### Description Logic Knowledge Base Exchange

Elena Botoeva supervisor: Diego Calvanese



Fakultät für Informatik Facoltà di Scienze e Tecnologie informatiche Faculty of Computer Science

PhD Final Examination April 10, 2014 Bolzano

# Outline

Introduction

### **2** Summary of Work

### 3 Results

 Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations

# Outline

### Introduction

2 Summary of Work

#### **3** Results

Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations





Viewed as a knowledge base:





The website after restructuring:



Elena Botoeva(FUB)





# Data Exchange [Fagin et al., 2003]





source instance

# Data Exchange [Fagin et al., 2003]



# Data Exchange [Fagin et al., 2003]



Mapping  $\mathcal{M}$  is a set of inclusions of conjunctive queries (CQs)

$$\forall x, y \ (q_1(x, y) \rightarrow \exists z \ q_2(x, z)).$$

 $I_1$  is a complete database instance.

 $I_2$  is an incomplete database instance.

$$\mathcal{M}: \quad \forall a, t. (\quad AuthorOf(a, t) \quad \rightarrow \quad \exists g. BookInfo(t, a, g) \quad )$$

<i>I</i> <sub>1</sub> :	AuthorOf		
	nabokov	lolita	
	tolkien	lotr	

$$\mathcal{M}: \forall a, t. ($$
 AuthorOf $(a, t) \rightarrow \exists g. BookInfo(t, a, g) )$ 

<i>I</i> <sub>1</sub> :	AuthorOf		<i>I</i> <sub>2</sub> :	BookInfo		nfo
	nabokov	lolita		lolita	nabokov	tragicomedy
	tolkien	lotr		lotr	tolkien	fantasy

#### $I_2$ is a solution for $I_1$ under $\mathcal{M}$ .

$$\mathcal{M}: \forall a, t. ($$
 AuthorOf $(a, t) \rightarrow \exists g. BookInfo(t, a, g) )$ 

<i>I</i> <sub>1</sub> :	AuthorOf		
	nabokov	lolita	
	tolkien	lotr	

<b>l</b> <sub>2</sub> :	BookInfo		
	lolita	nabokov	tragicomedy
	lotr	tolkien	fantasy

<i>I</i> <sub>2</sub> :	BookInfo		
	lolita	nabokov	$null_1$
	lotr	tolkien	$null_2$

 $I_2$  is a solution for  $I_1$  under  $\mathcal{M}$ .  $I'_2$  is a universal solution for  $I_1$  under  $\mathcal{M}$ .

$$\mathcal{M}: \forall a, t. ($$
 AuthorOf $(a, t) \rightarrow \exists g. BookInfo(t, a, g) )$ 

<i>I</i> <sub>1</sub> :	AuthorOf		
	nabokov	lolita	
	tolkien	lotr	

<b>l</b> <sub>2</sub> :	BookInfo		
	lolita	nabokov	tragicomedy
	lotr	tolkien	fantasy

<b>I</b> '_2 :	BookInfo		
	lolita	nabokov	$null_1$
	lotr	tolkien	null <sub>2</sub>

 $I_2$  is a solution for  $I_1$  under  $\mathcal{M}$ .  $I'_2$  is a universal solution for  $I_1$  under  $\mathcal{M}$ .

 $\Rightarrow$  there is a homomorphism from  $I'_2$  to  $I_2$ .

## Incomplete Data and Knowledge Exchange

A framework for Data Exchange with incomplete data was proposed by Arenas et al. [2011].



## Incomplete Data and Knowledge Exchange

A framework for Data Exchange with incomplete data was proposed by Arenas et al. [2011].



# Incomplete Data and Knowledge Exchange

A framework for Data Exchange with incomplete data was proposed by Arenas et al. [2011].



We specialize this framework to Description Logics, and in particular to DL-Lite<sub>R</sub>.

Elena Botoeva(FUB)

### Description Logic *DL-Lite*<sub>R</sub>

*Description Logics (DLs)* are decidable fragments of First-Order Logic, used as Knowledge Representation formalisms.

