

# Inconsistency-Tolerant Querying of Description Logic Knowledge Bases

M2 KNOWLEDGE REPRESENTATION

---

Meghyn Bienvenu (LaBRI - CNRS & *Université de Bordeaux*)

# HANDLING DATA INCONSISTENCIES

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

**Standard semantics:** everything is implied - **not informative!**

- when  $\mathcal{K}$  unsatisfiable,  $\text{cert}(q, \mathcal{K})$  contains all possible tuples

# HANDLING DATA INCONSISTENCIES

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

**Standard semantics:** everything is implied - **not informative!**

- when  $\mathcal{K}$  unsatisfiable,  $\text{cert}(q, \mathcal{K})$  contains all possible tuples

Two approaches to inconsistency handling:

- **resolve the inconsistencies**
  - preferable, but not always applicable!
- live with the inconsistencies - **adopt alternative semantics**
  - **meaningful answers** to queries despite inconsistencies

Note: focus on case where errors in ABox (**assume TBox reliable**)

# EXAMPLE: REASONABLE INFERENCES

TBox  $\mathcal{T}_{\text{univ}}$ :

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

Consider following ABoxes:

$\mathcal{A}_1$	=	$\{\text{Prof}(\text{anna}), \text{Lect}(\text{anna}), \text{Fellow}(\text{alex})\}$
$\mathcal{A}_2$	=	$\{\text{Prof}(\text{anna}), \text{Fellow}(\text{alex}), \text{Lect}(\text{alex})\}$

Which assertions would be reasonable to infer from these two KBs?

$\text{Prof}(\text{anna})$	$\text{Lect}(\text{anna})$	$\text{Fac}(\text{anna})$
$\text{Fellow}(\text{alex})$	$\text{Lect}(\text{alex})$	$\text{Fac}(\text{alex})$

# EXAMPLE: REASONABLE ANSWERS

TBox  $\mathcal{T}_{\text{univ}}$ :

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

ABox  $\mathcal{A}_{\text{univ}}$ :

Prof(anna), Lect(anna), Fellow(anna), Prof(kim), Lect(kim),  
Fellow(julie), teaches(csc343, julie), Fellow(alex), teaches(alex, csc486)

Question: what are **reasonable answers** for our example queries?

$q_1(x) = \text{Fac}(x)$	$q_2(x) = \exists y \text{teaches}(x, y)$
$q_3(x) = \exists y \text{Fac}(x) \wedge \text{teaches}(x, y)$	$q_4(x, y) = \text{Fac}(x) \wedge \text{teaches}(x, y)$

# ALTERNATIVE SEMANTICS

In general: **no single best way** to define answers for inconsistent KBs  
 $\Rightarrow$  consider **many different inconsistency-tolerant semantics**

Formally: a **semantics  $\mathcal{S}$**  associates a set of **query answers** to **every KB and query**

- if  $\mathcal{K}$  is **satisfiable**, should return **certain answers**
- for **unsatisfiable**  $\mathcal{K}$ , can give **different answers** than classical semantics

Write  $\mathcal{K} \models_{\mathcal{S}} q(\vec{a})$  if  **$\vec{a}$  answer to  $q$  w.r.t.  $\mathcal{K}$  under semantics  $\mathcal{S}$**   
(and use  $\mathcal{K} \models q(\vec{a})$  for certain answer semantics, i.e.  $\vec{a} \in \text{cert}(q, \mathcal{K})$ )

Consider different ways of **comparing semantics**

# CONSISTENCY PROPERTIES

An ABox  $\mathcal{A}$  is  $\mathcal{T}$ -consistent if the KB  $(\mathcal{T}, \mathcal{A})$  is satisfiable

Call  $\mathcal{C} \subseteq \mathcal{A}$  is a (consistent)  $\mathcal{T}$ -support of  $q(\vec{a})$  if:

(i)  $\mathcal{C}$  is  $\mathcal{T}$ -consistent      (ii)  $(\mathcal{T}, \mathcal{C}) \models q(\vec{a})$

