisms  
 $\rightsquigarrow$  decide via **mosaic procedure**; “hard-code” into automaton



# Harvest



## Theorem

In  $\mathcal{ELI} \dots \text{Horn-ALCHIF}$ ,  $\mathcal{T}_1 \models_{\Gamma, \Sigma}^{\text{CQ}} \mathcal{T}_2$  can be decided in time single exponential in  $|\mathcal{T}_1|$  and double exponential in  $|\mathcal{T}_2|$ .  
The problem is 2ExpTime-complete.

## Corollary:

The same holds for  $\Sigma$ -inseparability and conservative extensions.



# Next ...

- 1 An overview of  $\Sigma$ -entailment
- 2  $\Sigma$ -entailment in FO fragments
- 3 Query entailment in expressive Horn description logics
- 4 Outlook**



# Outlook

## Deductive $\Sigma$ -entailment in FO fragments:

What happens if we add

- guarded counting quantifiers,
- transitive relations or equivalence relations,
- fixed points?

Finite-model version of conservative extensions?

## $\Sigma$ -query entailment:

Extension to Datalog<sup>±</sup> languages (aka existential rules),  
in particular to frontier-guarded TGDs?



# Thank you.



## An example

$\mathcal{T}_1$  (in FO notation):  $\forall x \left( \text{PhdStud}(x) \rightarrow \exists y \left( \text{adv}(y, x) \wedge \text{Prof}(y) \right) \right)$

“every PhD student is advised by some prof”

$\mathcal{T}_2 = \mathcal{T}_1 \wedge \forall x_1 x_2 y \left( \text{adv}(x_1, y) \wedge \text{adv}(x_2, y) \rightarrow x_1 = x_2 \right)$

“everyone has  $\leq 1$  advisor”

$\Gamma = \{\text{PhdStud}, \text{adv}\}, \quad \Sigma = \{\text{Prof}\}$

$\mathcal{T}_1 \not\models_{\Gamma, \Sigma}^{\text{CQ}} \mathcal{T}_2$ :

- $\Gamma$ -ABox  $\mathcal{A} = \{\text{PhdStud}(\text{john}), \text{adv}(\text{mary}, \text{john})\}$
- $\Sigma$ -CQ  $q(a) = \text{Prof}(\text{mary})$
- $(\mathcal{T}_2, \mathcal{A}) \models \text{Prof}(\text{mary})$  but  $(\mathcal{T}_1, \mathcal{A}) \not\models \text{Prof}(\text{mary})$



# Uniform interpolation

**Remember:**  $\psi$  is a **uniform  $\Sigma$ -interpolant** of  $\varphi$  if

- 1  $\text{sig}(\psi) \subseteq \Sigma$ ,
- 2  $\psi \equiv_{\Sigma} \varphi$

**Uniform interpolant recognition problem (UIRP):**

Given  $\varphi, \psi, \Sigma$ , is  $\psi$  a uniform  $\Sigma$ -interpolant of  $\varphi$ ?

**Easy reductions:** to deciding  $\Sigma$ -entailment,  
“backwards” from deciding conservative extensions

## Corollary

UIRP is **2ExpTime-complete** in  $\text{GF}^2$  and **undecidable** in all extensions of  $\text{FO}^2$  or  $\text{GF}^3$  with Craig interpolation, e.g., **GNF**.

There is **no decidable extension** of  $\text{FO}^2$  and of  $\text{GF}^3$  that has **effective uniform interpolation**.