DLs talk about the domain of interest by means of

- concepts (unary predicates): Author, Book, A, B
- roles (binary predicates): AuthorOf, WrittenBy, P, R

DL-Lite<sub>R</sub> is a light-weight DL that asserts

- concept inclusions and disjointness of atomic concepts A, the domain ∃P and the range ∃P<sup>-</sup> of atomic roles P, Book ⊑ ∃WrittenBy,
- role inclusions and disjointness of atomic roles P and their inverses P<sup>-</sup>, AuthorOf ⊆ WrittenBy<sup>-</sup>,
- ground facts Author(nabokov), AuthorOf(nabokov,lolita), A(a), P(a, b).

### Description Logic *DL-Lite*<sub>R</sub>

Description Logics (DLs) are decidable fragments of First-Order Logic, used as Knowledge Representation formalisms.

DLs talk about the domain of interest by means of

- concepts (unary predicates): Author, Book, A, B
- roles (binary predicates): AuthorOf, WrittenBy, P, R

DL-Lite<sub>R</sub> is a light-weight DL that asserts

- L-Lite<sub>R</sub> is a light-weight DL that associate
  concept inclusions and disjointness of atomic concepts A, the domain ∃P and the range ∃P<sup>-</sup> of atomic roles P, Book ⊑ ∃WrittenBy,
  role inclusions and disjointness of atomic roles P and the role inclusions and disjointness of atomic roles P and the role atomic roles P and the role writtenBy<sup>-</sup>,

  - ground facts Author(nabokov), AuthorOf(nabokov,lolita), A(a), P(a, b).

## Description Logic *DL-Lite*<sub>R</sub>

Description Logics (DLs) are decidable fragments of First-Order Logic, used as Knowledge Representation formalisms.

DLs talk about the domain of interest by means of

- concepts (unary predicates): Author, Book, A, B
- roles (binary predicates): AuthorOf, WrittenBy, P, R

DL-Lite<sub>R</sub> is a light-weight DL that asserts

- Concept inclusions and disjointness of atomic concepts A, the domain ∃P and the range ∃P<sup>-</sup> of atomic roles P, Book ⊆ ∃WrittenBy,
  role inclusions and disjointness of atomic roles P and D<sup>-</sup> Author Of ⊆ WrittenBy<sup>-</sup>

  - ground facts Author(nabokov), AuthorOf(nabokov,lolita), A(a), P(a, b).

Satisfiability check over a *DL-Lite*<sub> $\mathcal{R}$ </sub> KB  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  can be done in polynomial time (in fact, in NLOGSPACE).

# Outline

Introduction

### 2 Summary of Work

#### **3** Results

Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations The Summary of my PhD Work

In this thesis, we

- Propose a general framework for exchanging Description Logic Knowledge Bases.
- Define and analyse relevant reasoning problems in this setting.
- Develop reasoning techniques and characterize the computational complexity of the problems.

# Outline

Introduction

2 Summary of Work

### 3 Results

Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations

## 1. Knowledge Base Exchange Framework



# 1. Knowledge Base Exchange Framework



# 2. Reasoning Problems

# 2. Reasoning Problems

#### universal solution

preserves all models

### universal UCQ-solution

preserves all answers to Unions of Conjunctive Queries

#### UCQ-representation

preserves all answers to UCQs, independently of the ABox

Elena Botoeva(FUB)

Solution

# 2. Reasoning Problems

universal solution

preserves all models

Single A.S. Restered on an is been nut

### universal UCQ-solution

preserves all answers to Unions of Conjunctive Queries

#### UCQ-representation

preserves all answers to UCQs, independently of the ABox

Elena Botoeva(FUB)

Solution



**Universal solutions** Membership Non-emptiness simple ABoxes PTIME-complete PTIME-complete extended ABoxes NP-complete PSPACE-hard, in ExpTIME

### **Universal** UCQ**-solutions** Membership Non-emptiness

simple ABoxes PSPACE-hard in EXPTIME extended ABoxes in EXPTIME PSPACE-hard

UCQ-**representations** Membership Non-emptiness Complexity NLOGSPACE-complete NLOGSPACE-complete

Elena Botoeva(FUB)

Universal solutions	simple ABoxes	extended ABoxes	
Membership	PTIME-complete	NP-complete	
Non-emptiness	PTIME-complete	PSPACE-hard, in EXPTIME	
games	ad-hoc	e automata	
Universal UCQ-solutions	simple ABoxes	extended ABoxes	
Membership	PSPACE-hard	in EXPTIME	
Non-emptiness	in EXPTIME	PSPACE-hard	