Semantics  $S$  satisfies the **CONSISTENT SUPPORT property** if whenever  $\mathcal{K} \models_s q(\vec{a})$ , there exists a  $\mathcal{T}$ -support  $\mathcal{C} \subseteq \mathcal{A}$  of  $q(\vec{a})$

- important for explaining / justifying query results to users

# CONSISTENCY PROPERTIES

An ABox  $\mathcal{A}$  is  **$\mathcal{T}$ -consistent** if the KB  $(\mathcal{T}, \mathcal{A})$  is satisfiable

Call  $\mathcal{C} \subseteq \mathcal{A}$  is a **(consistent)  $\mathcal{T}$ -support** of  $q(\vec{a})$  if:

(i)  $\mathcal{C}$  is  **$\mathcal{T}$ -consistent**      (ii)  $(\mathcal{T}, \mathcal{C}) \models q(\vec{a})$

Semantics  $S$  satisfies the **CONSISTENT SUPPORT property** if whenever  $\mathcal{K} \models_S q(\vec{a})$ , there **exists a  $\mathcal{T}$ -support  $\mathcal{C} \subseteq \mathcal{A}$  of  $q(\vec{a})$**

- important for **explaining / justifying query results to users**

Semantics  $S$  satisfies the **CONSISTENT RESULTS property** if for every KB  $\mathcal{K}$ , there exists a model  $\mathcal{I}$  of  $\mathcal{T}$  such that  $\mathcal{K} \models_S q(\vec{a})$  implies  $\mathcal{I} \models q(\vec{a})$ .

- set of **query results** is **jointly consistent with TBox**
- **safe to combine query results**



# CONSISTENCY PROPERTIES

An ABox  $\mathcal{A}$  is  **$\mathcal{T}$ -consistent** if the KB  $(\mathcal{T}, \mathcal{A})$  is satisfiable

Call  $\mathcal{C} \subseteq \mathcal{A}$  is a **(consistent)  $\mathcal{T}$ -support** of  $q(\vec{a})$  if:

- (i)  $\mathcal{C}$  is  **$\mathcal{T}$ -consistent**      (ii)  $(\mathcal{T}, \mathcal{C}) \models q(\vec{a})$

Semantics  $S$  satisfies the **CONSISTENT SUPPORT property** if whenever  $\mathcal{K} \models_S q(\vec{a})$ , there **exists a  $\mathcal{T}$ -support  $\mathcal{C} \subseteq \mathcal{A}$  of  $q(\vec{a})$**

- important for **explaining / justifying query results to users**

Semantics  $S$  satisfies the **CONSISTENT RESULTS property** if for every KB  $\mathcal{K}$ , there exists a model  $\mathcal{I}$  of  $\mathcal{T}$  such that  $\mathcal{K} \models_S q(\vec{a})$  implies  $\mathcal{I} \models q(\vec{a})$ .

- set of **query results** is **jointly consistent with TBox**
- **safe to combine query results**

Note: **neither property implies the other**

# COMPARING DIFFERENT SEMANTICS

Given two semantics  $S$  and  $S'$ , we say that:

- $S'$  is an **under-approximation** (or: **sound approximation**) of  $S$  just in the case that

$$\mathcal{K} \models_{S'} q(\vec{a}) \Rightarrow \mathcal{K} \models_S q(\vec{a})$$

- $S'$  is an **over-approximation** (or: **complete approximation**) of  $S$  just in the case that

$$\mathcal{K} \models_S q(\vec{a}) \Rightarrow \mathcal{K} \models_{S'} q(\vec{a})$$

**Consistency properties** are **preserved by under-approximations**:

$S'$  is an **under-approximation** of  $S$  &  $S$  satisfies  $P \Rightarrow S'$  also satisfies  $P$

here  $P \in \{\text{CONSISTENT SUPPORT, CONSISTENT RESULTS}\}$

Many semantics are based upon the notion of repair

**Repair** of an ABox  $\mathcal{A}$  w.r.t. a TBox  $\mathcal{T}$   
= inclusion-maximal subset of  $\mathcal{A}$  that is  $\mathcal{T}$ -consistent