UCQ-**representations** Membership Non-emptiness Complexity NLOGSPACE-complete NLOGSPACE-complete

Elena Botoeva(FUB)

**Universal solutions** Membership Non-emptiness simple ABoxes PTIME-complete PTIME-complete extended ABoxes NP-complete PSPACE-hard, in ExpTIME



**Universal solutions** Membership Non-emptiness simple ABoxes PTIME-complete PTIME-complete extended ABoxes NP-complete PSPACE-hard, in ExpTIME

### **Universal** UCQ**-solutions** Membership Non-emptiness

simple ABoxes PSPACE-hard in EXPTIME extended ABoxes in ExpTIME PSPACE-hard

### UCQ-**representations** Membership Non-emptiness

Complexity NLOGSPACE-complete NLOGSPACE-complete





# Outline

Introduction

2 Summary of Work

**3** Results



Universal Solutions Universal UCQ-solutions UCQ-representations

### The Essence of Knowledge Base Exchange

A mapping is a triple  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is a set of  $DL\text{-}Lite_{\mathcal{R}}$  inclusions from  $\Sigma_1$  to  $\Sigma_2$ 

 $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  is a *DL-Lite*<sub> $\mathcal{R}$ </sub> knowledge base over  $\Sigma_1$  (source KB)  $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a *DL-Lite*<sub> $\mathcal{R}$ </sub> knowledge base over  $\Sigma_2$  (target KB)


## The Essence of Knowledge Base Exchange

A mapping is a triple  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is a set of DL-Lite<sub> $\mathcal{R}$ </sub> inclusions from  $\Sigma_1$  to  $\Sigma_2$  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  is a DL-Lite<sub> $\mathcal{R}$ </sub> knowledge base over  $\Sigma_1$  (source KB)  $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a DL-Lite<sub> $\mathcal{R}$ </sub> knowledge base over  $\Sigma_2$  (target KB)



For a KB  $\mathcal{K}$ , we denote by  $\mathcal{U}_{\mathcal{K}}$  the canonical model of  $\mathcal{K}$  (chase in databases).

- $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff\*  $\mathcal{T}_2 = \emptyset$  and  $\mathcal{U}_{\mathcal{A}_2}$  is  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .
- $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .
- $\mathcal{T}_2$  is a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*\*</sup> for each source ABox  $\mathcal{A}_1$ ,  $\mathcal{U}_{\langle \mathcal{T}_2 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  is  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

- $\star$  plus one more condition with little technical impact
- \*\* plus one more condition for checking consistency

Elena Botoeva(FUB)

## The Essence of Knowledge Base Exchange

A mapping is a triple  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is a set of  $DL\text{-Lite}_{\mathcal{R}}$  inclusions from  $\Sigma_1$  to  $\Sigma_2$  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  is a  $DL\text{-Lite}_{\mathcal{R}}$  knowledge base over  $\Sigma_1$  (source KB)  $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a  $DL\text{-Lite}_{\mathcal{R}}$  knowledge base over  $\Sigma_2$  (target KB)



For a KB  $\mathcal{K}$ , we denote by  $\mathcal{U}_{\mathcal{K}}$  the canonical model of  $\mathcal{K}$  (chase in databases).

- $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff\*  $\mathcal{T}_2 = \emptyset$  and  $\mathcal{U}_{\mathcal{A}_2}$  is  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .
- $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .
- $\mathcal{T}_2$  is a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*\*</sup> for each source ABox  $\mathcal{A}_1$ ,  $\mathcal{U}_{\langle \mathcal{T}_2 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  is  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

#### We present our **6** techniques that check the existence of the homomorphisms.

- \* plus one more condition with little technical impact
- $\star\star$  plus one more condition for checking consistency

Elena Botoeva(FUB)

Let  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ , where













# Outline

Introduction

2 Summary of Work

**3** Results

 Technical Development Universal Solutions Universal UCQ-solution UCQ-representations

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ ,
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ .

<sup>\*</sup> plus one more condition of no technical interest

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ .

\* plus one more condition of no technical interest

Elena Botoeva(FUB)

Description Logic Knowledge Base Exchange

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ . via Reachability Games on graphs

<sup>\*</sup> plus one more condition of no technical interest

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ . via Reachability Games on graphs

For a KB  $\mathcal{K}$ , an ABox  $\mathcal{A}$ , and a signature  $\Sigma$ ,  $\checkmark$ we construct a reachability game (G, F) such that

there exists a  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}}$  to  $\mathcal{U}_{\mathcal{A}}$ 

iff

 $\begin{array}{|c|c|}\hline \textbf{Duplicator} & \textbf{has a strategy against} & \textbf{Spoiler} \\ \hline \textbf{to avoid } F & \textbf{in G.} \end{array}$ 

<sup>\*</sup> plus one more condition of no technical interest

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ . via Reachability Games on graphs

For a KB  $\mathcal{K}$ , an ABox  $\mathcal{A}$ , and a signature  $\Sigma$ ,  $\checkmark$ we construct a reachability game (G, F) such that

there exists a  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}}$  to  $\mathcal{U}_{\mathcal{A}}$ 

iff

 $\begin{array}{|c|c|}\hline \textbf{Duplicator} & \textbf{has a strategy against} & \textbf{Spoiler} \\ \hline \textbf{to avoid } F & \textbf{in G.} \end{array}$ 



\* plus one more condition of no technical interest

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ . via Reachability Games on graphs

For a KB  $\mathcal{K}$ , an ABox  $\mathcal{A}$ , and a signature  $\Sigma$ ,  $\checkmark$ we construct a reachability game (G, F) such that

there exists a  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}}$  to  $\mathcal{U}_{\mathcal{A}}$ 

iff

 $\begin{array}{|c|c|}\hline \textbf{Duplicator} & \textbf{has a strategy against} & \textbf{Spoiler} \\ \hline \textbf{to avoid } F & \textbf{in G.} \end{array}$ 



\* plus one more condition of no technical interest

 $\mathcal{A}_2$  is a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup> there exist

- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\mathcal{A}_2}$  to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ , EASY
- a  $\Sigma_2$ -homomorphism from  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  to  $\mathcal{U}_{\mathcal{A}_2}$ . via Reachability Games on graphs

For a KB  $\mathcal{K}$ , an ABox  $\mathcal{A}$ , and a signature  $\Sigma$ ,  $\checkmark$ we construct a reachability game (G, F) such that

there exists a  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}}$  to  $\mathcal{U}_{\mathcal{A}}$ 

iff



\* plus one more condition of no technical interest









 $\mathcal{U}_{\mathcal{K}}$  is  $\Sigma\text{-homomorphically embeddable into }\mathcal{U}_{\mathcal{A}}.$ 



Duplicator has a strategy against Spoiler to avoid F in G from  $a \mapsto a$  iff  $\mathcal{U}_{\mathcal{K}}$  is  $\Sigma$ -homomorphically embeddable into  $\mathcal{U}_{\mathcal{A}}$ .



 $\fbox{\begin{tabular}{|c|c|c|c|} \hline \begin{tabular}{c} Duplicator \\ \hline \begin{tabular}{c} Duplicator \\ \hline \begin{tabular}{c} A & A \\ \hline \begin{tabular}{c} U_{\mathcal{K}} \\ \hline \begin{tabular}{c} S & A \\ \hline \begin{tabular}{c} A & A \\ \hline$ 



















Duplicator has a strategy against Spoiler to avoid F in G from  $a \mapsto a$  iff  $\mathcal{U}_{\mathcal{K}}$  is  $\Sigma$ -homomorphically embeddable into  $\mathcal{U}_{\mathcal{A}}$ .





#### Elena Botoeva(FUB)

#### Non-emptiness for Extended Universal Solutions is in $\operatorname{ExpTIME}$

There exists a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup>  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  is  $\Sigma_2$ -homomorphically embeddable to a finite subset of itself.

#### Non-emptiness for Extended Universal Solutions is in $\operatorname{ExpTIME}$

There exists a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup>  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  is  $\Sigma_2$ -homomorphically embeddable to a finite subset of itself.

For KBs  $\mathcal{K}_1$ ,  $\mathcal{K}_2$ , and a signature  $\Sigma$ , we construct a two-way alternating automaton (2ATA)  $\mathbb{A}$  that accepts a tree encoding a finite subset  $\mathcal{C}$  of  $\mathcal{U}_{\mathcal{K}_2}$  such that

 $\mathcal{U}_{\mathcal{K}_1}$  is  $\Sigma$ -homomorphically embeddable into  $\mathcal{C}$ .