Intuition: different **ways of achieving consistency** while **retaining as much of the original data as possible**

Denote by  $Rep(\mathcal{A}, \mathcal{T})$  the **set of repairs of  $\mathcal{A}$  w.r.t.  $\mathcal{T}$**

- abbreviate to  $Rep(\mathcal{K})$  when  $\mathcal{K} = (\mathcal{T}, \mathcal{A})$

Every KB has **at least one repair**

- inconsistent KB  $\Rightarrow$  **typically multiple repairs**

# EXAMPLE: REPAIRS

Reconsider the TBox  $\mathcal{T}_{\text{univ}}$ :

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

and ABox  $\mathcal{A}_{\text{univ}}$ :

Prof(anna), Lect(anna), Fellow(anna), Prof(kim), Lect(kim),  
Fellow(julie), teaches(csc343, julie), Fellow(alex), teaches(alex, csc486)

Recall the minimal  $\mathcal{T}_{\text{univ}}$ -inconsistent subsets:

$\{\text{Prof(anna), Lect(anna)}\}$        $\{\text{Prof(anna), Fellow(anna)}\}$   
 $\{\text{Lect(anna), Fellow(anna)}\}$        $\{\text{Prof(kim), Lect(kim)}\}$   
 $\{\text{Fellow(julie), teaches(csc343, julie)}\}$

Question: **How many repairs of  $\mathcal{A}_{\text{univ}}$  w.r.t.  $\mathcal{T}_{\text{univ}}$ ?**

## EXAMPLE: REPAIRS (CONT.)

Twelve repairs of  $\mathcal{A}_{\text{univ}}$  w.r.t.  $\mathcal{T}_{\text{univ}}$ :

$$\begin{aligned}\mathcal{R}_1 &= \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_2 &= \{\text{Lect}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_3 &= \{\text{Fellow}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_4 &= \{\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_5 &= \{\text{Lect}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_6 &= \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_7 &= \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_8 &= \{\text{Lect}(\text{anna}), \text{Lect}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_9 &= \{\text{Fellow}(\text{anna}), \text{Prof}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{10} &= \{\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{11} &= \{\text{Lect}(\text{anna}), \text{Prof}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{12} &= \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}}\end{aligned}$$

where the ABox  $\mathcal{A}_{\text{Int}}$  that is common to all the repairs is as follows:

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{teaches}(\text{alex}, \text{csc486})\}$$

**Repair:**  $\subseteq$ -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

**Repair:**  $\subseteq$ -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

**Plausible answers:** hold no matter which repair is chosen

# PLAUSIBLE ANSWERS: AR SEMANTICS

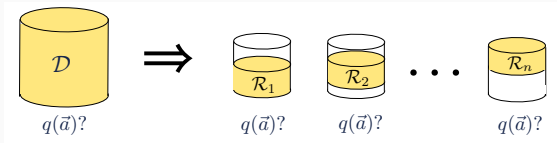
**Repair:**  $\subseteq$ -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

**Plausible answers:** hold no matter which repair is chosen

**AR semantics:** query each repair separately, intersect results

$$\mathcal{K} \models_{\text{AR}} q(\vec{a}) \Leftrightarrow (\mathcal{T}, \mathcal{B}) \models q(\vec{a}) \text{ for every repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$





## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$ , as every repair contains one of  $\text{Prof}(\text{anna})$ ,  $\text{Lect}(\text{anna})$ , and  $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$ , as every repair contains  $\text{Prof}(\text{kim})$  or  $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$ , as every repair contains  $\text{Fellow}(\text{alex})$

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$ , as every repair contains one of  $\text{Prof}(\text{anna})$ ,  $\text{Lect}(\text{anna})$ , and  $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$ , as every repair contains  $\text{Prof}(\text{kim})$  or  $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$ , as every repair contains  $\text{Fellow}(\text{alex})$

These are the **only answers under AR semantics**:

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$ , as every repair contains one of  $\text{Prof}(\text{anna})$ ,  $\text{Lect}(\text{anna})$ , and  $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$ , as every repair contains  $\text{Prof}(\text{kim})$  or  $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$ , as every repair contains  $\text{Fellow}(\text{alex})$

These are the **only answers under AR semantics**:

- $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{julie})$  as  $(\mathcal{T}_{\text{univ}}, \mathcal{R}_7) \not\models \text{Fac}(\text{julie})$
- can similarly show  $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{CSC486})$  and  $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{CSC343})$

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_2 = \exists y \text{teaches}(x, y)$ , we have:

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_2 = \exists y \text{teaches}(x, y)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{kim})$ , as every repair contains  $\text{Prof}(\text{kim})$  or  $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{alex})$ , as every repair contains  $\text{teaches}(\text{alex}, \text{csc486})$

## EXAMPLE: AR SEMANTICS

Reconsider our **example KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

For the query  $q_2 = \exists y \text{teaches}(x, y)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{kim})$ , as every repair contains  $\text{Prof}(\text{kim})$  or  $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{alex})$ , as every repair contains  $\text{teaches}(\text{alex}, \text{csc486})$

These are the **only answers under AR semantics**:

- $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{anna})$  as  $(\mathcal{T}_{\text{univ}}, \mathcal{R}_3) \not\models \exists y \text{teaches}(\text{anna}, y)$
- can similarly show **julie**, **csc486**, and **csc343** are **not answers**

# PLAUSIBLE ANSWERS: AR SEMANTICS

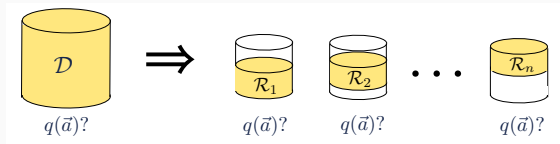
**Repair:**  $\subseteq$ -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

**Plausible answers:** hold no matter which repair is chosen

**AR semantics:** query each repair separately, intersect results

$$\mathcal{K} \models_{\text{AR}} q(\vec{a}) \iff (\mathcal{T}, \mathcal{B}) \models q(\vec{a}) \text{ for every repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$



Satisfies both **CONSISTENT SUPPORT** and **CONSISTENT RESULTS**



# SUREST ANSWERS: IAR SEMANTICS

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

# SUREST ANSWERS: IAR SEMANTICS

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

**IAR semantics:** query the **intersection of the repairs**

$$\mathcal{K} \models_{\text{IAR}} q(\vec{a}) \quad \Leftrightarrow \quad (\mathcal{T}, \mathcal{D}) \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \mathcal{B}$$

## EXAMPLE: IAR SEMANTICS

Reconsider our **example KB**  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

Intersection of the repairs of  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$ :

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{teaches}(\text{alex}, \text{csc486})\}$$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

# EXAMPLE: IAR SEMANTICS

Reconsider our **example KB**  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

Intersection of the repairs of  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$ :

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{teaches}(\text{alex}, \text{csc486})\}$$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

$$\cdot \mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex}), \text{ as } (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{Int}}) \models \text{Fac}(\text{alex})$$

# EXAMPLE: IAR SEMANTICS

Reconsider our **example KB**  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$

Intersection of the repairs of  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$ :

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{teaches}(\text{alex}, \text{csc486})\}$$

For the query  $q_1(x) = \text{Fac}(x)$ , we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$ , as  $(\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{Int}}) \models \text{Fac}(\text{alex})$

This is the **only answer to  $q_1$**  under **IAR semantics**:

- **anna** and **kim** are **no longer considered answers** since needed to **reason by cases** (e.g., kim is either Prof or Lect)

# SUREST ANSWERS: IAR SEMANTICS

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

**IAR semantics:** query the **intersection of the repairs**

$$\mathcal{K} \models_{\text{IAR}} q(\vec{a}) \iff (\mathcal{T}, \mathcal{D}) \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \mathcal{B}$$