#### Non-emptiness for Extended Universal Solutions is in $\operatorname{ExpTIME}$

There exists a universal solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff<sup>\*</sup>  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$  is  $\Sigma_2$ -homomorphically embeddable to a finite subset of itself.

For KBs  $\mathcal{K}_1$ ,  $\mathcal{K}_2$ , and a signature  $\Sigma$ , we construct a two-way alternating automaton (2ATA)  $\mathbb{A}$  that accepts a tree encoding a finite subset  $\mathcal{C}$  of  $\mathcal{U}_{\mathcal{K}_2}$  such that

 $\mathcal{U}_{\mathcal{K}_1}$  is  $\Sigma$ -homomorphically embeddable into  $\mathcal{C}$ .







There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff


There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff

#### there is a $\Sigma$ -homomorphism from $\mathcal{U}_{\mathcal{K}_1}$ to $\mathcal{C}$ ? q<sub>h</sub> P'h $\mathbb{R}$ P'P'awo bw<sub>P</sub> ā P'P'**bw**<sub>P</sub>w<sub>P</sub> Wo A run of $\mathbb{A}$ over $T_{\mathcal{C}}$ $\mathbb{S}^{1\cdot 1\cdot 1}$ $\mathcal{U}_{\mathcal{K}}^{\Sigma}$ $C^{\Sigma}$ $T_{\mathcal{C}}$

**2** 2ATA  $\mathbb{A}$  to check  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}_1}$  to a finite subset  $\mathcal{C}$  of  $\mathcal{U}_{\mathcal{K}_2}$ 

There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff

**2** 2ATA  $\mathbb{A}$  to check  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}_1}$  to a finite subset  $\mathcal{C}$  of  $\mathcal{U}_{\mathcal{K}_2}$ 



There exists an accepting run of the 2ATA  $\mathbb{A}$  over  $T_{\mathcal{C}}$  iff

**2** 2ATA  $\mathbb{A}$  to check  $\Sigma$ -homomorphism from  $\mathcal{U}_{\mathcal{K}_1}$  to a finite subset  $\mathcal{C}$  of  $\mathcal{U}_{\mathcal{K}_2}$ 





#### decidable in exponential time

Elena Botoeva(FUB)

# Outline

Introduction

2 Summary of Work

**3** Results

 Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations

## Membership for Universal UCQ-Solutions is in $\operatorname{ExpTIME}$

 $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

Membership for Universal UCQ-Solutions is in  $\operatorname{ExpTIME}$ 

 $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

For KBs  $\mathcal{K}_1$ ,  $\mathcal{K}_2$ , and a signature  $\Sigma$ , we use reachability games for checking whether  $\mathcal{U}_{\mathcal{K}_1}$  is finitely  $\Sigma$ -homomorphically embeddable into  $\mathcal{U}_{\mathcal{K}_2}$ .

Membership for Universal UCQ-Solutions is in  $\operatorname{ExpTIME}$ 

 $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

For KBs  $\mathcal{K}_1$ ,  $\mathcal{K}_2$ , and a signature  $\Sigma$ , we use reachability games for checking whether  $\mathcal{U}_{\mathcal{K}_1}$  is finitely  $\Sigma$ -homomorphically embeddable into  $\mathcal{U}_{\mathcal{K}_2}$ .

Now,  $\mathcal{U}_{\mathcal{K}_2}$  is in general infinite.

- we cannot use the game  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}_1}, \mathcal{U}_{\mathcal{K}_2}) = (G_i, F_i)$ , a straightforward extension of  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}}, \mathcal{U}_{\mathcal{A}})$ , as  $G_i$  is in general infinite.
- so we define a game  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}_1}, \mathcal{G}_{\mathcal{K}_2}) = (G_f, F_f)$ , where  $G_f$  is of exponential size and the states have a more complicated structure involving

 $\{u_1,\ldots,u_k\}\mapsto w.$ 

 $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$  is a universal UCQ-solution for  $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$  iff  $\mathcal{U}_{\langle \mathcal{T}_2, \mathcal{A}_2 \rangle}$  is finitely  $\Sigma_2$ -homomorphically equivalent to  $\mathcal{U}_{\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle}$ .