**Under-approximation** of the **AR semantics**

**Satisfies** both **CONSISTENT SUPPORT** and **CONSISTENT RESULTS**

## POSSIBLE ANSWERS: BRAVE SEMANTICS

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

# POSSIBLE ANSWERS: BRAVE SEMANTICS

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

**Brave semantics:** query the repairs, take union of their answers

$$\mathcal{K} \models_{\text{brave}} q(\vec{a}) \iff (\mathcal{T}, \mathcal{B}) \models q(\vec{a}) \text{ for some repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$



## EXAMPLE: BRAVE SEMANTICS

Reconsider the **KB**  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$  and **query**  $q_1(x) = \text{Fac}(x)$  .

## EXAMPLE: BRAVE SEMANTICS

Reconsider the KB  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$  and query  $q_1(x) = \text{Fac}(x)$ .

Moving from AR to brave semantics yields an additional answer:

- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{anna})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{kim})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{alex})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{julie})$   $(\mathcal{T}_{\text{univ}}, \mathcal{R}_i) \models q_1(\text{julie})$  for  $1 \leq i \leq 6$

## EXAMPLE: BRAVE SEMANTICS

Reconsider the KB  $\mathcal{K}_{\text{univ}} = (\mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}})$  and query  $q_1(x) = \text{Fac}(x)$ .

Moving from AR to brave semantics yields an additional answer:

- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{anna})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{kim})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{alex})$  AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{julie})$   $(\mathcal{T}_{\text{univ}}, \mathcal{R}_i) \models q_1(\text{julie})$  for  $1 \leq i \leq 6$

These are the only answers to  $q_1$  under brave semantics:

- csc486 and csc343 cannot be obtained as answers from any repair

# POSSIBLE ANSWERS: BRAVE SEMANTICS

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

**Brave semantics**: query the repairs, take union of their answers

$$\mathcal{K} \models_{\text{brave}} q(\vec{a}) \Leftrightarrow (\mathcal{T}, \mathcal{B}) \models q(\vec{a}) \text{ for some repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$

**Over-approximation** of the **AR semantics**

- ... and **every semantics that satisfies CONSISTENT SUPPORT**

**Does not satisfy CONSISTENT RESULTS**

Why?

# SEMANTICS BASED UPON PREFERRED REPAIRS

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

# SEMANTICS BASED UPON PREFERRED REPAIRS

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

Formally: use **preference relation**  $\preceq$  to compare repairs

- compare w.r.t. **cardinality** ( $\leq$ )
- partition ABox into **priority levels**  $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ 
  - compare **level-by-level using set inclusion** ( $\subseteq_{\mathcal{P}}$ )
  - compare **level-by-level using cardinality** ( $\leq_{\mathcal{P}}$ )
- assign **weights to ABox assertions**
  - compare repairs by **total weight** ( $\leq_w$ )

# SEMANTICS BASED UPON PREFERRED REPAIRS

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

Formally: use **preference relation**  $\preceq$  to compare repairs

- compare w.r.t. **cardinality** ( $\leq$ )
- partition ABox into **priority levels**  $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ 
  - compare **level-by-level using set inclusion** ( $\subseteq_{\mathcal{P}}$ )
  - compare **level-by-level using cardinality** ( $\leq_{\mathcal{P}}$ )
- assign **weights to ABox assertions**
  - compare repairs by **total weight** ( $\leq_w$ )

**AR / IAR / brave semantics** based upon **most preferred repairs**  
( $\preceq$ -AR,  $\preceq$ -IAR,  $\preceq$ -brave)

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy:**



Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy:**

- every **minimal support for  $q(\vec{a})$**  contains **at most  $|q|$  assertions**

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy:**

- every **minimal support for  $q(\vec{a})$**  contains **at most  $|q|$  assertions**
- every **minimal  $\mathcal{T}$ -inconsistent subset** has **cardinality at most two**

# COMPLEXITY ANALYSIS FOR DL-LITE

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy:**

- every **minimal support for  $q(\vec{a})$**  contains **at most  $|q|$  assertions**
- every **minimal  $\mathcal{T}$ -inconsistent subset** has **cardinality at most two**
- **CQ answering**, instance checking, and KB consistency can be performed by **FO query rewriting** (so in  **$AC^0$  in data complexity**)

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy**:

- every **minimal support for  $q(\vec{a})$**  contains **at most  $|q|$  assertions**
- every **minimal  $\mathcal{T}$ -inconsistent subset** has **cardinality at most two**
- **CQ answering**, instance checking, and KB consistency can be performed by **FO query rewriting** (so in  $AC^0$  in data complexity)
- **CQ answering** is **NP-complete for combined complexity**

# COMPLEXITY ANALYSIS FOR DL-LITE

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite <sub>$\mathcal{R}$</sub>**  and **all DL-Lite dialects that satisfy**:

- every **minimal support for  $q(\vec{a})$**  contains **at most  $|q|$  assertions**
- every **minimal  $\mathcal{T}$ -inconsistent subset** has **cardinality at most two**
- **CQ answering**, instance checking, and KB consistency can be performed by **FO query rewriting** (so in  $AC^0$  in data complexity)
- **CQ answering** is **NP-complete for combined complexity**
- **instance checking** is **NL-complete in combined complexity**

# BAD NEWS: INTRACTABILITY OF AR SEMANTICS

Theorem CQ and IQ answering under AR semantics are coNP-complete in data complexity

# BAD NEWS: INTRACTABILITY OF AR SEMANTICS

Theorem CQ and IQ answering under AR semantics are coNP-complete in data complexity

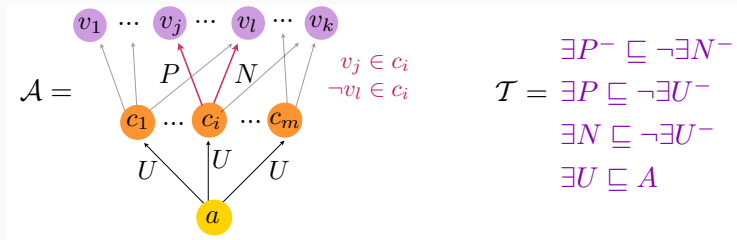
Upper bound: guess  $\mathcal{A}' \subseteq \mathcal{A}$ , verify  $\mathcal{A}'$  is repair and  $(\mathcal{T}, \mathcal{A}') \not\models q(\vec{a})$

# BAD NEWS: INTRACTABILITY OF AR SEMANTICS

Theorem **CQ and IQ answering** under **AR semantics** are **coNP-complete** in **data complexity**

Upper bound: **guess**  $\mathcal{A}' \subseteq \mathcal{A}$ , **verify**  $\mathcal{A}'$  is **repair** and  $(\mathcal{T}, \mathcal{A}') \not\models q(\vec{a})$

Lower bound: **reduction from UNSAT**  $\varphi = c_1 \wedge \dots \wedge c_m$  over  $v_1, \dots, v_k$



Can show  $\varphi$  **unsatisfiable**  $\Leftrightarrow \mathcal{T}, \mathcal{A} \models_{\text{AR}} A(a)$



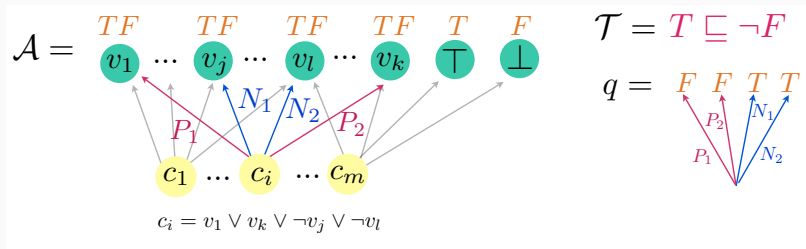
## BAD NEWS: INTRACTABILITY OF AR SEMANTICS (CONT.)