For KBs  $\mathcal{K}_1$ ,  $\mathcal{K}_2$ , and a signature  $\Sigma$ , we use reachability games for checking whether  $\mathcal{U}_{\mathcal{K}_1}$  is finitely  $\Sigma$ -homomorphically embeddable into  $\mathcal{U}_{\mathcal{K}_2}$ .

Now,  $\mathcal{U}_{\mathcal{K}_2}$  is in general infinite.

- we cannot use the game  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}_1}, \mathcal{U}_{\mathcal{K}_2}) = (G_i, F_i)$ , a straightforward extension of  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}}, \mathcal{U}_{\mathcal{A}})$ , as  $G_i$  is in general infinite.
- so we define a game  $G_{\Sigma}(\mathcal{G}_{\mathcal{K}_1}, \mathcal{G}_{\mathcal{K}_2}) = (G_f, F_f)$ , where  $G_f$  is of exponential size and the states have a more complicated structure involving

 $\{u_1,\ldots,u_k\}\mapsto w.$ 

Hence, we obtain an EXPTIME upper bound.

# Outline

Introduction

2 Summary of Work

**3** Results

 Technical Development Universal Solutions Universal UCQ-solutions UCQ-representations





such that for each ABox  $\mathcal{A}_1$  and each query q



such that for each ABox  $\mathcal{A}_1$  and each query  $\textbf{\textit{q}}$ 



Elena Botoeva(FUB)











such that for each ABox  $\mathcal{A}_1$  and each query q



 $\begin{array}{l} \text{Membership for UCQ-representations is in NLOGSPACE} \\ \mathcal{T}_2 \text{ is a UCQ-representation of } \mathcal{T}_1 \text{ under } \mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12}) \text{ iff:} \end{array}$ 

Let  $\mathcal{A}_1 = \{X(a)\}.$ 









 $\begin{array}{l} \text{Membership for UCQ-representations is in NLOGSPACE} \\ \mathcal{T}_2 \text{ is a UCQ-representation of } \mathcal{T}_1 \text{ under } \mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12}) \text{ iff:} \end{array}$ 



Let  $\mathcal{A}_1 = \{X(a), Y(a)\}$ 

3)  $\langle \mathcal{T}_1 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle$  is inconsistent iff  $\langle \mathcal{T}_2 \cup \mathcal{T}_{12}, \mathcal{A}_1 \rangle$  is inconsistent

Elena Botoeva(FUB)

Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}$ , and  $\mathcal{T}_1 = \{A \sqsubseteq B\}$ .

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):

Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}, \text{ and } \mathcal{T}_1 = \{A \sqsubseteq B\}.$ 

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):



We have that  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . So  $\mathcal{T}_2$  should be such that  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models A \sqsubset B'$ . Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}, \text{ and } \mathcal{T}_1 = \{A \sqsubseteq B\}.$ 

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):



We have that  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . So  $\mathcal{T}_2$  should be such that  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ .
Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}, \text{ and } \mathcal{T}_1 = \{A \sqsubseteq B\}.$ 

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):



We have that  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . So  $\mathcal{T}_2$  should be such that  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}, \text{ and } \mathcal{T}_1 = \{A \sqsubseteq B\}.$ 

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):



We have that  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . So  $\mathcal{T}_2$  should be such that  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ .

If  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models C \sqsubseteq B'$ , then it should be the case  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models C \sqsubseteq B'$ .

Non-emptiness for UCQ-representations is in NLOGSPACE

Let  $\Sigma_1 = \{A(\cdot), B(\cdot), C(\cdot)\}, \Sigma_2 = \{A'(\cdot), B'(\cdot), C'(\cdot)\}, \text{ and } \mathcal{T}_1 = \{A \sqsubseteq B\}.$ 

Is there  $\mathcal{T}_2$ , a UCQ-representation for  $\mathcal{T}_1$  under  $\mathcal{M} = (\Sigma_1, \Sigma_2, \mathcal{T}_{12})$ , where  $\mathcal{T}_{12}$  is as follows (gray arrows):



We have that  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ . So  $\mathcal{T}_2$  should be such that  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models A \sqsubseteq B'$ .

If  $\mathcal{T}_2 \cup \mathcal{T}_{12} \models C \sqsubseteq B'$ , then it should be the case  $\mathcal{T}_1 \cup \mathcal{T}_{12} \models C \sqsubseteq B'$ .