In fact: **CQ answering** is **coNP-hard** for simple TBox  $\mathcal{T} = \{T \sqsubseteq \neg F\}$

# BAD NEWS: INTRACTABILITY OF AR SEMANTICS (CONT.)

In fact: **CQ answering** is **coNP-hard** for simple TBox  $\mathcal{T} = \{T \sqsubseteq \neg F\}$

**Reduction from 2+2UNSAT:**  $\varphi = c_1 \wedge \dots \wedge c_m$  over  $v_1, \dots, v_k, \top, \perp$   
each clause has **two positive** and **two negative literals**



Can show  $\varphi$  **unsatisfiable**  $\Leftrightarrow \mathcal{T}, \mathcal{A} \models_{\text{AR}} q$

# GOOD NEWS: IAR AND BRAVE

For IAR and brave semantics,  
have **same low data complexity** as classical semantics

Theorem **CQ answering** under **IAR semantics** are  
in  $AC^0$  in **data complexity**

Theorem **CQ answering** under **brave semantics** are  
in  $AC^0$  in **data complexity**

# GOOD NEWS: IAR AND BRAVE

For IAR and brave semantics,  
have **same low data complexity** as classical semantics

Theorem **CQ answering** under **IAR semantics** are  
in  $AC^0$  in **data complexity**

Theorem **CQ answering** under **brave semantics** are  
in  $AC^0$  in **data complexity**

Can use **FO-query rewriting** to **compute IAR- and brave-answers**

## EXAMPLE: REWRITING FOR IAR SEMANTICS

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

## EXAMPLE: REWRITING FOR IAR SEMANTICS

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

(Normal) rewriting of  $q_2(x) = \exists y \text{teaches}(x, y)$  w.r.t.  $\mathcal{T}_{\text{univ}}$ :

$$q'_2(\mathbf{x}) = \text{Prof}(\mathbf{x}) \vee \text{Lect}(\mathbf{x}) \vee \exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y})$$

## EXAMPLE: REWRITING FOR IAR SEMANTICS

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

(Normal) rewriting of  $q_2(x) = \exists y \text{teaches}(x, y)$  w.r.t.  $\mathcal{T}_{\text{univ}}$ :

$$q'_2(\mathbf{x}) = \text{Prof}(\mathbf{x}) \vee \text{Lect}(\mathbf{x}) \vee \exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y})$$

Rewriting of  $q_2$  for IAR semantics:

$$\begin{aligned} q''_2(\mathbf{x}) = & \text{Prof}(\mathbf{x}) \wedge (\neg \text{Lect}(x) \wedge \neg \text{Fellow}(x) \wedge \neg \text{Course}(x) \wedge \neg \exists z.\text{teaches}(z, x)) \vee \\ & \text{Lect}(\mathbf{x}) \wedge (\neg \text{Prof}(x) \wedge \neg \text{Fellow}(x) \wedge \neg \text{Course}(x) \wedge \neg \exists z.\text{teaches}(z, x)) \vee \\ & \exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y}) \wedge (\neg \text{Prof}(y) \wedge \neg \text{Lect}(y) \wedge \neg \text{Fellow}(y)) \end{aligned}$$

## EXAMPLE: REWRITING FOR BRAVE SEMANTICS

Idea: modify UCQ-rewriting to ensure each disjunct can only match  $\mathcal{T}$ -consistent subset of ABox



## EXAMPLE: REWRITING FOR BRAVE SEMANTICS

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match  $\mathcal{T}$ -consistent subset of ABox**

Modified **TBox**  $\mathcal{T}'_{\text{univ}}$ : add  $\exists \text{teaches} \sqsubseteq \text{Fac}$  to  $\mathcal{T}_{\text{univ}}$

## EXAMPLE: REWRITING FOR BRAVE SEMANTICS

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match  $\mathcal{T}$ -consistent subset of ABox**

Modified **TBox**  $\mathcal{T}'_{\text{univ}}$ : add  $\exists \text{teaches} \sqsubseteq \text{Fac}$  to  $\mathcal{T}_{\text{univ}}$

**(Normal) rewriting** of  $q_2(x) = \exists y \text{teaches}(x, y)$  w.r.t.  $\mathcal{T}'_{\text{univ}}$ :

$$q'_2(\mathbf{x}) = \text{Prof}(\mathbf{x}) \vee \text{Lect}(\mathbf{x}) \vee \exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y})$$

## EXAMPLE: REWRITING FOR BRAVE SEMANTICS

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match  $\mathcal{T}$ -consistent subset of ABox**

Modified **TBox**  $\mathcal{T}'_{\text{univ}}$ : add  $\exists \text{teaches} \sqsubseteq \text{Fac}$  to  $\mathcal{T}_{\text{univ}}$

(Normal) rewriting of  $q_2(x) = \exists y \text{teaches}(x, y)$  w.r.t.  $\mathcal{T}'_{\text{univ}}$ :

$$q'_2(\mathbf{x}) = \text{Prof}(\mathbf{x}) \vee \text{Lect}(\mathbf{x}) \vee \exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y})$$

Rewriting of  $q_2$  for **brave semantics**:

$$q'_2(\mathbf{x}) = \text{Prof}(\mathbf{x}) \vee \text{Lect}(\mathbf{x}) \vee (\exists \mathbf{y}.\text{teaches}(\mathbf{x}, \mathbf{y}) \wedge \mathbf{x} \neq \mathbf{y})$$

to **disallow** using assertions of the form  $\text{teaches}(a, a)$

# COMPLEXITY LANDSCAPE FOR DL-LITE

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	in $AC^0$	in $AC^0$	NP	NL
AR	coNP	coNP	$\Pi_2^P$	coNP
IAR	in $AC^0$	in $AC^0$	NP	NL
brave	in $AC^0$	in $AC^0$	NP	NL

Note: IQs is for “instance queries”, aka instance checking

# TOWARDS PRACTICAL SYSTEMS FOR INCONSISTENCY HANDLING

CQAPri first system for AR query answering in DL-Lite

Implements **hybrid approach**:

- compute IAR and brave answers polytime
  - gives upper and lower **bounds on AR answers**
- use SAT solvers to **identify remaining AR answers**
- three categories of answers : **possible, likely**, (almost) **sure**

# TOWARDS PRACTICAL SYSTEMS FOR INCONSISTENCY HANDLING

CQAPri first system for AR query answering in DL-Lite

Implements **hybrid approach**:

- compute IAR and brave answers polytime
  - gives upper and lower **bounds on AR answers**
- use SAT solvers to **identify remaining AR answers**
- three categories of answers : **possible, likely**, (almost) **sure**

**Encouraging experimental results:**

- in most cases, IAR and brave enough to decide if tuple is AR-answer  $\Rightarrow$  **few calls to SAT solvers**
- **SAT encodings** are typically **small and easy to solve**

Lightweight DL  $\mathcal{EL}_{\perp}$ : constructors  $\top, \perp, \sqcap, \exists r.C$

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	P	P	NP	P
AR	coNP	coNP	$\Pi_2^P$	coNP
IAR	coNP	coNP	$\Delta_2^P[O(\log n)]$	coNP
brave	NP	NP	NP	NP

Observe: **IAR** and **brave** are **no longer tractable**

- **no bound** on size of **minimal  $\mathcal{T}$ -inconsistent subsets**

Expressive DL  $\mathcal{ALC}$ : constructors  $\top, \perp, \neg, \sqcap, \sqcup, \exists r.C, \forall r.C$

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	coNP	coNP	EXP	EXP
AR	$\Pi_2^P$	$\Pi_2^P$	EXP	EXP
IAR	$\Pi_2^P$	$\Pi_2^P$	EXP	EXP
brave	$\Sigma_2^P$	$\Sigma_2^P$	EXP	EXP

Observe:

- IAR and brave no easier than AR
- increased data complexity, no increase in combined complexity