#### Publications

#### **Conference Publications**

- E. Botoeva, R. Kontchakov, V. Ryzhikov, F. Wolter, and M. Zakharyaschev. Query inseparability for description logic knowledge bases. In Proc. of the 14th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2014). To appear.
- M. Arenas, E. Botoeva, D. Calvanese, and V. Ryzhikov. Exchanging OWL 2 QL knowledge bases.
  In Proc. of the 23rd Int. Joint Conf. on Artificial Intelligence (IJCAI 2013), pages 703-710, 2013.
- M. Arenas, E. Botoeva, D. Calvanese, V. Ryzhikov, and E. Sherkhonov. Exchanging description logic knowledge bases.
  In Proc. of the 13th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2012), pages 563-567.

#### Workshop Publications

- M. Arenas, E. Botoeva, D. Calvanese, and V. Ryzhikov. Computing solutions in OWL 2 QL knowledge exchange. In Proc. of the 26th Int. Workshop on Description Logic (DL 2013), volume 1014, pages 4-16, 2013.
- M. Arenas, E. Botoeva, D. Calvanese, V. Ryzhikov, and E. Sherkhonov. Representability in DL-Liter knowledge base exchange. In Proc. of the 25th Int. Workshop on Description Logic (DL 2012), volume 846, 2012.
- M. Arenas, E. Botoeva, and D. Calvanese. Knowledge base exchange.
  In Proc. of the 24th Int. Workshop on Description Logic (DL 2011), volume 745, 2011.

#### Publications cont.

#### **Technical Reports**

 M. Arenas, E. Botoeva, D. Calvanese, and V. Ryzhikov. Exchanging OWL 2 QL knowledge bases (extended version). CoRR Technical Report arXiv:1304.5810, arXiv.org e-Print archive, 2013. Available at http://arxiv.org/abs/1304.5810.

#### Under Submission

 M. Arenas, E. Botoeva, D. Calvanese, and V. Ryzhikov. Knowledge base exchange: The case of OWL 2 QL. Under submission to a journal.

# Thank you for your attention!

Universal solutions Membership Non-emptiness simple ABoxes PTIME-complete PTIME-complete extended ABoxes NP-complete PSPACE-hard, in EXPTIME

Universal UCQ-solutions Membership Non-emptiness simple ABoxes PSPACE-hard in EXPTIME

extended ABoxes in EXPTIME PSPACE-hard

UCQ-**representations** Membership Non-emptiness Complexity NLOGSPACE-complete NLOGSPACE-complete

Elena Botoeva(FUB)

Description Logic Knowledge Base Exchange

Marcelo Arenas, Jorge Pérez, and Juan L. Reutter. Data exchange beyond complete data. pages 83–94, 2011.

Ronald Fagin, Phokion G. Kolaitis, Renée J. Miller, and Lucian Popa. Data exchange: Semantics and query answering. pages 207–224, 2003.

Elena Botoeva(FUB)

Description Logic Knowledge Base Exchange

# Knowledge Base Exchange: Example



$$\mathcal{T}_{1}: \begin{array}{c} \exists Author Of \sqsubseteq Author \\ Author \sqsubseteq \exists Ta \times Number \end{array}$$



# Knowledge Base Exchange: Example



#### $\mathcal{A}_2$ is a universal solution for $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$ under $\mathcal{M}$ (with extended ABoxes).

Elena Botoeva(FUB)

Description Logic Knowledge Base Exchange

# Knowledge Base Exchange: Example



$$\mathcal{T}_{1}: \begin{array}{c} \exists AuthorOf \sqsubseteq Author\\ Author \sqsubseteq \exists TaxNumber \end{array}$$

 $\mathcal{T}_{2}: \quad \begin{array}{l} \exists Written By^{-} \sqsubseteq \exists SSN \\ \exists Written By \sqsubseteq \exists BookGenre \end{array}$ 



#### $\langle \mathcal{T}_2, \mathcal{A}_2 \rangle$ is a universal-UCQ solution for $\langle \mathcal{T}_1, \mathcal{A}_1 \rangle$ under $\mathcal{M}$ (with simple ABoxes).

Elena Botoeva(FUB)

Description Logic Knowledge Base Exchange











































